

# **MOLECULAR ORIGIN OF MESODERM: CONSERVATION OF T-BOX GENES IN NON-BILATERIAN ANIMALS**

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie  
vorgelegt der  
Philosophisch-Naturwissenschaftlichen Fakultät  
der Universität Basel

von

**Cosimo Martinelli**

aus

**Italien**

Basel, 2003

Zoologisches Institut  
Universität Basel  
Biozentrum/Pharmazentrum  
Klingelbergstrasse 50/70  
CH-4056 Basel  
Switzerland

# TABLE OF CONTENTS

<b>SUMMARY</b> .....	3
<b>GENERAL INTRODUCTION</b> .....	4
An overview on basal metazoans .....	5
Molecular phylogeny and relationships among basal metazoans .....	21
T-box genes .....	25
<b>CHAPTER 1</b>	
<b>Distinct expression patterns of the two T-box homologues <i>Brachyury</i> and <i>Tbx2/3</i> in the placozoan <i>Trichoplax adhaerens</i></b>	
Abstract .....	32
Introduction .....	33
Materials and methods .....	36
Results and discussion .....	39
<b>CHAPTER 2</b>	
<b>T-box genes in basal metazoans: further indications for the monophyletic origin of metazoans</b>	
Abstract .....	45
Introduction .....	46
Materials and methods .....	48
Results and discussion .....	51
<b>CHAPTER 3</b>	
<b>Role of the homeobox gene <i>Not</i> in body remodelling of the basal metazoan <i>Trichoplax adhaerens</i></b>	
Abstract .....	58
Introduction .....	59
Materials and methods .....	61
Results and discussion .....	64
<b>References</b> .....	69
<b>Appendices</b> .....	78
<b>Acknowledgments</b> .....	97
<b>Curriculum vitae</b> .....	98

## SUMMARY

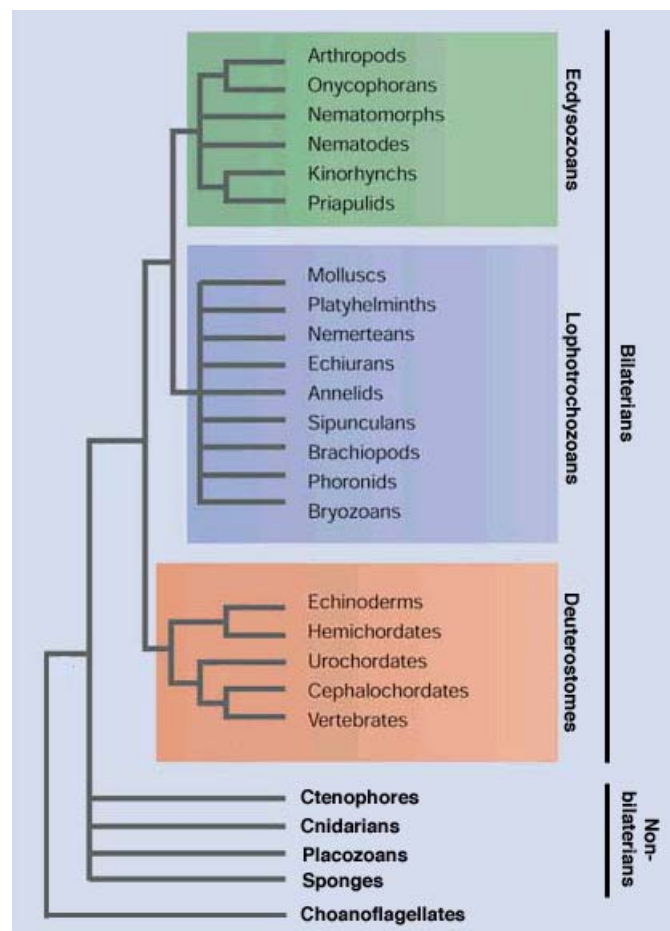
Several members of the T-box gene family were characterized from cnidarians, ctenophores, sponges and placozoans, which are considered as the four metazoan phyla at the basis of animal evolution leading to bilaterians. T-box genes code for transcription factors with high sequence similarity in a region of about 180 amino acids called T-box domain, which is involved in DNA binding. T-box genes have important roles in many developmental processes such as mesoderm formation, antero-posterior axis formation, muscle and limb formation. Thus, T-box genes have been considered key genes to study developmental processes linked to evolutionary aspects. In this study it is shown that the four extant non-bilaterian phyla have members of well-defined T-box gene subfamilies such as *Brachyury* or *Tbx2/3*. These are clear homologs of higher evolved bilaterian genes, suggesting the presence of a complex T-box gene family already in the hypothetical common ancestor of all metazoans (Urmetazoa).

The data collected in this study were also used to test the phylogenetic relationship of the four basal metazoan phyla in relation to bilaterian animals. Molecular phylogeny has redefined the evolutionary tree, introducing the ecdysozoa and lophotrochozoa as new clades of protostomes, but did not solve the relationship and the branching order of many phyla, especially of the four phyla at the base of animal evolution. In this study most of the investigations were done to clarify the position of the placozoan *Trichoplax adhaerens*, which is an enigmatic organism that has been interpreted either as a derived cnidarian or as a primitive non-bilaterian within its own phylum named Placozoa. The phylogenetic analyses reported here confirm that *Trichoplax* forms an own phylum, but does not really improve the knowledge about its phylogenetic position. An important improvement for the investigation of the biology of *Trichoplax* was achieved by adapting in situ hybridisation methods. Experiments with T-box genes such as *Brachyury* and *Tbx2/3*, homeobox genes such as *Not* and *Trox-2*, and additional reference genes suggest that *Trichoplax* has more than the four cell types described in the literature and that the process of regeneration could be regulated by a homeobox gene of the *Not* class.

# GENERAL INTRODUCTION

## AN OVERVIEW ON BASAL METAZOANS

Basal metazoans or non-bilaterians are terms often used to unite the four animal phyla at the base of the evolutionary tree, namely the Porifera, Placozoa, Ctenophora and Cnidaria. Sometimes also the term diploblast is used to compare these animals to triploblasts. According to most textbooks only triploblasts or bilaterians are formed by ectoderm, endoderm and mesoderm in between. However, independent description of mesoderm like structures exist for sponges, the placozoan *Trichoplax adhaerens*, ctenophores, and cnidarians. How this often morphological descriptions correspond to molecular correlations and could be compared to the position of non-bilaterians in relation to the better studied bilaterian phyla (Fig. 1) is discussed in the first part of this introduction.



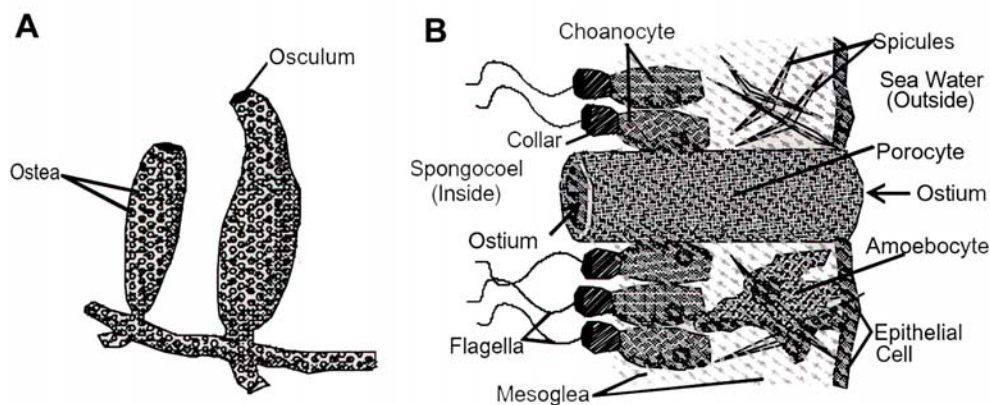
**Fig. 1.** Proposed phylogeny of the animal kingdom, based primarily on 18S rDNA sequences (Holland 1999). All animals can be divided into two basic groups according to their bauplan: the bilaterians and the non-bilaterians.

## PORIFERA: THE SIMPLEST ANIMAL ORGANIZATION

Sponges are among the simplest multicellular animals, although depending on the species from 10 to 15 cell types have been counted. They lack organs and tissues and all the cells exhibit considerable independence. The sponges are an ancient group with a fossil record extending back to the Precambrian (Li et al. 1998). Sponges have often been considered to be the result of an independent line of evolution, without much relationship to other metazoans, however, recent molecular phylogenetic studies are supporting a view where sponges share the same ancestor with all the others metazoans (Müller et al. 2001).

### The main characters of sponges

Sponges are characterized by numerous canals and chambers, which open to the outside by pores giving this phylum its name (Fig. 2A). Many sponges are asymmetrical, but some exhibit radial symmetry (Ruppert and Barnes 1994). The sponge is made up of two single-cell-deep layers and an intermediate mesenchyme called mesohyl (Fig. 2B). The outer layer, called sac, consists of flattened polygonal cells called pinacocytes and the inner layer consists of flagellated cells called collar cells or choanocytes. These two layers do not form a true epithelia, in fact no junction among cells have been detected in both layers, and it is controversial whether the outer layer forms a syncytial frame (Harrison and De Vos 1991). The middle layer, the mesohyl, consists of gelatinous protein or/and carbohydrate material, a range of mobile cells, and a skeleton formed by a network of fibrous flexible protein called spongin and needle-like structures called spicules. Spicules are usually made of calcium carbonate or oxides of silicon, the chemical composition and the shape of the spicules are used for classification (Ruppert and Barnes 1994).



**Fig. 2.** General body plan of a simple sponge. (A) Part of a colony, (B) longitudinal section including details of choanocytes and an amoebocyte. The inside of the sponge (spongocoel) is to the left in B.

### **How sponges get food: an elaborate filtering system**

The body plan of a typical sponge is shown in figure 2. Water enters from numerous small pores called ostia, carrying suspended plankton and other potential food. The ostia are surrounded by specialized cells called porocyte that open and close to control water flow. For the sponge depicted in figure 2, water flows directly into an open chamber called spongocoel, and then water leaves the spongocoel by a larger opening, the osculum. The osculum size is controlled by a special type of mesohyl cell called myocyte, which displays similarity to a smooth muscle cell in shape and contractility, however, unlike the true muscle cells the myocytes surrounding an osculum do not touch each other (Bagby 1966). The choanocytes have a tubular collar facing the spongocoel, a flagellum extends from the centre of this collar and the movement of the flagella create currents that force water through the sponge duct system. Suspended food particles in the water are drawn through the collar from below, trapped on the outside of the collar, and then by phagocytotic activity inside the cell. The choanocytes do not digest the captured food themselves, instead it is passed to a second cell type, an amoebocyte that is waiting in the mesohyl, and the amoebocytes carry the food to other cells that require nutrition. Thus, it is partly compensate for the sponge's lack of a circulatory system (Harrison and De Vos 1991).

The amoebocytes carry on other jobs, for example, they can undergo developmental changes to turn into any other cell type that may be required and this allows for growth, repair, and reproduction of the sponge. In addition, amoebocytes are responsible for producing the sponge skeleton.

### **The four extant classes of sponges**

Living poriferans have traditionally been classified into four classes: Calcarea, Hexactinellida, Demospongiae and Sclerospongiae.

Within these classes three types of sponge structure are recognized: the asconoid, the most primitive, is regular, tube-shaped, and radially symmetrical; the syconoid is a more irregular structure that displays some degree of folding of the body wall, but still maintain a basic radial symmetry; the leuconoid is highly irregular, displays the greatest degree of folding of the body wall, and has lost radial symmetry. In the leuconoid sponges choanocytes line the pockets formed by the convoluted body wall (Ruppert and Barnes 1994).

Members of the class Calcarea are characterized by a skeleton composed of crystals of calcium carbonate. Calcareous sponges are small, usually only a few centimetres high, and are generally dull in appearance, although several species are brightly coloured and the spicules

often protrude through the sac, giving to the organism a rough texture. Members of this class are among the simplest sponges, and all three morphological types - asconoid, syconoid, and leuconoid - are represented. There are approximately 150 known species, exclusively marine and shallow-water inhabitant (Ruppert and Barnes 1994).

Hexactinellida (glass sponges) are deep-sea sponges. They lack an epidermal covering, and their skeletons are composed of spicules of silica. The spicules are more complex than needle-shaped, in fact they have six points or some multiple thereof and often form a scaffold. Glass sponges are pale in colour and are cup or basket shaped. The spongocoel is large, and the osculum is covered by a grating of fused spicules. When the living tissue is removed, the cylindrical skeletons often have the appearance of spun glass. The body plan of Hexactinellida is between syconoid and leuconoid (Ruppert and Barnes 1994).

To the class of Demospongiae belong most sponges. It includes sponges with a skeleton made up of silicon-containing spicules or spongin fibres or both. In the latter case, the spongin provides a matrix in which the spicules are embedded. The Demospongiae vary in size from small, encrusting forms to very large, irregular masses. All are leuconoid; many have bright colours (Ruppert and Barnes 1994).

Members of the Sclerosponges are quite rare today and they are mostly known from fossils. There are a few modern species, which are only found on coral reefs in the West Indies and Pacific, where they contribute to the structure of the reefs (Chombard et al. 1997).

### **Reproduction in sponges**

Pieces of sponge are able to regenerate into whole new sponges. Asexual reproduction occurs by budding or by fragmentation. The buds may remain attached to the parent or separate from it, and each bud develops into a new individual. Freshwater sponges, as well as several marine species, form resistant structures called gemmules that can withstand adverse conditions such as drying or cold and later develop into new individuals. Gemmules are aggregates of sponge tissue and food, covered by a hard coating containing spicules or spongin fibres. Sexual reproduction also occurs. Most sponges are hermaphroditic, but in some species the sexes are separate. Sexual reproduction is timed to environmental cues such as water temperature, tides, and the phase of the moon. During the mating season sperms are released into the water and eventually enter the porocytes of the female sponge. The sperms are later taken up by choanocytes, and then passed on to amoebocytes, which carry them to egg cells located in the mesohyl (Harrison and De Vos 1991).

### **Distinct patterns of development in sponges**

The development of sponges is not yet well understood, it is not even clear if the cellular movement during the larval formation can be compared to the process of gastrulation of other animals, in fact some authors argue that an adult sponge does not have any gut, so they cannot undergo gastrulation (Ereskovsky 2000). Other authors consider the choanocytes as the gut epithelium, for them gastrulation in sponges occurs during the metamorphosis by inversion of ectoderm and endoderm. Thus, after metamorphosis the ciliated cells of the ectoderm give rise to the choanocytes and the cells of the endoderm to the pinacocytes (Simpson 1984). For others those layers that invert at metamorphosis are already formed from differentiated cells and gastrulation occurs during larva formation by cellular movement (Leys and Degnan 2002). Four main types of larva are found in the porifera (Leys and Degnan 2002, and references therein): the coeloblastula, a ciliated hollow blastula-like larva (Amano and Hori 2001); the amphiblastula, a hollow blastula-like larva with ciliated cells confined in one hemisphere, typical of the class Calcarea; the paranchymella, a ciliated egg-shaped larva that is typically solid, but may have internal chambers; and the trichimella of hexactinellid sponges, which has multiciliated cells in a skirt around the middle and syncitial tissues like the adults. Most of the sponge larvae are paranchymella type and look like the planula of cnidarians and typically settle on the anterior pole and then undergo metamorphosis (Leys and Degnan 2002). The mechanisms of metamorphosis in paranchymella are not yet clear. Experiments done with radioactive cell tracers do not show any inversion of germ layers (Misevic et al. 1990), while natural markers show inversion of those layers (Amano and Hori 1996).

### **Morphological evidences for a third embryonic layer in a sponge larva**

In a recent study a fluorescent cell marker that remains locked inside the cells for several days was used to study the developmental processes of the sponge *Reniera sp.*, a desmosponge with paranchymella larvae (Leys and Degnan 2002). This sponge has large eggs of 300  $\mu\text{m}$ , the fertilization is internal and the cleavages are irregular. Blastomeres differentiate early, before any cellular movement, by formation of three types of cells: macromeres, with yolk inclusion; and two kinds of micromeres, one with dark cytoplasm, called pigmented cells, and the other with pale cytoplasm. Gastrulation begins with the movement of the pigmented cells towards one pole that will become the posterior pole and those cells can form a ring (Leys and Degnan 2001). After additional movements a late stage embryo has a three-layered appearance. The centre of the embryo consists of amoeboid cells in a dense matrix, whereas



the outer layer is composed of columnar epithelial cells with long cilia. Between the inner and outer layers there is a thin layer of cells that lie perpendicular to the columnar epithelial cells and form a circumferential sheath around the whole embryo. It is thought that this third layer originates from micromeres that migrate during gastrulation from the periphery of the embryo to their final position at the late-stage embryo (Leys and Degnan 2002). In the later phases of *Reneiera* development, the fate of the cells belonging to this thin layer placed between ectoderm and endoderm is not clear. Leys and Degnan (2002) suggest that this middle thin layer needs further investigation. On the basis of these morphological observations it is possible to speculate that this thin layer corresponds to the mesoderm of higher animals.

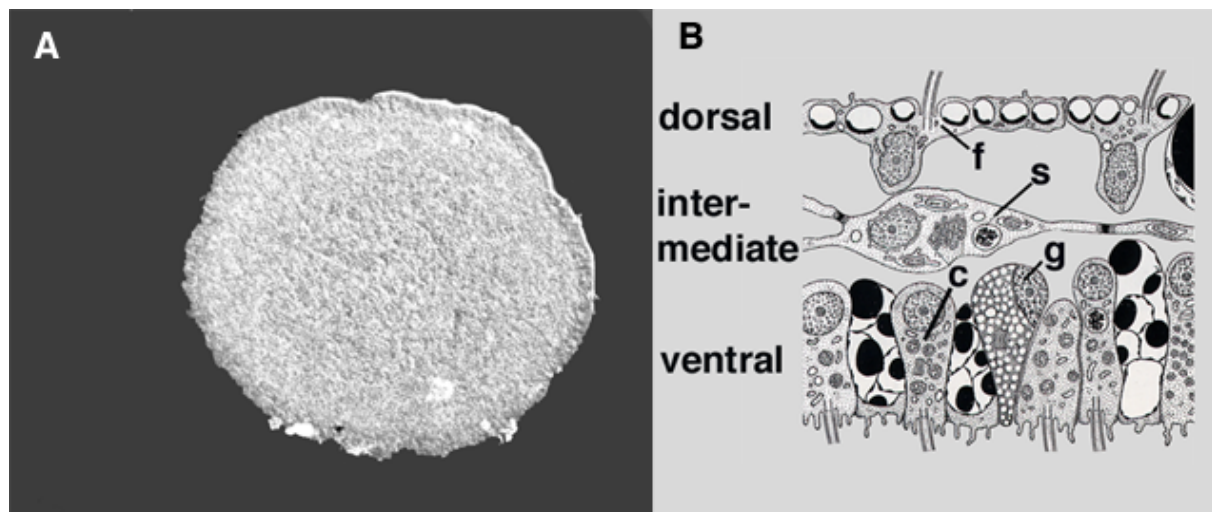
Molecular data could help to better understand whether there is an evolutionary link between this embryonic sponge middle layer and higher metazoan mesoderm. Some genes encoding transcription factors or components of signalling pathways that are characteristic of higher metazoans have been found in sponges (Manuel and Le Parco 2000; Wimmer et al. 1999), but their expression patterns have not been studied yet in sponge development.

## **PLACOZOA: AN ENIGMATIC PHYLUM**

*Trichoplax* is an enigmatic organism described for the first time by F. E. Schulze (1883), who was not able to put it in any known phyla due to its peculiar characteristics. Already in these early studies the main features, such as the small size of only a few millimetres with a constant thickness of 0.02 mm, no symmetry axes, no organs, a dorso-ventrally three-layered body and continuous changes in shape were observed (Fig. 3A). The description of this new animal type in German had an immediate echo in the scientific community and an English report appeared in the first volume of *Science* (Minot 1883). An apparently related species was observed in Naples and named *Treptoplax reptans* (Monticelli 1896), but since this original description it was not found again. Later, *Trichoplax* was described as the planula larva of the hydrozoan *Eleutheria krohi* (Krumbach 1907). Although this work was criticized (Schubotz 1912), the idea of *Trichoplax* as a hydrozoan larva was accepted and reported in textbooks (Hyman 1940). *Trichoplax* disappeared from the literature until 1971, when new observations suggested that *Trichoplax* is not a larva (Grell 1971; Miller 1971). *Trichoplax* is now thought to be the only species from the phylum Placozoa (Grell and Ruthmann 1991; Ruppert and Barnes 1994; Syed and Schierwater 2002).

### The body organization of *Trichoplax*

The body structure of *Trichoplax* consists of an upper thin epithelial layer with flattened mononuclear cells, a basal epithelial layer with two cell types, column-shaped cells with cilia and round-shaped cells, called gland cells, without cilia. Between these two layers there are star-shaped cells, that form a syncytial net (Fig. 3B). *Trichoplax* move by changing their shape similar to amoebae, due to the contraction of the microtubular system of the star-shaped cells (Thiemann and Ruthmann 1989). A nervous system seems to be missing, but with an antibody against RFamide, which recognizes neuropeptides in many animal phyla, a few isolated cells can be detected (Schuchert 1993a). The feeding system is located in the ventral layer; particles already digested by enzymes produced from the gland cells can be incorporated by endocytosis (Ruthmann et al. 1986). Apparently, *Trichoplax* is missing a structured extracellular matrix and a recognizable basal lamina (Grell and Ruthmann 1991). The epithelial cells are connected by belt desmosomes in both epithelial layers, no other type of junction has been described (Ruthmann et al. 1986). Feulgen cytometry has shown that *Trichoplax* is the species with the lowest amount of DNA of all the Metazoa (Ruthmann and Wenderoth 1975).



**Fig. 3.** General body plan of *Trichoplax adhaerens*. (A) A typical animal shows an arbitrary body shape without any symmetry axis. Animals vary in size from 0.5 to 5 mm and contain more than 10'000 cells. (B) Depiction of a *Trichoplax* section shows the “three layers” body structure (after Grell and Ruthmann 1991). *f*, flattened dorsal cell; *s*, star-shaped cell of the intermediate layer; *c*, column-shaped cell; *g*, gland cell.

### Reproduction and regeneration in *Trichoplax*

Still little is known about reproduction of *Trichoplax*. In aquaria, *Trichoplax* does it irregularly by fission. Under special circumstances it was observed to build hollow swarmer spheres, stolons and eggs developing to up to the 64 cells blastula stage (Grell 1971; Ruthmann et al. 1981; Thiemann and Ruthmann 1991). Although the three layers of *Trichoplax* could be compared to ectoderm, mesoderm, and endoderm of higher animal, as

Schulze noticed in 1883, the presence of three layers at embryonic stages could be proven or disproved only by comparing developmental processes.

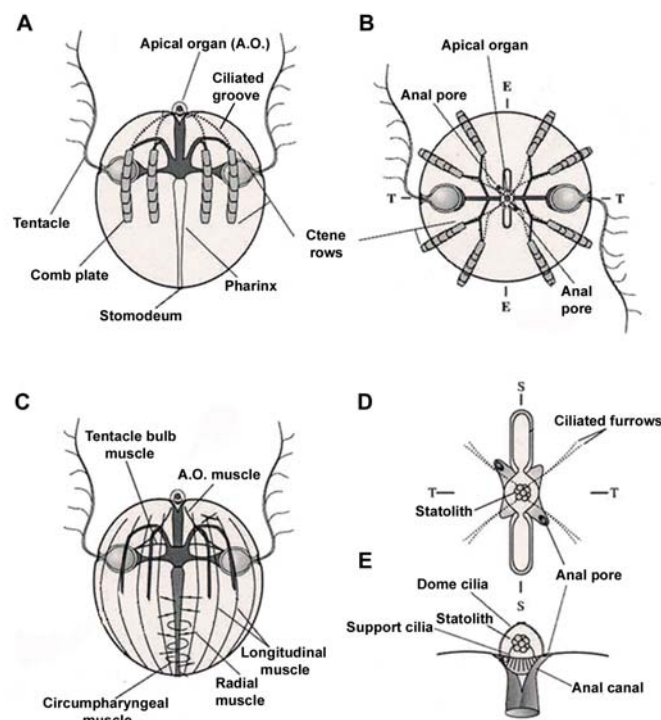
As most of the lower metazoans, *Trichoplax* has a regeneration potential inducible by cutting a piece of the body. No data are available regarding the dynamic of wound closure, so it is not clear whether in *Trichoplax* the regeneration follow a morphallactic process where no proliferation is required or an epimorphic process. Only the morphological changes during regeneration have been observed. The wound closure process is fast, just after the cut the central part of the body contracts, while the two edges of the wound get close by a sliding movement the external regions on the central part of the body. In half an hour the two lips of the wound touch each other and need 10 minutes to fuse each other, so less than one hour is required to close the wound (Schwartz 1984). Moreover, it has been shown that fragments containing both peripheral and central regions of the body are able to regenerate, while fragments containing only a central part of the body or a peripheral part are not able to regenerate (Schwartz 1984).

## **CTENOPHORA: HIGH COMPLEXITY IN A BIRADIAL BODY ORGANIZATION**

Ctenophores are a rather enigmatic group of marine carnivores. Virtually all species are holopelagic and feed with the aid of a pair of tentacles located on opposite sides of the mouth, although one derived group, the beroids, has lost their tentacles and another taxa, the platyctenes, has lost their comb plates and became benthic (Podar et al. 2001).

The major longitudinal axis of all ctenophores is called the oral-aboral axis, along which run two planes of symmetry, the tentacular and the esophageal axes (Fig. 4A and B, Martindale and Henry 1999, and references therein). Most of the sensory and locomotory structures of ctenophores are cilia-derived. Ctenophores have eight longitudinal arrays of ectodermally derived ctene (comb) plates and each comb plate is composed of thousands of cilia in longitudinal arrays that form “paddles”. The synchronized beating of these comb plates along each comb row accounts for the movements of these animals through the water column. The control of comb plate beating in individual comb rows is accomplished in large parts by a central sensory structure called the apical organ located at the aboral pole of the animal. The apical organ is a richly innervated statocyst or "gravity amplifier" (Fig. 4D and E). The statocyst is housed under a cupola of non-motile dome cilia and is composed of a group of mineral-containing lithocytes placed upon four groups of balancing cilia. Each group of balancing cilia is functionally connected to a pair of ctene rows by the ciliated grooves in

each of the four quadrants. The position of the lithocytes atop the balancing cilia regulates the beating of the rows of comb plates. The apical organ also has neural connections to both tentacles (Hernandez-Nicaise 1991). The cells thought to mediate detection of light and regulate spawning are modified ciliated cells in the floor of the apical organ that contain lamellate bodies and synapse onto adjacent cells in the apical organ (Hernandez-Nicaise 1991, and references therein). The apical organ, with its high concentration of synapses integrates sensory information and initiates the appropriate motor and behavioural output (Hernandez-Nicaise 1991, and references therein). Although these data suggest that the apical organ with its centralized function could be homologue to the bilaterian brain, *forkhead*, a gene typically expressed in the bilaterian brain, is not expressed in the apical organ (Yamada and Martindale 2002). Thus, from the molecular point of view, the evolutionary relationship between the apical organ and the bilaterian brain is doubtful based on the genes tested so far.



**Fig. 4.** General body plan of common ctenophores. (A) Lateral view, (B) aboral view of A, two anuses and an apical organ are typical of ctenophores, (C) Anatomical scheme of the muscle structures of the jelly body, (D) and (E) are enlarged depiction of the apical organ from aboral and lateral view, respectively (after Martindale and Henry 1999).

### Distinct body organization within two ctenophore classes

There are two generally recognised classes of ctenophores: Tentaculata and Nuda. The class Tentaculata is characterized by the presence of two tentacles, which stem from the aboral margin. This class also contains species that are flattened at either the tentacular (or lateral) axis, or the oral-aboral (or dorso-ventral) axis. Examples include *Pleurobrachia pileus* (Order

Cydropida), which has a characteristic ovoid shape (Ruppert and Barnes 1994) and *Cestum veneris* (Order Cestida), which has a laterally flattened form resembling a belt. It grows to nearly a metre in length, and is relatively common in Mediterranean waters (Ruppert and Barnes 1994). *Vallicula multiformis* (Order Platyctenida) is an example of a dorso-ventrally flattened form, which has entirely lost the typical rows of cilia, but retains two tentacles. It moves over marine algae and sediment by crawling, much as a turbellarian flatworm might do (Freeman 1967). *Gastrodes* is unusual in being parasitic in nature, living in the body of the pelagic Thaliacean *Salpa*. *Gastrodes* is also believed by some to provide a direct link between the Ctenophora and Cnidaria, having a planula larval stage (Harbison 1986). *Tjalfiella* (Order Platyctenida) has been found at depths of around 500 m and is another unusual ctenophore, appearing to be entirely sessile in habit (Harbison and Madin 1982)

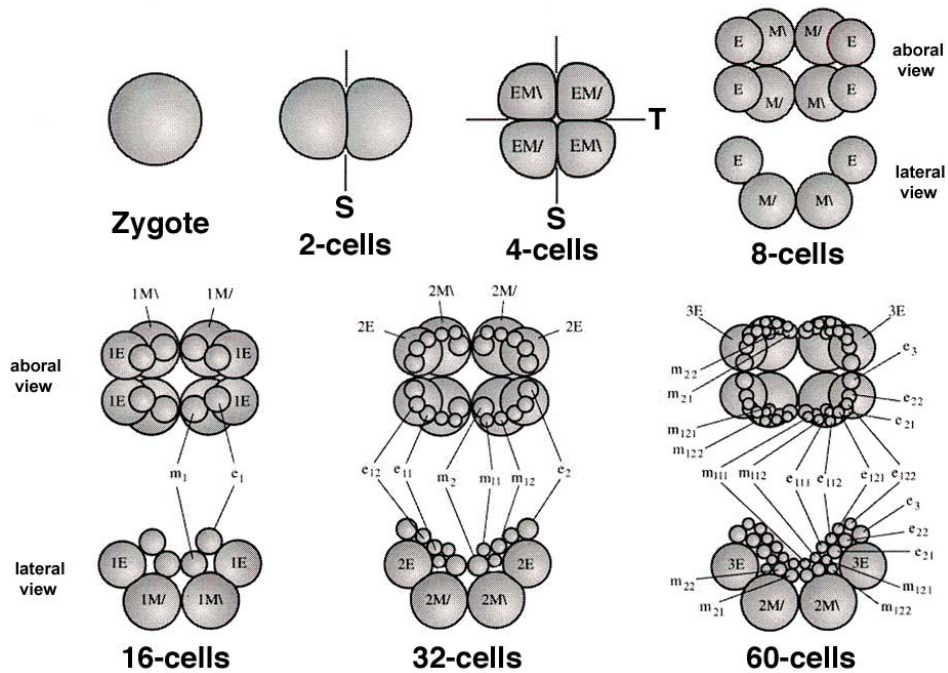
The class Nuda is represented by just one order, which contains the species *Beroe*. Members of this class are generally flattened in the tentacular plane, possess no tentacles, and have a heavily branched gastro-vascular system. *Beroe* (Order Beroida) is a complex form that has a body wall covered with series of canals originating from its gastro-vascular tract. *Beroe* is pinkish in colouration, roughly thimble shaped, and has a large muscular mouth used to engulf other ctenophoran species that constitute its prey (Ruppert and Barnes, 1994)

### **The unique developmental pattern of ctenophores**

Adult ctenophores are hermaphrodite. Eggs and spermatozoans develop internally in discrete bands and, when mature, they are released via the mouth for external fertilisation. Development of the embryo is by determinate cleavage and growth is via a cydippid larva, which is essentially an exact miniature of its parents. The ctenophores are noted as having a substantial capacity for regeneration (Ruppert and Barnes 1994).

Ctenophore development is unique and does not bear obvious similarities to embryogenesis in any other extant organisms (Fig. 5). First cleavage is unipolar, passes through the future sagittal plane and defines the oral pole (Freeman 1977). Second cleavage passes through the future tentacular plane and gives rise to four equal-sized cells. Each of the first four blastomeres divides to give rise to E (end) and M (middle) blastomeres, which occupy distinct positions in the embryo (Fig. 5). As shown in figure 4, each E blastomere gives rise to a series of three smaller micromeres at the aboral pole (called e1, e2 and e3), while the M blastomeres generate two aboral micromeres (m1 and m2). Following a period of rapid cell division and gastrulation, differentiated cell types begin to appear. Embryogenesis

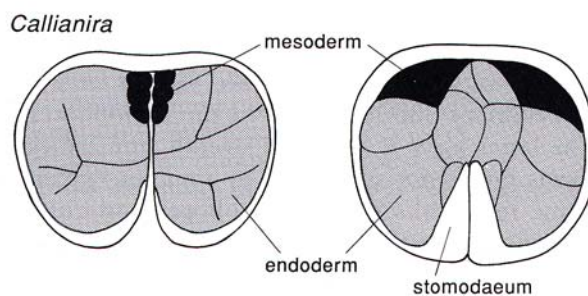
in most ctenophores is rapid, and direct, with a juvenile *Mnemiopsis leidyi* developing in less than 24 hours at 18-20°C (Martindale and Henry 1999).



**Fig. 5.** Description of ctenophore development from zygote to 60 cells stage. The first cleavage plane is sagittal, while the second cleavage plane is perpendicular and corresponds to the tentacular plane. At the third cleavage originate two distinct cell lineages, E and M, which one characterized by a peculiar position due to a “diagonal” cleavage. Each blastomers of the 60 cells stage embryo give rise to a defined cell population in the adult animal (after Martindale and Henry 1999).

***Mnemiopsis leidyi*: establishing the fate of early blastomers**

An embryonic fate map for the first 60 cells has been generated using intracellular markers for the ctenophore *Mnemiopsis leidyi* (Martindale and Henry 1999).



**Fig. 6.** Early and late gastrula of the ctenophore *Callianira liabata*. A third germ layer is highlighted (after Nielsen 2001).

During gastrulation the aboral micromeres proliferate and spread over the mesendodermal macromeres by epiboly, with the blastopore becoming the adult mouth. All

muscle cells are derived from micromeres born at the oral pole of endomesodermal precursors (2M and 3E macromeres) and the layer of cells originating from these micromeres could be considered the equivalent of the mesoderm. A similar pattern for ctenophore development was already identified more than a century ago in *Callianira bialiata* (Fig. 6, Nielsen 2001, and references therein). A unique pattern of development is observed in the 2M lineage. Unlike any of the other cells injected in similar experiments, the 2M macromeres displayed a phenomenon known as “diagonal development” or “diagonal determination.” In other words, the pair of 2M cells located along the anal axis (i.e., 2M $\backslash$ ) exhibit different fates from those located along the orthogonal axis (i.e., 2M/). 2M $\backslash$  macromeres are located along the so-called ctenophore “anal axis” and generate a pair of endodermally derived anal canals that connect the gut cavity to the external seawater in diagonally opposed quadrants at the aboral pole (Martindale and Henry 1995). The anal canals and their external orifices are relatively small and exist in two of the four quadrants, each of which is clearly derived from a single 2M $\backslash$  macromere. Although the 2M/ macromeres also give rise to endoderm, they do not contribute to the formation of the anal canals. Instead, they give rise to distinct sets of muscle cells, including longitudinal body wall muscles, circumpharyngeal muscles, tentacle to apical organ muscles, and the longitudinal muscles, which run along the external surface of the pharynx. In addition, they give rise to a population of wandering multipolar mesenchymal cells (“stellate” mesenchyme), which may have phagocytotic activity. All of these mesodermal cell types are notably absent in derivatives of 2M $\backslash$  macromeres (Martindale and Henry 1999).

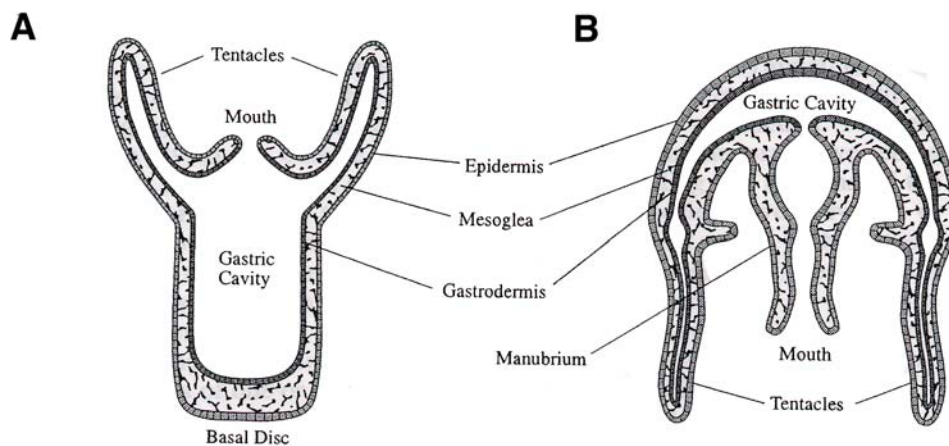
The tentacles form from two ectodermal patches that invaginate and fuse with muscle precursors born from the oral end of the E and M lineages (Martindale and Henry 1999). The tentacles are generated continuously from the central region of the tentacle bulbs and exit the animal through the ectodermal tentacle sheathes. The apical organ arises from the coalescence of components from all four quadrants toward the aboral pole as the ctene rows have begun to form following gastrulation (Martindale and Henry 1999).

## **CNIDARIA: A WIDE VARIETY OF LIFE FORMS CHARACTERIZED BY A UNIQUE CELL TYPE, THE NEMATOCYTE**

There are four widely recognised classes of cnidarians: Hydrozoa, Scyphozoa, Cubozoa and Anthozoa. They are represented by polyps such as sea anemones and corals, and by medusa such as jellyfish. A polypoid or a medusoid cnidarian is a radially or biradially symmetrical, uncephalized animal with a single body opening, the mouth. The mouth is surrounded by

tentacles that are studded with microscopic stinging capsules known as nematocysts that are the agents of offence and defence. The possession of intrinsic nematocysts is the defining characteristic of the phylum, nematocysts are the most diverse and widespread of three types of cnidae, hence the preferred name of the phylum (Ruppert and Barnes 1994).

The body and tentacles of cnidarians consist of two cell layers, the inner layer called gastrodermis and the outer layer called epidermis (Fig. 7). Between the two cell layers is the mesoglea, which ranges from little more than a glue to bind the layers, for example in *Hydra*, to the vast bulk of the animal, for example in jellyfish of the Class Scyphozoa. The body has a single sac-like body space, the coelenteron, which communicates with the surrounding medium through the mouth. The coelenteron, also termed the gastrovascular cavity, serves for gas exchange and digestion.



**Fig. 7.** Body plans of two life stages of cnidarians. (A) Cross section of a polyp and (B) cross section of a medusa. Both stages possess an outer epidermis and an inner gastrodermis; these two layers are separated by an acellular tissue called mesoglea (after Martindale and Henry 1998).

All cnidarians are carnivorous, with cnidae and tentacles active in prey capture. Because polyps are typically sessile, and only medusae possess sensory structures, cnidarians are generally believed to be passive predators, feeding on prey items that blunder into their tentacles. Some cnidarians can absorb dissolved organic matter directly from seawater, but it is not known how widespread this ability is (Schlichter 1975). Living within the tissues of anthozoans of many species and hydrozoans and scyphozoans of a small number of species are unicellular algae from which the animals derive reduced carbon. Dinoflagellate symbionts, termed zooxanthellae, are by far the most common algal symbionts; they are exclusively marine, while Green algal symbionts, termed zoochlorellae, occur in both marine and freshwater cnidarians (Ruppert and Barnes 1994).



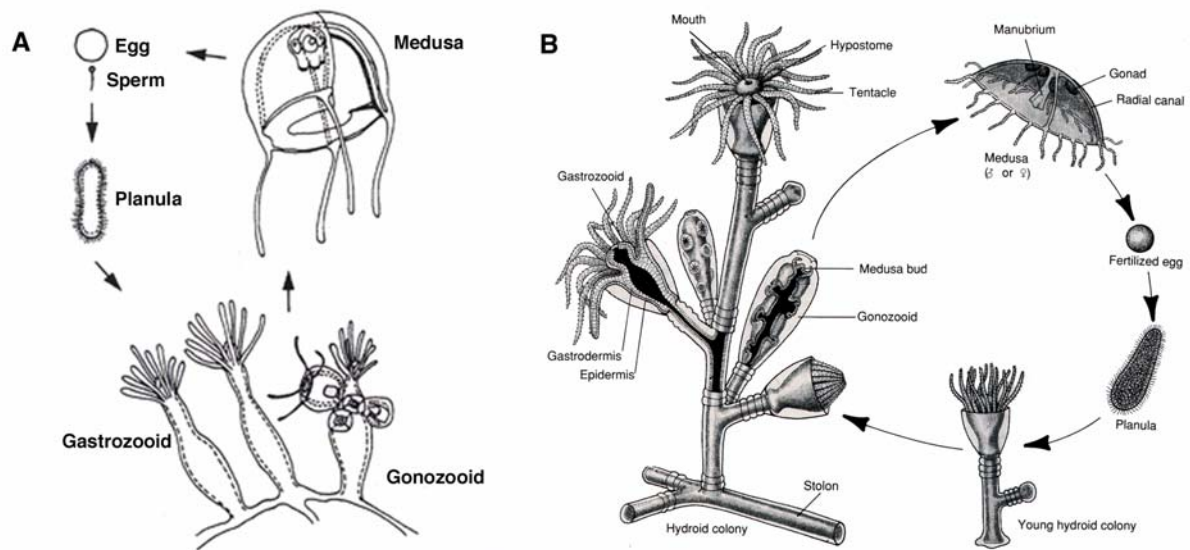
### The life stages of cnidarians

The text-book depictions of the typical cnidarian life cycle is an alternation between a polyp and a medusa termed metagenesis, the former is the asexual stage and the latter is the sexually reproductive stage (Fig. 8). In fact, an attribute of the entire class Anthozoa is the absence of a medusa. At least some individuals of all anthozoans species form gametes; those of some species may reproduce vegetatively as well. The other three classes - Cubozoa, Hydrozoa, and Scyphozoa - are often grouped as the "Medusozoa" because the medusa phase is present in all of them (Collins 2002). Indeed, the medusa dominates the life cycle of members of the classes Cubozoa and Scyphozoa. Life cycles of the hydrozoans are the most diverse in the phylum: although the polyp is the more conspicuous and persistent stage in most taxa, some lack the medusa phase, whereas others lack the polyp phase (Galliot 2000). *Hydra*, which is considered by many researches as a model organism of this phylum, is utterly atypical: it lacks a medusa, it has aggregations of gametogenic tissue that function as gonads, and it is among only a handful of freshwater cnidarian species.

Polyps are usually sessile, and colonial in nature, their body plan basically comprises a thin-walled cylinder, which has an opening at one end. This is known as the hypostome, and is surrounded by tentacles bearing nematocysts. Colonies are not always comprised of homologous units, but may be polymorphic, with specific specialised functions, for example gastrozooids and medusozoids in *Podocoryne carnea* (Fig. 8A). The anatomy of the medusa is more complex. Most of the differences are found in the medusa bell, which not only can carry complicated sense organs such as lens eyes, statocysts, and nerve rings but also consists of two nonmyoepithelial cell layers and additionally a third layer of epithelial mononucleated striated muscle cells, so medusal musculature can be well developed, and capable of rapid contractions (Ruppert and Barnes 1994). The hypostome is termed manubrium in the medusa stage. The mouth gives access to the gastric cavity within which digestion occurs and from the gastric cavity develops a radial and ring canal system that allows diffusion of nutrients all over the bell.

The cnidarian larva is the planula, a pear-shaped, entirely ciliated animal. A typical cnidarian life-cycle is described in figure 8. Male and female medusa spawn freely into the sea, where fertilization occurs and a planula develops. At metamorphosis, the planula settles on and attaches to the substratum, where it metamorphoses into a polyp. The primary polyp produces additional polyps asexually, by budding, stolonial outgrowth, or some other process, to form a clone or a colony. At the appropriate time, determined by factor such as size of the colony or environmental conditions medusae are produced asexually. In Cubozoa, an entire

polyp metamorphoses into a medusa. They are released to take up a pelagic existence and the cycle begins anew.



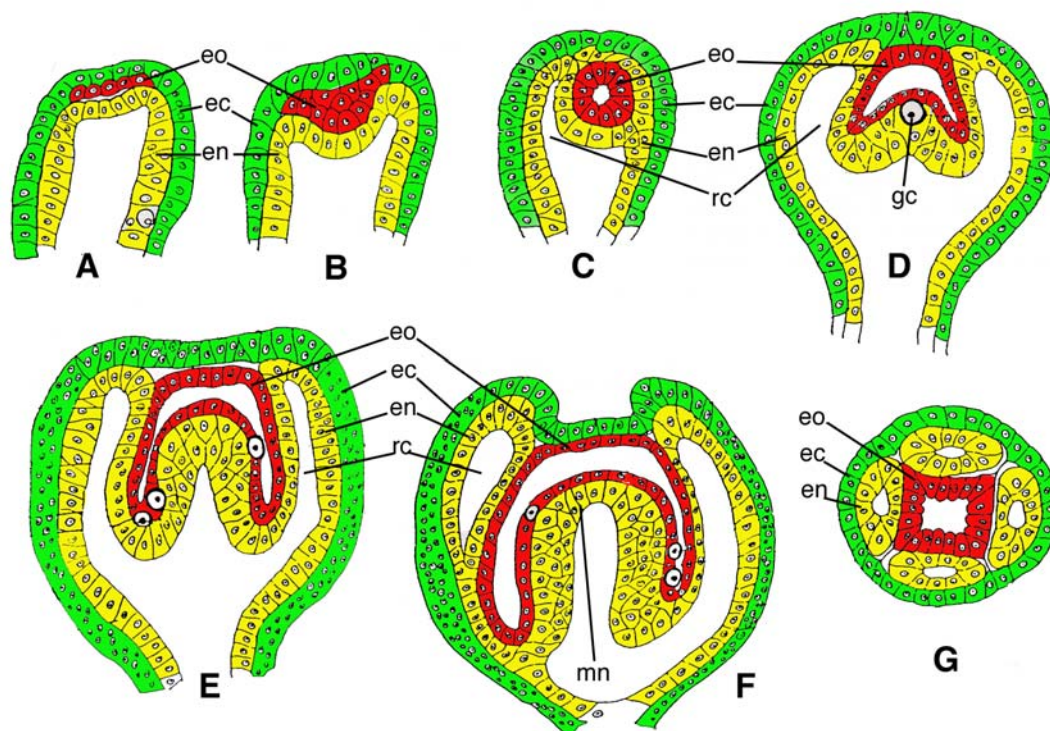
**Fig. 8.** Typical hydrozoan life cycles. (A) The lifecycle of *Podocoryne carnea* comprises a pelagic medusa stage and a benthic, colonial polyp stage. The transformation of the swimming planula to the polyp is called metamorphosis, the medusa will arise from the polyp through budding. (B) The life cycle of *Obelia sp.* is represented in many textbooks (after Rupert and Barnes 1994).

### Reproduction in cnidarians

Cnidarians have both asexual and sexual way of reproduction. Asexual reproduction is by budding, depending upon the species and the life history stage, budding may produce a polyp, a medusa, or a frustule, which is similar to a planula larva without cilia that separates from the parent and develops into a polyp. Polyps may bud off either polyps or medusae, some medusae bud off medusae. There are several ways of sexual reproduction. The gametes could be released and the fertilization is external, or the fertilization can be internal in the coelenteron. In each case temperature and light play a key role in the timing of maturation and release of gametes. Cleavages are total and the blastomer cell fate cannot be determined. The planula develops by a process of gastrulation, can occur in several ways and usually ends in a bleared swimming larva (Tardent 1978).

### The third germ layer in hydrozoans: the entocodon

In hydrozoans the medusa forms by budding from polyps. During this process appears a third cell layer between ectoderm and endoderm. This mesoderm-like layer is called entocodon and separates from the ectodermal layer early in medusa development (Fig. 9).



**Fig. 9.** Sketch for early developmental stages of a hydrozoan medusa bud. (A) At the first stage only few cells of the entocodon originate from the ectoderm, (B) the second stage is characterized by a rapid proliferation of cells in the entocodon, (C) at the third stage the entocodon forms a cavity and the endoderm starts forming radial canals, (D) during the fourth stage the three layers are convoluted similarly to the adult medusa, (E) there is a continued proliferation of cells at the fifth stage, (F) the continuous invagination of endoderm give rise to the main body structure of the medusa, the manubrium. (G) Cross section of D. *eo*, entocodon; *ec*, ectoderm; *en*, endoderm; *rc*, ring canal; *gc*, germ cell; *mn*, manubrium (after Hyman 1940).

The entocodon is located between the distal ectodermal and the endodermal tissue, and is separated from both layers by an extracellular matrix. Entocodon cells in early bud stages are embryonic in appearance and highly proliferative. Later the entocodon forms a coelom-like cavity, which finally connects to the outside by the developing velar opening. In older bud stages mitotic activity in the bell gradually stops and the outer wall of the entocodon differentiates into the striated muscle while the inner wall forms the smooth muscle of the feeding and sex organ of the animal. Thus, in hydrozoans the striated muscle derives from the entocodon, while in bilaterians the striated muscle tissue originates from mesoderm. This similarity together with the formation of a coelom-like cavity led to the idea that the entocodon could be a mesoderm-like layer (Boero et al. 1998). With the aim to find molecular evidence for the homology of entocodon and mesoderm several studies were done with the hydrozoan *Podocoryne carnea* and several mesoderm-specific genes known from bilaterians.

**Molecular origin of mesoderm: the entocodon genes of *Podocoryna carnea***

The molecular analysis of muscle development in *Podocoryne* has demonstrated that the structural genes for a *tropomyosin* and a *myosin heavy chain* are structurally and functionally conserved and specific for the striated muscle tissue (Yanze et al. 1999). Furthermore, the presence of the homeobox gene *Otx* (Müller et al. 1999), a head and gastrulation regulator in bilaterians, in jellyfish striated muscle and the basic helix-loop-helix (bHLH) factor *Twist* during the formation of the entocodon in medusa development (Spring et al. 2000), indicates that genes with specific roles in mesoderm patterning of bilaterians are already present in the common ancestor with bilaterians. Next to homeodomain and bHLH transcription factors, the best-studied regulatory genes are members of the T-box, MADS-box and zinc finger families, such as *Brachyury*, *Mef2*, and *Snail*, respectively (Spring et al. 2002). These three gene families are involved at different levels in the specification of the mesodermal and myogenic lineage of bilaterian animals from *Drosophila* to vertebrates.

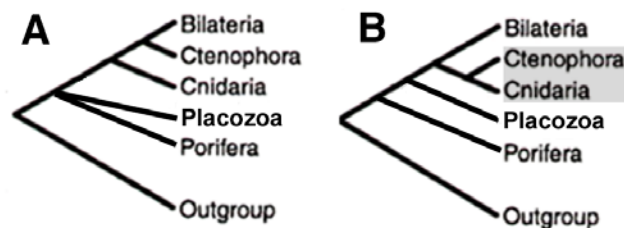
*Podocoryne Twist* is expressed in the early embryo until the myoepithelial cells of the larva differentiate and then again during medusa development. There, the gene is detected first when the myoepithelial cells of the polyp dedifferentiate to form the medusa bud and later *Twist* is found transiently in the entocodon, a mesoderm-like cell layer that differentiates into the smooth muscle and striated muscle of the bell. In later bud stages and the medusa, expression is seen where non-muscle tissues differentiate (Spring et al. 2000).

To investigate the hypothesis that the entocodon of jellyfish is homologous to the mesoderm of bilaterians, a *Podocoryne* homolog of *Brachyury*, *Mef2*, and *Snail* were isolated and expression patterns were studied throughout the life cycle and specifically during muscle development. The results demonstrate that all three genes are expressed during myogenic differentiation (Spring et al. 2002). Additionally, as is true for their bilaterian cognates, they appear to have other functions as well. The sequence and expression data demonstrate that the genes are structurally and functionally conserved and even more similar to humans or other deuterostomes than to protostome model organisms such as *Drosophila* or *Caenorhabditis elegans* (Spring et al. 2002). The data further strengthen the hypothesis that the common ancestor of cnidarians and bilaterians already used the same regulatory and structural genes and comparable developmental patterns to build muscle systems.

## MOLECULAR PHYLOGENY AND RELATIONSHIPS AMONG BASAL METAZOANS

In order to study the evolution of various molecular, cellular and developmental characters it is essential to determine the early branching order of animal phyla. Thus, both the phylogenetic relationships among early animal phyla and the evolutionary pathway that the metazoans followed from unicellular organisms are central topics in many researches.

Cladistic analyses of morphology tend to suggest a step-wise arrangement where either Porifera or Placozoa diverge first, then Cnidaria, and finally Ctenophora, from the lineage leading to Bilateria (Fig. 10A; Nielsen 2001). Until the arrival of molecular analyses, morphological data provided many arguments. For several reasons cnidarians and ctenophores were treated as a single clade named coelenterates (Fig 10B; Hyman 1940). Although this view is rejected by molecular data, there are at least three major issues in relation to the phylogenetic status of diploblasts that remain poorly resolved: whether all the metazoans including diploblasts and triploblasts are monophyletic, whether the diploblasts are monophyletic or paraphyletic, and how the diploblast phyla are related (Cavalier-Smith et al. 1996; Christen et al. 1991; Collins 2002; Field et al. 1988; Kim et al. 1999; Kobayashi et al. 1996; Medina et al. 2001; Wainright et al. 1993). Therefore, additional data are needed to improve our understanding of the early phylogenetic history of animals.



**Fig. 10.** Two alternative evolutionary hypotheses for non-bilaterians (after Collins 1998).

### Monophyletic origin of metazoans

In view of similarities in the organization of the epithelial and mesenchymal tissues and connective tissue, including collagen and other molecules (for review, see Müller et al 2001), between sponges and cnidarians as well as higher animals, the arguments for the monophyly of Metazoa are very strong. However, analyses of the origin of metazoans using methods of molecular phylogeny with rRNA genes as a phylogenetic index had revived the controversy as to whether all metazoans can be viewed as successive offshoots within a single

monophyletic unit, or different processes of cell aggregation led to parallel radiation and different patterns of body plan organization.

The early publications of phylogenetic trees based on partial sequences of animal *18S rRNA* genes (Field et al. 1988) and *28S rRNA* genes (Christen et al. 1991) led to the idea that the two animal subkingdoms, Radiata and Bilateria, originated independently from different flagellate protozoan ancestors. Recently these conclusions have been widely criticized by several studies based on complete gene sequences of *18S rRNA* genes and the protein coding genes *elongation factor 1 alpha* and *Hsp70* (Borchiellini et al. 1998; Kim et al. 1999; Kobayashi et al. 1996; Peterson and Eernisse 2001; Wainright et al. 1993). All these studies support the monophyly of the animal kingdom including sponges, although the bootstrap value for monophyly is very weak.

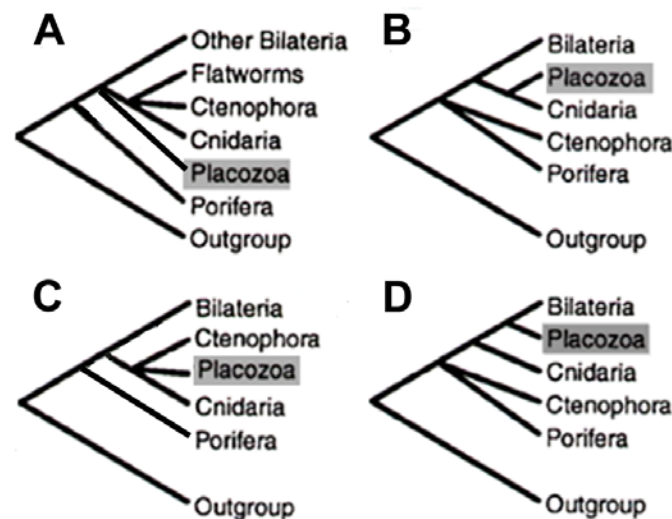
The monophyletic origin of animals was supported as well from the discovery of a receptor tyrosine kinase (RTK) from a sponge (Schacke et al. 1994), suggesting that the evolution of RTKs for cell-cell signalling could have played a role in the evolution of early metazoans. Moreover, the discovery of a RTK in choanoflagellates demonstrates that at least one family of metazoan signalling molecules was present before the origin of animals (King and Carol 2001). Phylogenetic and molecular genetic evidence (Wainright et al. 1993, Snell et al. 2001), together with the ultrastructural similarities between choanoflagellates and sponge choanocytes (Hibberd 1975), suggest that choanoflagellates may best represent the eukaryotic form of life immediately ancestral to the evolution of animals (Fig. 1).

Thus, although some steps have been done to fix the monophyletic origin of metazoans, further comparisons of the expressed genes in basal metazoans could infer a better view on the origin of metazoans. For instance, up to now no transcription factor family involved in development of all animal phyla but absent in fungi or plants was found.

### **Enigmatic evolutionary position of the placozoan *Trichoplax adhaerens***

Even if the monophyletic origin of metazoans including sponges is assumed to be correct, the early history of metazoans and the relationship between sponges, placozoans, ctenophores, cnidarians, as well as bilaterians is not resolved. A clear example of this confusion is the phylogenic position of the placozoan *Trichoplax adhaerens* (Fig. 11). On the basis of morphological features *Trichoplax* has always been thought to be a basal animal and one of the most convincing hypothesis places *Trichoplax* together with Eumetazoa, to the exclusion of sponges, in a group called Epitheliozoa (Fig. 11A; Ax 1996). In the last decade, with the rise of molecular phylogenetics, several analyses were carried out with the *18S rRNA* and *28S*

*rRNA* genes to determine the relationship of animal phyla. All possible combinations of grouping *Trichoplax* with others non-bilaterians were reported. *Trichoplax* has been suggested to be a sister group of cnidarians (Fig. 11B; Philippe et al. 1994; Wainright et al. 1993) or it has been suggested to belong to the cnidarians itself (Bridge et al. 1995). Moreover, *Trichoplax* has been grouped with ctenophores, as a sister group of cnidarians (Fig. 11C; Odorico and Miller 1997) and most recently as a sister group of the bilaterians, to the exclusion of cnidarians, ctenophores and sponges (Fig. 11D; Collins 1998; Collins 2002). Further analysis of *Trichoplax* genes and comparison with sponges, cnidarians, ctenophores, and bilaterians may reveal the presence of additional shared genes with potential significance to the evolutionary pathway and the origin of metazoans.



**Fig. 11.** Four alternative evolutionary hypotheses for the Phylum Placozoa (after Collins 1998).

### Phylogeny within sponges and cnidarians

Besides the evolutionary position of basal phyla, there are also arguments about the relationship of different classes within these phyla. It is not even clear whether all these phyla have a monophyletic origin, putting in question the origin of metazoans itself. While the monophyletic origin of ctenophores is accepted from both morphological and molecular points of view (Podar et al 2001) and placozoans consist of only one specie, sponges and cnidarians are discussed the position of classes belonging to sponges and cnidarians are most discussed.

Although cladistic analyses of morphological data have concurred in suggesting that Porifera are monophyletic, two opposing views have arisen from these studies (Böger 1988; Mehl and Reisinger 1991). One hypothesis holds that the two groups of cellular sponges, Calcarea and Demospongia, are the sister group of the syncytial sponges, Hexactinellida

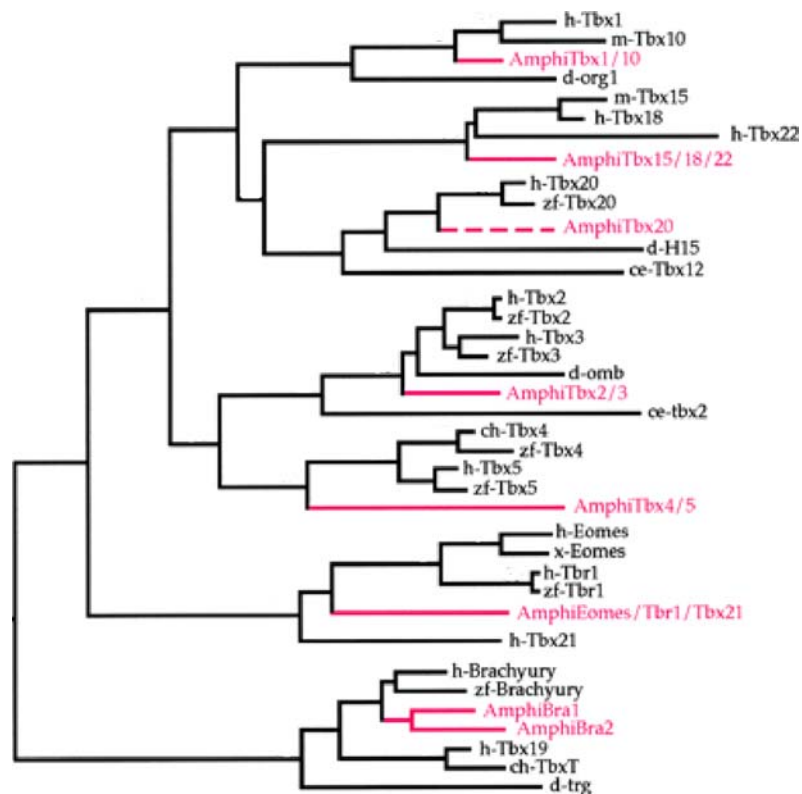
(Mehl and Reiswig 1991). The contrasting view is that the two sponge groups with siliceous spicules (Hexactinellida and Demospongia) form a sister group to sponges with calcareous spicules, Calcarea (Böger 1988). An alternative to both these hypotheses, weakly suggested by *18S rRNA* data, is that Calcarea is actually more closely related to nonsponge animals than it is to the other sponges (Cavalier-Smith et al. 1996, Peterson and Eernisse 2001; Collins 2002). Further evidence that Calcarea may be the sister to Eumetazoa comes from amino acid sequences of protein kinase C (Schuetze et al. 1999). These studies differ from those derived from 18S data by faintly suggesting that Hexactinellida may be the sister group to all other metazoans (Medina et al. 2001). Anyway, combination of 28S and 18S complete sequences unequivocally contradict this hypothesis and the proposal that Calcarea and Demospongia form a clade (Medina et al 2001), leaving two hypotheses pertaining to sponge phylogeny for further consideration. Porifera may be monophyletic, with Calcarea sister to the siliceous sponges, or Porifera may be paraphyletic, with Calcarea forming a clade with Eumetazoa. Nevertheless, three independent sets of molecular data show sponges to be paraphyletic 18S, 28S, and protein kinase C, implying that this alternative is most likely. If true, then an animal with sponge characteristics is in the direct ancestry of non-sponge animals.

Evolutionary relationships among the four extant cnidarian classes and specially the position of the anthozoans were the subject of many debates until the rise of molecular phylogeny (Hyman 1940, Schuchert 1993b). Anthozoans were alternatively considered the most basal or the most derived group. The former hypothesis assumes that the polyp is the original body form, with the medusa (and metagenesis) being derived. The latter perspective is that, in the life cycle, the medusa is gametogenic, and so constitutes the definitive, or adult, stage, with the polyp being a persistent larva. Thus, the polyp may have evolved secondarily, and loss of the original body form, the medusa, places Anthozoa as the most derived form. A comprehensive morphological cladistic analysis by (Schuchert 1993b) supports the basal position of Anthozoa with the Scyphozoa and Cubozoa being more closely related to each other than to Hydrozoa. 18S rDNA also supports the basal position of anthozoans, but do not resolve the relationships among Scyphozoa, Cubozoa and Hydrozoa (Bridge et al. 1995, Collins 2002).



## T-BOX GENES

T-box genes encode transcription factors that share a conserved DNA-binding domain of 180–200 amino acids, called the T-box domain. In the past ten years cloning and expression analysis of an increasing number of T-box genes from all metazoans brought progress in understanding the various roles that T-box genes have in animal development. It has been demonstrated that they are involved in the cellular movement during gastrulation, in the formation of the mesoderm and its derivatives and several of them have been linked to heritable human diseases. No T-box gene, however, can be recognized in genomes of fungi, plants or parasitic protists (Papaioannou 2001). The T-box domain is able to recognize half of a palindromic binding site in DNA (Müller and Herrmann 1997). On the basis of sequence homology several T-box subfamilies can be recognized. There is still some confusion about the classification of the 5 to 8 subfamilies and the nomenclature in different species (Papaioannou 2001; Ruvinsky et al. 2000), but at least comparisons of amphioxus and vertebrates suggests a conservative evolution of this gene family (Fig. 12).



**Fig. 12.** T-box gene family. By comparing human, amphioxus, *Drosophila* and *C. elegans* T-box genes 7 subfamily can be recognized. The Tbx subfamily and additional *C. elegans* genes have been excluded (after Ruvinsky et al. 2000).

**Evolutionary conserved roles of *Brachyury***

*Brachyury*, or *T*, was the first T-box gene cloned (Herrmann et al. 1990). *Brachyury* was originally identified in mice; heterozygous mutant animals have a short tail, while homozygous individuals die *in utero* and lack a notochord and mesoderm posterior to somite 7 (Dobrovolskaïa-Zavadskaïa 1927; Gluecksohn-Schoenheimer 1938). Homologs of *Brachyury* have been isolated from other vertebrate species, including *Xenopus*, zebrafish and chicken, and these homologs display similar expression patterns during embryogenesis. The function of *Brachyury* also appears to be conserved. In particular, the zebrafish gene *no tail* (*ntl*), which was originally identified during screens for early embryonic lethal mutations (Walker and Steisinger 1983), proved to be a homologue of *Brachyury* (Schulte-Merker et al. 1994). Experiments in *Xenopus* demonstrate not only that *Brachyury* is necessary for normal mesoderm formation, but also that it is sufficient to induce mesoderm when misexpressed in prospective ectodermal tissue (Cunliffe and Smith 1992). *Brachyury* homologues were found as well in several invertebrates, with a similar role in posterior mesoderm specification. The *Drosophila* *Brachyury* homologue *brachyenteron* (*byn*) as well as the homologues of other insects are expressed mainly in the hindgut, but are not present during the formation of mesoderm (Kispert et al. 1994). Still *byn* induces the formation of one muscle/pericardial precursor cell in each parasegment. The functions of *byn* in mesodermal development of *Drosophila* may be comparable to the roles of the vertebrate *Brachyury* genes during gastrulation (Kusch and Reuter 1999). There is no clear orthologue identifiable in the model organism *Caenorhabditis elegans*, although there are even more, but highly derived T-box genes found in the complete genome sequence of this worm (The *C. elegans* Sequencing Consortium 1998). However, *C. elegans* *mab-9* mutants are defective in hindgut and male tail development showing a phenotype comparable to *Drosophila* *byn* (Woollard and Hodgkin 2000). *mab-9* has been cloned and encodes a member of the T-box family the closest relatives of *mab-9* are the human gene TBX20 and the *Drosophila* gene *H15*, that clearly belong to a different T-box subfamily than *Brachyury*.

Among bilaterians several species show a similar pattern of *Brachyury* expression. For instance, in the mollusc *Patella vulgata*, as soon as the four-fold symmetry has been broken, *Brachyury* is expressed in the D quadrant, and remains a marker of the posterior pole of the AP axis up to the end of larval development, until the whole axis is laid down (Lartillot et al. 2002). During embryogenesis of the polychaete *Platynereis dumerilii* (Arendt et al. 2001), where *Brachyury* expression is first detected at 8 h of development in vegetal cells around the closing blastopore, than in the post-gastrula larva, *Brachyury* demarcates the outline of the

closed, slit-like blastopore. From anterior to posterior, the stained cells represent developing ventral foregut (stomodaeum), ventral midline cells and developing hindgut (proctodaeum) surrounded by a few posterior mesodermal cells. In the late trochophora larva, *Brachyury* expression is restricted to the ventral portion of the developing foregut and to the developing hindgut and this closely resembles *Brachyury* expression in the ciliary larvae of basal deuterostomes (Arendt et al 2001).

On a broader scale, it appears that the mesoderm formation linked to the progressive growth of the AP axis, in an anterior to posterior sequence and through the activity of a posterior growth zone, is observed in diverse phyla across Bilateria as, for instance, insects, polychaetes and chordates. In vertebrates, the posterior pole of the developing axis is the organizer, in the form of the superior lip of the blastopore of fish and frog, or the node of amniotes, and later on, the chordo-neural hinge of the tail bud. *Brachyury* is expressed in the vertebrate organizer during the major part of embryonic development, and genetic studies in mouse and fish show that this expression is necessary for AP axis formation. This striking similarity is in favour of the hypothesis that the progressive growth of the AP axis represents a conserved developmental process which, already in Urbilateria (the last common ancestor of insects and vertebrates), would have taken the form of an organizing posterior growth zone, controlled by a genetic system involving *Brachyury*, among others (Arendt et al 2001).

*Brachyury* homologues have been identified also in cnidarians. In *Hydra vulgaris*, a simple fresh-water polyp with only two cell layers, *Brachyury* is expressed in the endoderm and is associated with the formation of the hypostome (Technau and Bode 1999). In the marine hydrozoan jellyfish *Podocoryne carnea*, however, *Brachyury* is expressed in the entocodon, a third mesoderm-like layer characteristic of the medusa bud development (Spring et al. 2002). In the anthozoan sea anemone *Nematostella vectensis*, *Brachyury* is expressed around the blastopore during the early stages of development and later in the mesenteries, invaginating septae of the polypoid sea anemone (Scholz and Technau 2003).

### **Involvement of other T-box gene subfamilies in mesoderm formation**

On the basis of genetic studies in species ranging from humans to worms three other T-box genes has emerged beside *Brachyury* that play roles in mesoderm formation. These are *Eomesodermin* (*Eomes*), *Tbx6* and *VegT*. *Eomes* was first identified in *Xenopus*, where it was shown to be involved in mesoderm development (Ryan et al. 1996). Mouse *Eomes* mutant embryos arrest soon after implantation, and tetraploid chimera analyses demonstrates that *Eomes* is required in the embryo for mesoderm formation (Russ et al. 2000). *Eomes* mutant

cells can occasionally adopt mesodermal fates in chimeras, suggesting that *Eomesodermin* may function to recruit cells into the primitive streak (Russ et al. 2000).

In the mouse embryo, *Tbx6* is expressed in the primitive streak, the paraxial mesoderm, and the tail bud (Chapman et al. 1996). In *Tbx6* mutant embryos, posterior paraxial mesoderm develops as neural tissue, suggesting that *Tbx6* is required in paraxial mesoderm to block neural development (Chapman and Papaioannou 1998). The zebrafish *tbx6* gene, although not considered the true ortholog of mouse *Tbx6* (Ruvinsky et al. 1998) is expressed very similarly to the mouse gene (Hug et al. 1997).

A third gene, *VegT*, was isolated initially in *Xenopus* (Zhang and King 1996). *VegT* transcripts are present maternally in the vegetal hemisphere of the embryo, and the gene is then zygotically activated throughout the mesoderm in a slightly broader domain than *Xbra* (Horb and Thomsen 1997). The zebrafish *spadetail* gene (named also *Tbx16*), originally identified by mutation, is a *VegT* ortholog (Griffin et al. 1998). As opposed to *VegT* in *Xenopus*, *spadetail* transcripts are not present in zebrafish eggs; *VegT/spadetail* is expressed in lateral mesoderm in both fish and frogs and in prechordal plate mesoderm in fish, but not frogs (Griffin et al., 1998; Ruvinsky et al., 1998). Homozygous *spadetail* mutant embryos lack trunk somites and later are deficient in trunk muscle; additional analyses demonstrate that *spadetail* is required in trunk somitic precursors for convergence movements and for muscle cell fate decisions (Yamamoto et al. 1998).

*Eomes* forms with *Tbr1* and *Tbx21* a highly conserved subfamily most closely related to the *Brachyury* subfamily. *Tbx6* and *VegT*, however, appear to be highly derived T-box genes with only one *Tbx6* homologue and no *VegT* equivalent in mammals

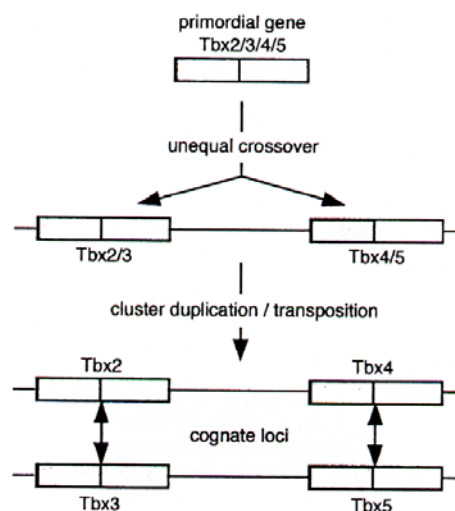
### **Developmental roles of Tbx2/3 and Tbx4/5 gene subfamilies in vertebrates**

A different T-box gene subfamily that has important roles in vertebrate limb development is formed from two cognate gene sets, *Tbx2-Tbx3* and *Tbx4-Tbx5*. Expression of these genes has been studied in the developing mouse (Gibson-Brown et al. 1996), in chick limb (Gibson-Brown et al. 1998), in *Xenopus* limb (Takabatake et al. 2000), in the developing and regenerating limb of the newt (Simon et al. 1997) and as well in zebrafish (Tamura et al. 1999). The expression patterns of these genes between species are remarkably similar, indicating a high degree of functional conservation of these cognate genes during evolution. However, differences in expression between species were also found and may reflect the morphological differences of the individual appendages.

*Tbx2* and *Tbx3* expression resolves during the early stages of forelimb/wing/fin and hind limb/wing/fin development to a similar and characteristic pattern of two proximo-distal bands of expressing cells, one in the anterior and one in the posterior margin of the limb bud mesenchyme. *Tbx4* and *Tbx5* are expressed many hours prior to overt outgrowth of the leg and wing, respectively, and appear to identify the area of cells of the lateral plate mesoderm that are destined to form the limb skeletal and cartilage elements. In addition, they remain remarkably limb-type-specific; *Tbx4* is expressed throughout the mesenchyme of the posterior limb/leg/fin buds, whereas *Tbx5* is expressed throughout the mesenchyme of the anterior limb/wing/fin buds. Two diseases, characterized by defects in limb development, are associated in humans with the *Tbx2/3* and *Tbx4/5* subfamilies, the Holt-Oram syndrome due to the *TBX5* mutations and the ulnar/mammary syndrome due to *TBX3* mutations (Papaioannou 2001).

### Hypothesis for the origin of *Tbx2/3* and *Tbx4/5* gene subfamilies

Members of the T-box gene family are found at different chromosomal locations, however, it was shown that in mouse *Tbx2* and *Tbx4* are tightly linked on chromosome 11, and the *Tbx3* and *Tbx5* are similarly linked on chromosome 5 (Agulnik et al. 1996). The clustered organization has been confirmed also in humans (Yi et al. 2000). Agulnik and co-workers have developed a model for the evolution of the *Tbx2/3/4/5* subfamily (Fig. 13; Agulnik et al. 1996).



**Fig. 13.** Origin of the cognate genes *Tbx2-Tbx3* and *Tbx4-Tbx5* in vertebrates (after Agulnik et al. 1996)

They propose that a single primordial *Tbx2/3/4/5* gene underwent an unequal crossing-over event to produce *Tbx2/3* and *Tbx4/5* progenitor genes, probably at a time close to divergence of the vertebrate and invertebrate lineages when complete genome duplications are

thought to have occurred (Holland et al. 1994). Later during vertebrate evolution, before the origin of tetrapods, a whole cluster duplication or transposition event gave rise to the *Tbx2-Tbx4* cluster on one chromosome and the *Tbx3-Tbx5* cluster on a second chromosome, generating two cognate gene pairs. In all species studied, the sequences encoding the T-domain are interrupted by either three or four introns (Wattler et al. 1998). The position of the introns combined with the primary protein sequence serve as a reliable feature for grouping similar genes into distinct subfamilies, and independently support the view that a tandem duplication and cluster duplication event created the *Tbx2-Tbx4* and *Tbx3-Tbx5* loci (Wattler et al. 1998).

### **Tbx2/3 gene subfamily in invertebrate**

Members of the *Tbx2/3* subfamily have been found in several invertebrates, but no members of the *Tbx4/5* subfamily have been identified in invertebrates, so far. In *Drosophila* the *Tbx2/3* homologue is *bifid* (*bi*; or *optomotor blind*) and it is involved in differentiation of the brain, the CNS, the wing and in patterning of adult abdominal segments, which are all ectodermally derived tissues (Kopp and Duncan 1997; Pflugfelder et al. 1992). In the nervous system, *bi* is found in both neurons and glia (Poeck et al. 1993). In the sea urchin *Lytechinus variegates*, *Lvtbx2/3* is involved in formation of the oral-aboral axis of the sea urchin embryo. In fact, an asymmetric distribution of *Lvtbx2/3* is observed in the aboral territories of all three germ layers and ectopic expression of *Lvtbx2/3* produced profound morphogenetic defects in derivatives of each germ layer with no apparent loss in specification events in those tissues (Gross et al. 2003). Thus, *Lvtbx2/3* functions as a regulator of morphogenetic movements in the aboral compartments of the ectoderm, endoderm and mesoderm. Besides these data on these two species there are no functional studies available, which does not allow formulating a hypothesis about the conservation of role during evolution between the ancestral *Tbx2/3* and derived *Tbx2*, *Tbx3* or even *Tbx4* and *Tbx5* genes.

### **Is the role of Tbx20 gene subfamily conserved in evolution?**

The *Tbx20* subfamily seems to be especially highly conserved during evolution. It was shown that *Drosophila H15* and zebrafish *tbx20* are both expressed early during heart formation, in strong support of previous work postulating that vertebrate and arthropod hearts are homologous structures with conserved regulatory mechanisms (Griffin et al. 2000). Orthologues of zebrafish *tbx20* were isolated from chick (Iio et al. 2001) and *Xenopus* (Brown et al. 2003) and the role in hearts development is conserved in both. This evolutionary view of

a conserved role for the Tbx20 subfamily is disturbed by the orthologue of *H15* in *C. elegans*, which is *mab-9* and has a function more comparable to *Brachyury* than to other T-box genes (Woollard and Hodgkin 2000).

Another T-box gene subfamily, the Tbx1/10 subfamily, had attracted attention due to the strong evidence suggesting *Tbx1* as the responsible gene for the major aortic arch defects, which characterize DiGeorge/velo-cardio-facial syndrome (Jerome and Papaioannou 2001; Lindsay et al. 2001). Among invertebrates a *Tbx1* homologue was found only in *Drosophila* (Porsch et al. 1998), but no functional information is as yet available about this gene.

# CHAPTER 1

## **Distinct expression patterns of the two T-box homologues *Brachyury* and *Tbx2/3* in the placozoan *Trichoplax adhaerens***

**Abstract** *Trichoplax adhaerens* is the only species known from the phylum Placozoa with one of the simplest metazoan body plans. In the small disc-like organism an upper and a lower epithelium can be distinguished with a less compact third cell layer in between. When *Trichoplax* was first described in 1883, the relation of these three cell layers with ectoderm, endoderm and mesoderm of higher animals was discussed. Still little is known about embryonic development of *Trichoplax*, however, genes thought to be specific for mesoderm in bilaterian animals turned out to be present in non-bilaterians already. Searching for a *Brachyury* homologue, two members of the T-box gene family were isolated from *Trichoplax*, *Brachyury* and a *Tbx2/3* homologue. The T-box genes encode a transcription factor family characterized by the DNA-binding T-box domain. T-box genes were found in all metazoans so far investigated, but in contrast to other transcription factors such as the homeobox family, T-box genes are not present in plants or fungi. The distinct expression patterns of two T-box genes in *Trichoplax* point to non-redundant functions already at the beginning of animal evolution. Since the expression patterns derived by in situ hybridization do not overlap with anatomical structures, it can be concluded that this simple animal has more than the four cell types described in the literature. This hidden complexity and the unresolved position in relation to Porifera, Cnidaria, Ctenophora and Bilateria highlight the necessity of the inclusion of *Trichoplax* in studies of comparative evolutionary and developmental biology.

**Keywords** T-box, *Brachyury*, Placozoa, *Trichoplax*, evolution



## Introduction

*Trichoplax adhaerens* is an enigmatic species described for the first time by Schulze (1883), who was not able to put it in any known phyla due to its peculiar characteristics. Already in these early studies the main features, such as the small size of only a few millimetres with a constant thickness of 0.02 mm, no symmetry axis, no organs, a dorso-ventrally three-layered body and continuous changes in shape were observed. The description of this new animal type in German had an immediate echo in the scientific community and an English report appeared in the first volume of *Science* (Minot 1883). An apparently related species was observed in Naples and named *Treptoplax reptans* by F. S. Monticelli, but since this original description it was not found again. Later, *Trichoplax* was described as the planula larva of the hydrozoan *Eleutheria krohi* (Krumbach 1907). Although this work was criticized (Schubotz 1912), the idea of *Trichoplax* as a hydrozoan larva was accepted and reported in textbooks (Hyman 1940). *Trichoplax* disappeared from the literature until 1971, when new observations suggested that *Trichoplax* is not a larva (Grell 1971; Miller 1971). *Trichoplax* is now thought to be the only species from the phylum Placozoa (Grell and Ruthmann 1991; Ruppert and Barnes 1994; Syed and Schierwater 2002).

The body structure of *Trichoplax* consists of an upper thin epithelial layer with stretched mononuclear cells, a basal epithelial layer with two cell types, column-shaped cells with cilia and round-shaped cells, called gland cells, without cilia. Between these two layers there are star-shaped cells, that form a syncytial net. *Trichoplax* move by changing their shape similar to amoebae, due to the contraction of the microtubular system of the star-shaped cells (Thiemann and Ruthmann 1989). A nervous system seems to be missing, but with an antibody against RFamide, which recognizes neuropeptides in many animal phyla, a few isolated cells can be detected (Schuchert 1993a). The feeding system is located in the ventral layer; particles already digested by enzymes produced from the gland cells can be incorporated by endocytosis (Ruthmann et al. 1986). Apparently, *Trichoplax* is missing a structured extracellular matrix and a recognizable basal lamina (Grell and Ruthmann 1991). The epithelial cells are connected by belt desmosomes in both epithelial layers and no other type of junction has been described (Ruthmann et al. 1986). Feulgen cytometry has shown that *Trichoplax* is the species with the lowest amount of DNA of all the Metazoa (Ruthmann and Wenderoth 1975).

As Schulze noticed in 1883, the three layers of *Trichoplax* could be compared to ectoderm, mesoderm, and endoderm of higher animal, but this could be proved or disproved

only by comparing developmental processes. Unfortunately, still little is known about reproduction of *Trichoplax*. In aquaria, it does it irregularly by fission. Under special circumstances it was observed to build hollow swarmer spheres, stolons and eggs developing to up to the 64 cells blastula stage (Grell 1971; Ruthmann et al. 1981; Thiemann and Ruthmann 1991). Without the natural clues to the regulation of the life cycle of this peculiar animal, it will be difficult to find the relevant life stages. We could address, however, the question whether so-called mesoderm-specific genes known from triploblasts are already present in *Trichoplax* and where and eventually when they are expressed. With that aim we decided to investigate *Brachyury* and the T-box transcription factor gene family.

Members of the T-box gene family have been found in all the animal phyla so far investigated (Papaioannou 2001; Technau 2001), but no T-box gene can be recognized in genomes of fungi, plants or parasitic protists. The main feature of the protein is the conserved region of about 180 amino acids, called T-box domain that is characteristic of the family. The T-box domain is able to recognize half of a palindromic binding site in DNA (Müller and Herrmann 1997). There is still some confusion about the classification of the 5 to 8 subfamilies and the nomenclature in different species (Papaioannou 2001; Ruvinsky et al. 2000), but at least comparisons of amphioxus and vertebrates suggests a conservative evolution of this gene family. Among the T-box family members the best studied is the founding member *Brachyury* or *T*, which is expressed during the gastrulation in the organizer and later along the AP axis of all the bilaterians that have been investigated so far (Arendt et al. 2001; Lartillot et al. 2002; Takada et al. 2002; Technau 2001). In chordates, *Brachyury* is involved in notochord formation (Amacher et al. 2002). *Brachyury* homologues have been identified also in cnidarians, which should be distinguished from bilaterians by the absence of mesoderm. In *Hydra vulgaris*, a simple fresh-water polyp with only two cell layers, *Brachyury* is expressed in the endoderm and is associated with the formation of the hypostome (Technau and Bode 1999). In the marine hydrozoan jellyfish *Podocoryne carnea*, however, *Brachyury* is expressed in the entocodon, a third mesoderm-like layer characteristic of the medusa bud development (Spring et al. 2002). In the anthozoan sea anemone *Nematostella vectensis*, *Brachyury* is expressed around the blastopore during the early stages of development and later in the mesenteries, invaginating septae of the polypoid sea anemone (Scholz and Technau 2003).

*Trichoplax* and Placozoa are often missing in phylogenetic studies and discussions on the relationship of animal phyla. In some studies based on 18S rRNA (Wainright et al. 1993; Philippe et al. 1994; Bridge et al. 1995; Odorico and Miller 1997; Collins 1998) or 28S rRNA

sequences (Christen et al. 1991) *Trichoplax* is included and appears to belong to the cnidarians or as sister-group of cnidarians, ctenophores or bilaterians. While this manuscript was in preparation, a further discussion of *Trichoplax* phylogeny based on mitochondrial rRNA sequence and structure confirming that Placozoa are not derived Cnidaria appeared (Ender and Schierwater 2003). Only two partial sequences of protein coding genes have been published so far; one a fragment of the homeobox gene *Trox-2* (Schierwater and Kuhn 1998), a homologue of the ParaHox gene *Gsx* and the other one a fragment of *Pax-B* (Gröger et al. 2000), a homologue of Pax2/5/8, both known from a wide variety of animals including non-bilaterians.

With the first in situ hybridizations of *Trichoplax* we present a molecular approach to the description of this elusive animal. Genes such as *Brachyury*, *Tbx2/3*, *elongation factor-1 alpha (EF1a)* or actin are highly conserved and often more similar to human sequences than to invertebrate model organisms. Not so surprisingly, there are also genes such as *Secp1* coding for a putative small secreted protein without recognizable sequence similarities. But together, they provide evidence that the gene repertoire of this simple animal is comparable to those of higher animals and the complex expression patterns seen with *Brachyury*, *Tbx2/3* and *Secp1* already indicate that *Trichoplax* consists of more than the four cell types described in the literature.

## Materials and methods

### Animals

*Trichoplax adhaerens* appear regularly on the glass walls of marine aquaria maintained to culture the hydrozoan jellyfish *Podocoryne carnea*. Animals are removed from the glass walls of the aquaria by a blow with a Pasteur pipette and transferred to a clean glass dish. They were left to adhere on the glass surface for about 2-3 hours and washed 4 times with Millipore filtered sea water. Individual specimens were selected for longer cultures, RNA and DNA extractions or in situ hybridization experiments.

### Molecular cloning and sequence comparison

Genomic DNA and total RNA were isolated from *Trichoplax* by using the TriReagent (Molecular Research Center) according to the manufacturer's recommendations. First strand cDNA was synthesized with the anchored oligo (dT) primer XT20V (5'-GGC AGG TCC TCG TTG ACT CGA GAC GT(20)(AGC)-3') by using the SMART RACE cDNA Amplification Kit (Clontech). With the Smart cDNA we were able to conduct homology PCR, 3' and 5' RACE to obtain the full length of *Brachyury*, *Tbx2/3* and *EF1a* homologues of *Trichoplax*. A 113 bp *Brachyury* fragment was amplified with the set of degenerated primers TF1 and TR1 (Spring et al. 2002), followed by TF2 (5'-TT(CT) GG(AGCT) CA(CT) TGG ATG-3') and TR2 (Spring et al. 2002); a 308 bp *Tbx2/3* fragment was amplified with the set of degenerated primers TF1 and TR1 followed by TF1 and TR2. For each PCR round standard conditions were used, except that the annealing temperature was 37°C for the first PCR round of 20 cycles. In the second PCR round the annealing temperatures were 37°C for 10 cycles and 45°C for 35 cycles. A 239 bp *EF1a* fragment was amplified with the degenerated primers uEF (5'-AAG TCA GT(AGCT) GA(AG) ATG CA(CT) CA(CT) GA-3') and uER (5'-GCA AT(AG) TG(AGCT) GC(AGCT) GT(AG) TG(AG) CA(AG) TC-3'); for the PCR standard condition with two different annealing temperature, 40°C for 10 cycles and 50°C for 40 cycles, were used. PCR products of the expected size were gel purified with a Qiaquick column (Qiagen), subcloned in the pCRII-TOPO vector (TOPO TA cloning Dual Promoter kit, Invitrogen) and sequenced on an ABI PRISM 310 genetic analyser (Applied Biosystems). Based on the sequences gene-specific primers were designed to carry out the 5' and 3' RACE as described (Müller et al. 1999; Yanze et al. 2001). Clones with the complete coding

sequences were isolated for *Brachyury* (1872 bp), *Tbx2/3* (2439 bp) and *EF1a* (1635 bp). The full length coding sequences of *Secp1* (737 bp) and an *actin* homologue (1351 bp) were found by random screening of clones, that were obtained by using as insert the PCR product on SMART cDNA amplified with the primers Nup1 (5'-AAG CAG TGG TAT CAA CGC AGA G-3') and X1 (5'-GGC AGG TCC TCG TTG ACT CG-3') subcloned in the pCRII-TOPO vector. The *Trichoplax Brachyury*, *Tbx2/3*, *Secp1*, *actin* and *EF1a* sequences described here have been submitted to the DNA databases with the accession numbers AJ549221-5, respectively. *Podocoryne Efla* is available under AJ549292.

Nucleotide and deduced amino acid sequences were analysed by using the GCG software package. BLAST searches (Altschul et al. 1997) were performed on the blast network service at the NCBI (<http://www.ncbi.nlm.nih.gov>). Sequences from representative species were retrieved from BLAST searches with *Trichoplax* and human homologues; a detailed list of accession numbers and the alignments are available on request. Multiple sequence alignments and phylogenetic trees based on the neighbour-joining method were generated with Clustal X (Jeanmougin et al. 1998) and maximum likelihood trees calculated with TREE-PUZZLE (Schmidt et al. 2002).

### In situ hybridisation

Whole mount in situ hybridisation experiments were carried out with adaptations to the protocol described for *Podocoryne carnea* (Yanze et al. 2001; Spring et al. 2002). Animals were fixed overnight at 4°C in Lavdowsky fixative supplemented with 0.2% glutaraldehyde; washing 3 times for 30 min in PBST (0.1% Tween20 in PBS); stepwise increase of the concentration from 15%, 30%, 45%, 60%, 75%, 90% and 100% of step-HB (5 x SSC; 50% formamide; 50 mg/l heparin; 0.1% Tween20) in 5 x SSC, shaking at each step 10 min at room temperature (RT); prehybridisation in hybridisation buffer (HB) (5 x SSC; 50% formamide; 50 mg/l heparin; 100 mg/l tRNA from *E. coli* (Sigma type XXI R4251); 0.1% Tween20) for 1h at 50°C; hybridisation with DIG-labelled antisense probe (approximately 20 ng/ml HB) over night at 50°C; washing 2 times for 20 min at 50°C with WS1 (5 x SSC; 50% formamide; 0.1% Tween20), WS2 (2 x SSC; 50% formamide; 0.1% Tween20), WS3 (2 x SSC; 0.1% Tween20) and WS4 (0.2 x SSC; 0.1% Tween20); incubation in PBST for 5 min, in blocking solution (1% blocking reagent; 100 mM maleic acid pH 7.5; 150 mM NaCl) for 1h at RT and with anti-DIG-AP Fab fragments 1:5000 in blocking solution overnight at 4°C; washing 3 times with PBST and equilibration in TMNT (100 mM Tris-HCl, pH 8; 50 mM MgCl<sub>2</sub> ; 100

mM NaCl; 0.1% Tween20) for 30 min at RT. Detection was done by incubation of the specimens in TMN (100 mM Tris-HCl, pH 8; 50 mM MgCl<sub>2</sub> ; 100 mM NaCl) with 340 mg/l NBT and 175 mg/l BCIP. When the chromogenic reaction was ready (2 to 4 hours), it was stopped by washing the specimens in PBST. DIG-RNA-labelled probes were synthesized according to the manufacture's recommendation (DIG-RNA-labelling, Roche). The DNA fragments for *Brachyury* and *Tbx2/3* probes were amplified with primers designed outside the T-box domain as following: TBra F6 (5'-GTC GAC ATC AAC GCC ATT AGG-3') and TBra R4 (5'-TAC TTT GCT ACT TGT TTG ATA-3'), TTbx F4 (5'-GAC ATG TGC AGA AGT AAC TAA AGG-3') and TTbx R7 (5'-TTC TTG AGT ACG GAA TTT TCT CG-3'). These fragments were subcloned in the pCRII-TOPO vector (TOPO TA cloning Dual Promoter kit, Invitrogen) and used as template for the probe synthesis. The equivalent was done for *Secp1* using the primer set TSec1 F1 (5'-TAA CTG TAA GGA CTG AAA AAT-3') and TSec1 R1 (5'-TTC GAA ATC CTT ATC GTG AAA C-3').

After in situ hybridisation some samples were processed for sectioning: postfixation in 4% paraformaldehyde at RT overnight; dehydration by incubation in ethanol 30%, 50%, 80%, 90%, 95% and 100% for 1 h at RT, respectively; incubation in 50% ethanol/histo-clear overnight at RT; incubation in histo-clear for 20 min at 58°C; transfer to 50% histo-clear/parablast for 1 h at 58°C; embedding twice in parablast, the first time 1h at 58°C and the second time overnight at 58°C; hardening in a 14°C water bath for 10 min; mounting on block carrier; cutting at microtome in slices of 10 µm (Microm HM 360); stretching sections on water warmed at 37°C for 1 h; mounting sections on slides previously washed in 50% ethanol/ether and dried for 1 h; washing the slides in histo-clear for 10 min at RT; mounting permanently with DPX (Fluka); covering with cover slides; hardening overnight at RT.

## Results and discussion

### Highly conserved *Brachyury* and *Tbx2/3* homologues from *Trichoplax*

Two distinct T-box family members were isolated from *Trichoplax* by homology PCR. One is clearly a member of the *Brachyury* subfamily and the other a member of the *Tbx2/3* subfamily that was known only from bilaterian animals. *Trichoplax Brachyury* is 70-80% identical to *Brachyury* subfamily members and less than 55% identical to other subfamilies in the T-box domain while little similarity can be detected outside of the T-box domain (Fig. 1A). The conservation of the amino acid residues involved in DNA binding and dimerization suggest that also *Trichoplax Brachyury* binds the typical consensus sequence as a dimer as determined by crystallographic studies (Müller and Herrmann 1997). Transactivation and repression domains were defined in the C-terminal half of mouse *Brachyury* (Kispert et al. 1995). Although there is little sequence conservation in the C-terminal part of *Brachyury* from *Trichoplax* to *Drosophila* or vertebrates, there are at least three conserved motifs containing proline, tyrosine or tryptophane (PY/W) in human and *Trichoplax Brachyury* (Fig. 1A).

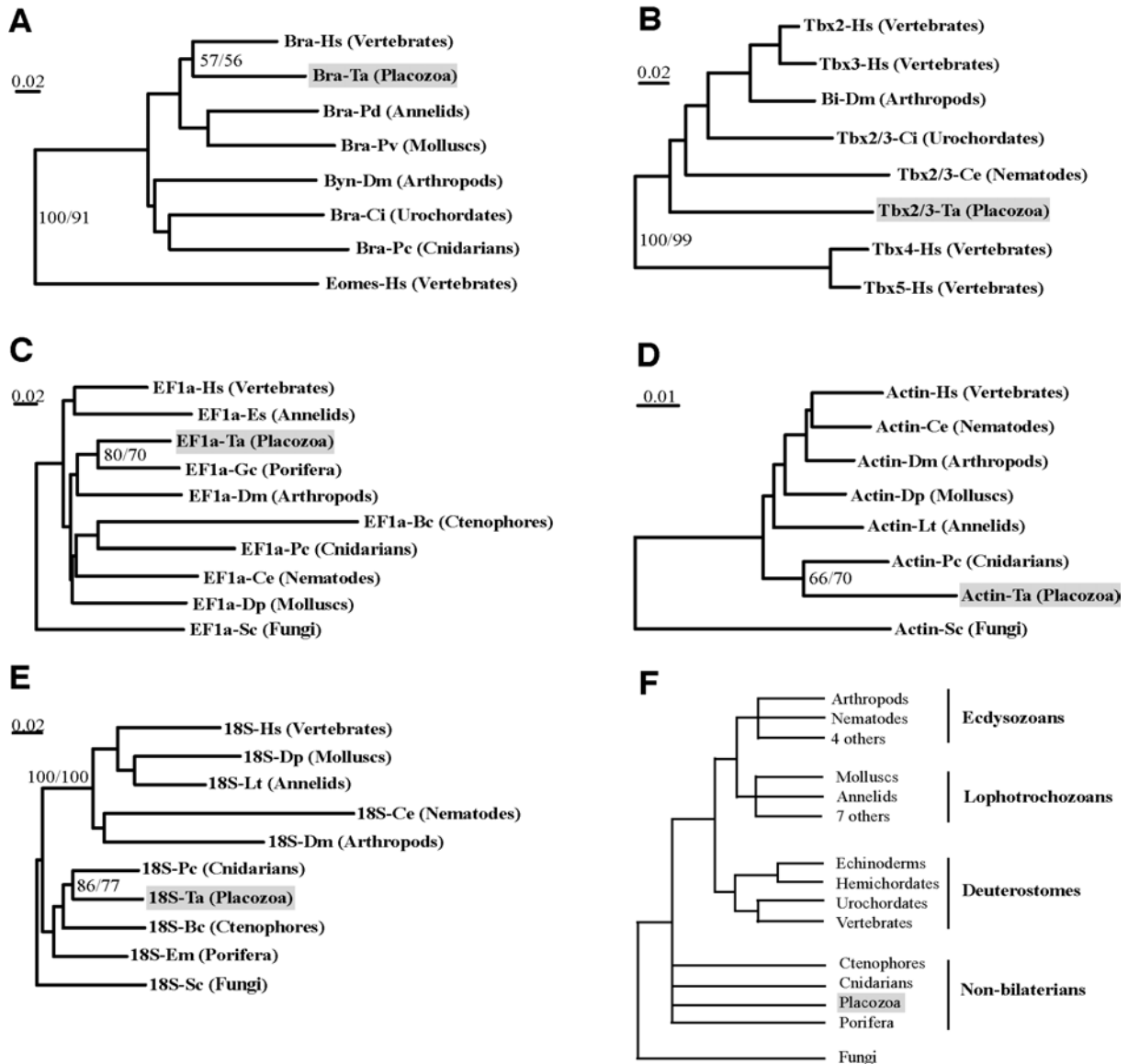
*Trichoplax Tbx2/3* is 60-70% identical to *Tbx2/3* subfamily members and up to 60% identical to *Tbx4/5* and less related subfamilies (Fig. 1B). The complexity of the T-box family in higher animals is best seen in the comparison of amphioxus family members to the corresponding vertebrate and invertebrate homologues (Ruvinsky et al. 2000). In this analysis the *Tbx2/3* subfamily is most related to the *Tbx4/5* subfamily specific for vertebrates and amphioxus and less related to the *Tbx1/10*, *Tbx15/18/22* or *Tbx20* subfamilies. These five subfamilies seem to be as similar as the *Brachyury* and *Eomes* subfamilies, while the *Tbx6* subfamily appears to be more diverged.

Comparison of the two known T-box domain structures of *Xenopus Brachyury* and human *TBX3* explains some of the differences observed with different subfamilies (Coll et al. 2002). *TBX3* does not need to dimerize to bind DNA and consistent with this, some of the residues required for *Brachyury* dimerization are not conserved in *TBX3*, and three out of four of these residues are identical in *Trichoplax* and human *Tbx2/3* family members (Fig. 1B). All residues involved in DNA binding are identical between *Trichoplax* and human *Tbx2/3* family members. The C-terminal part does not show any significant similarity with other members of the subfamily, although this region should contain the activator and repressor activities of the protein (Paxton et al. 2002). There is residual sequence similarity at





apparently high conservation of *Brachyury* from *Trichoplax* to humans there is no clear orthologue identifiable in the model organism *Caenorhabditis elegans*, although there are even more, but highly derived T-box genes found in the complete genome sequence of this worm (The *C. elegans* Sequencing Consortium 1998).



**Fig. 2.** Phylogenetic trees based on neighbour-joining and maximum likelihood methods generated with Clustal X and TREE-PUZZLE. (A) *Trichoplax* and human Brachyury form a weak clade within the well-supported *Brachyury* family. (B) *Trichoplax* Tbx2/3 is at the base of all known Tbx2/3 family members but clearly separated from other sub-families. (C) *Trichoplax* EF1a groups with a sponge, but not with other non-bilaterians. (D) *Trichoplax* and cnidarian actins form a clade at the base of animal actins. (E) *Trichoplax* 18S rRNA forms a clade with cnidarians within other non-bilaterians. (F) Simplified phylogeny of the animal kingdom based on 18S rRNA sequences (Holland, 1999). Numbers on the branches indicate the percentage of 1000 bootstrap replicates that support the topology shown calculated with Clustal X or TREE-PUZZLE. The bar represents the number of substitutions per site. Human eomesodermin was included as outgroup for the Brachyury family and TBX4 and TBX5 for the Tbx2/3 family, respectively. For the other genes, yeast homologues were used for rooting the trees. Abbreviations: Hs, *Homo sapiens*; Ta, *Trichoplax adhaerens*; Pd, *Platynereis dumerilii*; Pv, *Patella vulgata*; Dm, *Drosophila melanogaster*; Ci, *Ciona intestinalis*; Pc, *Podocoryne carnea*; Ce, *Caenorhabditis elegans*; Es, *Enchytraeus sp.*; Gc, *Geodia cydonium*; Bc, *Beroe cucumis*; Dp, *Dreissena polymorpha*; Sc, *Saccharomyces cerevisiae*; Lt, *Lumbricus terrestris*; Em, *Ephydatia muelleri*.

The *Tbx2/3* tree shows that the *Trichoplax* homologue belongs to this subfamily, even to the exclusion of the *Tbx4/5* subfamily found in vertebrates. An ancestral tandem duplication is thought to have occurred before the subsequent cluster duplication in vertebrates (Agulnik et al. 1996; Ruvinsky et al. 2000). Thus, a *Tbx2/3-Tbx4/5* like gene pair was probably already present close to the Hox cluster in an amphioxus-like animal at the beginning of vertebrate evolution. Now, the *TBX2-TBX4* gene pair is linked to the HoxB cluster and *TBX3-TBX5* to the HoxC cluster in humans. No *Tbx4/5* gene was found in *Trichoplax* or any other invertebrate below amphioxus indicating that the tandem duplication was an evolutionary late event.

On the basis of morphological features the placozoan *Trichoplax* has always been thought to be a basal animal and was often placed between sponges and the other animals. In the last decade, with the rise of molecular phylogenetics, several analyses were carried out with the *18S rRNA* gene to determine the relationship of animal phyla. All possible combinations of grouping *Trichoplax* with other non-bilaterians were reported. *Trichoplax* has been suggested to be a sister group of cnidarians (Wainright et al. 1993; Philippe et al. 1994) or it has been suggested to belong to the cnidarians itself (Bridge et al. 1995). Moreover, *Trichoplax* has been grouped with ctenophores, as a sister group of cnidarians (Odorico and Miller 1997) and most recently as a sister group of the bilaterians, to the exclusion of cnidarians, ctenophores and sponges (Collins 1998; Collins 2002). The best way to summarize the knowledge on the phylogenetic position of non-bilaterians and *Trichoplax* is still an unresolved tree (Fig. 2F; Holland 1999).

Since phylogenetic analysis of Brachyury does not conform to any reasonable expectation and for *Tbx2/3* there are no other non-bilaterian sequences available, we performed comparable analyses with some well-studied reference genes such as *EF1a*, *actin* and *18S rRNA*. *Trichoplax EF1a*, *actin* and the ends of the previously published *18S rRNA* (Wainright et al. 1993) were sequenced in this study for quality control reasons. Unfortunately, in spite of many phylogenetic studies with these genes, there are no complete actin sequences from sponges or ctenophores available. Representative species were selected from the ecdysozoan and lophotrochozoan subgroup of the protostomes and from deuterostomes to fit best with the available data from T-box genes.

In the *EF1a* tree *Trichoplax* groups with a sponge (Fig. 2C). It could make sense to think of Porifera and Placozoa as the simplest multicellular animals as a group, but this group should then be at the base of the tree and not a sister-group to arthropods in the middle of badly sorted bilaterian and non-bilaterian phyla. In the actin tree *Trichoplax* and cnidarians

are a sister-group to the bilaterians, but the lack of actin sequences from sponges and ctenophores does not allow resolving the relationship within non-bilaterian phyla (Fig. 2D). In an *18S rRNA* with a comparable set of species *Trichoplax* is placed in a group with the other non-bilaterians (Fig. 2E). In contrast to more extensive studies (Collins 1998; 2002), however, *Trichoplax* appears here as a sister-group to cnidarians and not to bilaterians. This indicates that the choice of species and number of different species within a phylum are important, which should also be considered when collecting additional data for T-box genes for improved phylogenetic studies.

The data that are presently available indicate that *Brachyury* and *EF1a* evolved more irregularly than *Tbx2/3*, *actin* and *18S rRNA*. Whether this is correlated with different functional conservation of *Brachyury* and *Tbx2/3* is difficult to judge when considering the apparent differences in reference genes such as *EF1a* and *actin*.

#### Distinct expression of *Brachyury* and *Tbx2/3*

In situ hybridisation revealed that *Trichoplax Brachyury* is expressed irregularly in individual animals in a few cells or groups of cells (Fig. 3A-D). While some animals lack expression completely, in larger animals *Brachyury* expression is localized to the edge of potential outgrowth zones. It was not possible yet to resolve the position of the signals in the three cell layers in sections.

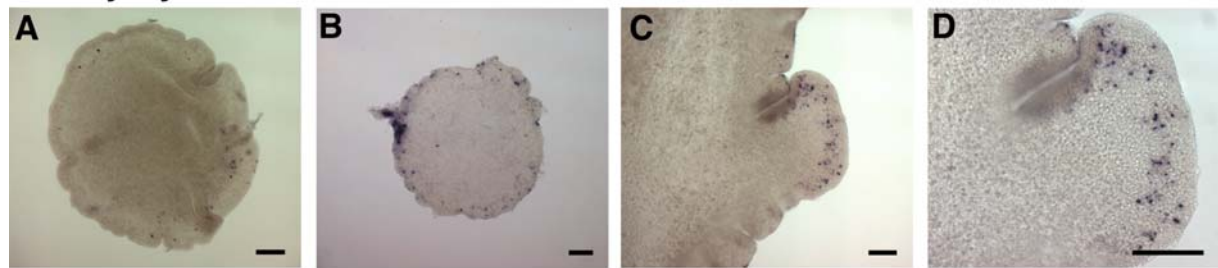
*Trichoplax Tbx2/3* is expressed in a very different pattern with strong signals at the periphery of attached animals (Fig. 3E, F). In a specimen that was not attached completely to the substrate, a differential *Tbx2/3* expression between the attached part and the floating part can be distinguished (Fig. 3G, M). The part of the body that was attached showed a strong signal in both the ventral and dorsal layers while the floating region had a signal only in the thin ventral layer (Fig. 3H).

Reference genes such as *actin* give a strong ubiquitous signal in comparable experiments (not shown). A more useful control was found with the gene coding for the putative small secreted protein *Secp1*, which gave a clear pattern of expression by in situ hybridisation, a well defined ring shaped signal that covered the edge of the animal (Fig. 3I). Analyses of the 150 amino acid sequence of *Secp1* reveals an N-terminal signal peptide of 19 residues with a potential signal cleavage site. The mature protein of 131 residues consists of

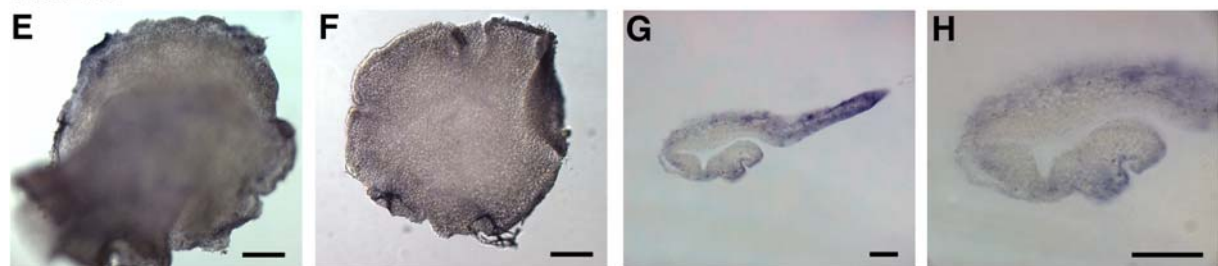
76 charged amino acids and shows no significant sequence similarity to known proteins. In sections, *Secp1* appears to be expressed in all three body layers (Fig. 3J).

Only four types of cells have been described in *Trichoplax* (Fig. 3K, L). From the expression patterns of *Trichoplax Brachyury*, *Tbx2/3* and *Secp1* we can conclude that there are at least three different kinds of cells with different gene expression profiles, which do not coincide with the morphologically distinguishable cell types. This hidden anatomical complexity and the clear presence of homologues of the Brachyury and *Tbx2/3* subfamilies already in *Trichoplax*, suggest that the common ancestor of bilaterians and more primitive animals was more complex than previously anticipated.

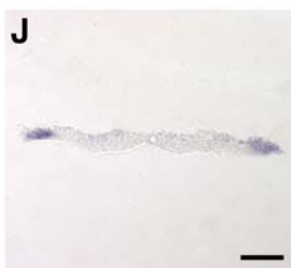
### **Brachyury**



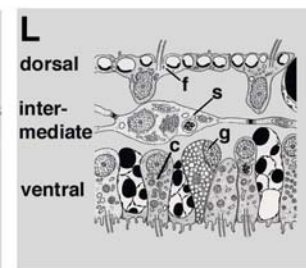
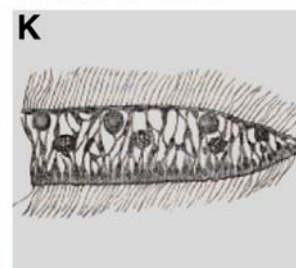
### **Tbx2/3**



### **Secp1**



### **Schulze 1883**



**Fig. 3.** Expression analysis by in situ hybridization of *Trichoplax* genes. *Brachyury* is expressed in a few isolated cells, mainly near the edge of the animal; (A, B) small round animals, (C) larger bifurcated animal with potential outgrowth zone, (D) higher magnification of C. (E, F) *Tbx2/3* is expressed at the periphery of animals, but not in all cells, (G) section of E shows the differential expression between the body region attached at the substrate and the floating body region, (H) higher magnification of G, in the floating region the *Tbx2/3* gene is mainly expressed in cells of the dorsal layer. (I) The gene for the putative secreted protein *Secp1* is expressed in a ring-shaped pattern near the edge of the animal. (J) Section of I, *Secp1* appears to be expressed in all three layers along the ring. (K) Sketch of the body structure of *Trichoplax* at an enlargement of 300 x (Schulze 1883). (L) Scheme according to Grell and Ruthmann (1991) with flat epithelial cells (f) in the dorsal layer, columnar epithelial (c) and non-flagellated gland cell (g) in the ventral layer and star-shaped cells (s) in the intermediate layer. The bars correspond to 0.3 mm.

## CHAPTER 2

### **T-box genes in basal metazoans: further indications for the monophyletic origin of metazoans**

**Abstract** Metazoan diversification involved several evolutionary steps in which new molecular, cellular and developmental characters arose from previous forms of life. Understanding these changes means to explain how body plan complexity increased during evolution. Such issues require knowledge of the evolutionary pathway that metazoans followed from unicellular organisms and from non-bilaterian to bilaterian organization. Moreover, whether a single organism or different organisms gave rise to all metazoans is still a matter of discussions. We therefore investigated all four non-bilaterian phyla by isolating T-box genes from sponges, placozoans, ctenophores and cnidarians. Up to now no T-box gene has been isolated outside metazoans. The T-box genes found in non-bilaterians belong to distinct T-box subfamilies recognizable in bilaterians, which suggests that metazoans originated from an ancestor, which had already more than one T-box gene. The evolutionary relationship of non-bilaterian phyla was also tested by sequencing *elongation factor-1-alpha* (*EF1a*) genes from a sponge and a ctenophore, respectively. Besides confirming the monophyletic origin of metazoans, the phylogenetic study conducted with EF1a does not solve the evolutionary relationship among non-bilaterians and further studies are needed.

**Key words:** T-box, *Brachyury*, *Tbx2/3*, *Tbx4/5*, *Elongation factor-1-alpha*, *aristaless*, evolution, metazoans, bilaterians, non-bilaterians, *Axinella verrucosa*, *Pleurobrachia pileus*, *Podocoryne carnea*, cnidarians, sponges, ctenophores, Cnidaria, Porifera, Ctenophora.

## Introduction

In view of similarities in the organization of the epithelial and mesenchymal tissues and connective tissue, including collagen and other molecules (Müller et al. 2001), between sponges and cnidarians as well as other animals, the arguments for the monophyly of metazoans are very strong. However, analyses of the origin of metazoans using methods of molecular phylogeny with rRNA genes as a phylogenetic index have recently revived the controversy as whether all metazoans can be viewed as successive offshoots within a single monophyletic unit, or different processes of cell aggregation led to parallel radiation and different patterns of body plan organization.

The early publications of phylogenetic trees based on partial sequences of animal *18S rRNA* genes (Field et al. 1988) and *28S rRNA* genes (Christen et al. 1991) suggested the idea that the two animal subkingdoms, Radiata and Bilateria, originated independently from different flagellate protozoan ancestors. Recently these conclusions have been widely criticized by several studies based on complete gene sequences. Evolutionary trees have been build with *18S rRNA* genes and with the protein coding genes *elongation factors-1-alpha* and *Hsp70* (Borchiellini et al. 1998; Kim et al. 1999; Kobayashi et al. 1996; Peterson and Eernisse 2001; Wainright et al. 1993), all these studies support the monophyly of the animal kingdom including sponges, although the bootstrap value for monophyly is very weak.

The monophyletic origin of the animal kingdom is supported as well by the discovery of a receptor tyrosine kinase (RTK) from a sponge (Schacke et al. 1994), suggesting that the evolution of RTKs for cell-cell signalling could have played a role in the evolution of early metazoans. Moreover, the discovery of a RTK in the choanoflagellate *Monosiga brevicollis* demonstrates that at least one family of metazoan signalling molecules was present before the origin of multicellular animals (King and Carroll 2001). Phylogenetic and molecular genetic evidence (Snell et al. 2001; Wainright et al. 1993), together with the ultrastructural similarities between choanoflagellates and sponge choanocytes (Hibberd 1975), suggest that choanoflagellates may best represent the eukaryotic form of life immediately anterior to the evolution of multicellular animals.

Although some studies have been done to confirm the monophyletic origin of metazoans, further comparisons of the expressed genes in non-bilaterians could infer a better view on the origin of metazoans. Therefore, T-box genes were isolated from all four non-bilaterian phyla, the sponge *Axinella verrucosa*, the placozoan *Trichoplax adhaerens*, the ctenophore *Pleurobrachia pileus* and the cnidarian *Podocoryne carnea*.

T-box genes encode transcription factors that share a conserved DNA-binding domain of 180–200 amino acids, called the T-box domain. In the past ten years cloning and expression analysis of an increasing number of T-box genes from all kind of metazoans brought progress in understanding the various roles that T-box genes have in animal development. It has been demonstrated that they are involved in the cellular movement during gastrulation, in the formation of the mesoderm and its derivatives and several of them have been linked to heritable human diseases. Still, no T-box gene can be recognized in genomes of fungi, plants or parasitic protists (Papaioannou 2001). The T-box domain is able to recognize half of a palindromic binding site in DNA (Müller and Herrmann 1997). On the basis of sequence homology several T-box subfamilies can be recognized. There is still some confusion about the classification of the 5 to 8 subfamilies and the nomenclature in different species (Papaioannou 2001; Ruvinsky et al. 2000), but at least comparisons of amphioxus and vertebrates suggests a conservative evolution of this gene family.

Most of the T-box genes isolated in this study are clearly belonging to T-box gene subfamilies recognizable in bilaterians, suggesting the presence of T-box genes and their subdivision in subfamilies already in the ancestral metazoans (Urmotazoa). Among the T-box genes isolated, we found in the cnidarian *Podocoryne* the first member of the Tbx4/5 subfamily of all invertebrates. We investigated by in situ hybridisation the expression of *Podocoryne Tbx4/5*. Recently, another T-box gene, *Brachyury*, was isolated and then characterized by in situ hybridisation in *Podocoryne* (Spring et al. 2002). The two genes show a distinct expression pattern, confirming the distinct roles for T-box genes already in non-bilaterians.

In addition we isolated *elongation factor -1-alpha (EF1a)* genes from *Axinella Trichoplax*, *Pleurobrachia* and *Podocoryne* as referent genes to compare to the phylogenetic analyses with T-box genes. The evolutionary history was already inferred through the amino acids sequences of EF1a (Kobayashy et al. 1996), we repeated this analyses with the new sequences containing representatives of all four non-bilaterians phyla, but without improving much the previous results.

## Materials and methods

### Animals

A body fragment of *Axinella verrucosa* (Porifera, Demospongiae) was brought in our laboratory from the Marine Biology Station of Roscoff. Cells from *Axinella* were isolated by cutting in little pieces the body fragment, then these pieces were washed with Millipore filtered sea and successively cells were scratched out with forceps. The cells were separated from the rest by a few seconds of centrifugation at 4°C, then the supernatant was removed and the cells treated for RNA and DNA extractions.

*Pleurobrachia pileus* (Ctenophora, Cydippida) was fished in Roscoff together with plankton. We were able to keep some of those specimens long enough to produce larval stages which were selected, washed with Millipore filtered sea water and then processed for RNA and DNA extractions. The adult animals survived about 10 days in our laboratory and no larval stages reached the adult stage.

Colonies of *Podocoryne carnea* (Cnidaria, Hydrozoa) were cultured in artificial seawater, the different life stages were collected and treated for RNA and DNA extraction as described elsewhere (Spring et al. 2002).

### Molecular cloning and sequence comparison

Genomic DNA and total RNA were isolated from *Axinella*, *Pleurobrachia* and *Podocoryne* by using the TriReagent (Molecular Research Center) according to the manufacturer's recommendations. First strand cDNAs were synthesized for each species with the anchored oligo (dT) primer XT20V (5'-GGC AGG TCC TCG TTG ACT CGA GAC GT(20)(AGC)-3') by using the SMART RACE cDNA Amplification Kit (Clontech). With the above-mentioned cDNAs we isolated by homology PCR, 3' and 5' RACE the full length of the following genes: *Tbx1/15/20* and *EF1a* from *Axinella*, *Brachyury*, *Tbx2/3* and *EF1a* from *Pleurobrachia* and *Tbx4/5* from *Podocoryne*. A 272 bp *Axinella Tbx1/15/20* fragment was amplified with the set of degenerated primers TF1 and TR1 (Spring et al. 2002), followed by TF2 (5'-TT(CT) GG(AGCT) CA(CT) TGG ATG-3') and TR1 (Spring et al. 2002), for each PCR round standard conditions were used, except that the annealing temperature was 37°C for the first PCR round of 20 cycles and in the second PCR round the annealing temperatures was 37°C for 10 cycles and 45°C for 35 cycles. A 113 bp *Pleurobrachia Brachyury* fragment



was amplified with the set of degenerated primers TF1 and TR1, followed by TF2 and TR2 by the same methodology described for *Axinella Tbx1/15/20*; a 185 bp *Pleurobrachia Tbx2/3* fragment was amplified with the set of degenerated primers TF3 (AAT TCA ATG CA(CT) AA (AG) TA(CT) (GC)A(AG) CC) and TR3 (CTG AAT CC(CT) TT(AGCT) GC(AG) AA(AGCT) GG(AG) TT); PCR was performed under standard conditions, except that the annealing temperatures, that were 37°C for 10 cycles and 45°C for 35 cycles. Two T-box gene fragments of 664 bp and 484 bp were isolated from *Podocoryne* genomic DNA with the set of primers TF3 and TR3 by performing PCR under the condition described for *Pleurobrachia Tbx2/3*. Both fragments contain an intron of 473 bp and 297 bp, respectively. The first fragment was present as well in *Podocoryne* cDNA and turned out to be a *Tbx4/5* homologue the other was never detected in cDNA and has the same degree of homology with *Tbx1* and *Tbx2/3* members. A 239 bp *EF1a* fragment was amplified from both *Pleurobrachia* and *Axinella* with the set of degenerated primers uEF (5'-AAG TCA GT(AGCT) GA(AG) ATG CA(CT) CA(CT) GA-3') and uER (5'-GCA AT(AG) TG(AGCT) GC(AGCT) GT(AG) TG(AG) CA(AG) TC-3'); for the PCR standard condition with two different annealing temperature, 40°C for 10 cycles and 50°C for 40 cycles, were used. PCR products of the expected size were gel purified with a Qiaquick column (Qiagen), subcloned in the pCRII-TOPO vector (TOPO TA cloning Dual Promoter kit, Invitrogen) and sequenced on an ABI PRISM 310 genetic analyser (Applied Biosystems). Based on the sequences gene-specific primers were designed to carry out the 5' and 3' RACE as described (Müller et al. 1999; Yanze et al. 2001). Clones from *Axinella* with the complete coding sequences were isolated for *Tbx1/15/20* (2607 bp) and *EF1a* (1648 bp); clones from *Pleurobrachia* with the complete coding sequences were isolated for *Brachyury* (1731 bp), *Tbx2/3* (2523 bp) and *EF1a* (1512 bp); a clone from *Podocoryne* with the complete coding sequences was isolated for *Tbx4/5* (1646 bp). The *Axinella Tbx1/15/20* and *EF1a*, the *Pleurobrachia Brachyury*, *Tbx2/3* and *EF1a* and the *Podocoryne Tbx4/5* and *Ef1a* sequences described here have been submitted to the DNA databases.

Nucleotide and deduced amino acid sequences were analysed by using the GCG software package. BLAST searches (Altschul et al. 1997) were performed on the blast network service at the NCBI (<http://www.ncbi.nlm.nih.gov>). Multiple sequence alignments and phylogenetic trees based on the neighbour-joining method were generated with Clustal X (Jeanmougin et al. 1998) and maximum likelihood trees calculated with TREE-PUZZLE (Schmidt et al. 2002).

### Whole-mount in situ hybridisation for *Podocoryne carnea*

Whole mount in situ hybridisation experiments were carried out with adaptations to the protocol described for *Podocoryne* elsewhere (Müller et al. 2003). After staining, specimens were extensively washed in PBST and investigated as whole mounts or as 10 µm sections of paraplast (Sigma) embedded specimens (Spring et al. 2002). The homeobox gene *Aristaless* (*Arx*) was used as control. The gene sequence of *Arx* was obtained by screening with the two hybrids system a cDNA library of *Podocoryne*, as described elsewhere (Müller et al. 2003).

DIG-RNA-labelled probes were synthesized according to the manufacture's recommendation (DIG-RNA-labelling, Roche). The DNA fragment for the *Podocoryne Tbx4/5* probe was amplified with primers designed outside the T-box domain as following: PcTbx2 F2 (5'-TCC GCA GAA GCA TTA CGT AAG AG-3') and PcTbx2 R4 (5'-AAA GTC AAT AGC TTC TTC ATA TTA C-3'); while the DNA fragment for the *Podocoryne Arx* probe was amplified with primers designed outside the homodomain as following: Arx F2 (5'-CGT CCG CAA AAC AAC GCA TCC TAC-3') and Arx R3 (5'-GCG ATG TAT GAG ACG ATT TGA CG-3').

## Results and discussion

### Conserved T-box genes from non-bilaterians

One T-box family member was isolated from the sponge *Axinella verrucosa*. Although this gene does not clearly belong to any of the known subfamilies from bilaterians, the T-box domain is 55-65% identical to Tbx2/3, Tbx4/5 and Tbx6/16 subfamilies and less than 48% identical to Brachyury and Eomes subfamilies. No homology is present outside the T-box domain.

Two distinct T-box family members were isolated from the ctenophore *Pleurobrachia pileus*. One is clearly a member of the Brachyury subfamily and the other a member of the Tbx2/3 subfamily. Within the T-box domain *Pleurobrachia Brachyury* is 70-80% identical to Brachyury subfamily members and less than 55% identical to other subfamilies while little similarity can be detected outside of the T-box. *Pleurobrachia Tbx2/3* is 60-70% identical to Tbx2/3 subfamily members and up to 60% identical to Tbx4/5 in the T-box domain, and less related to other subfamilies.

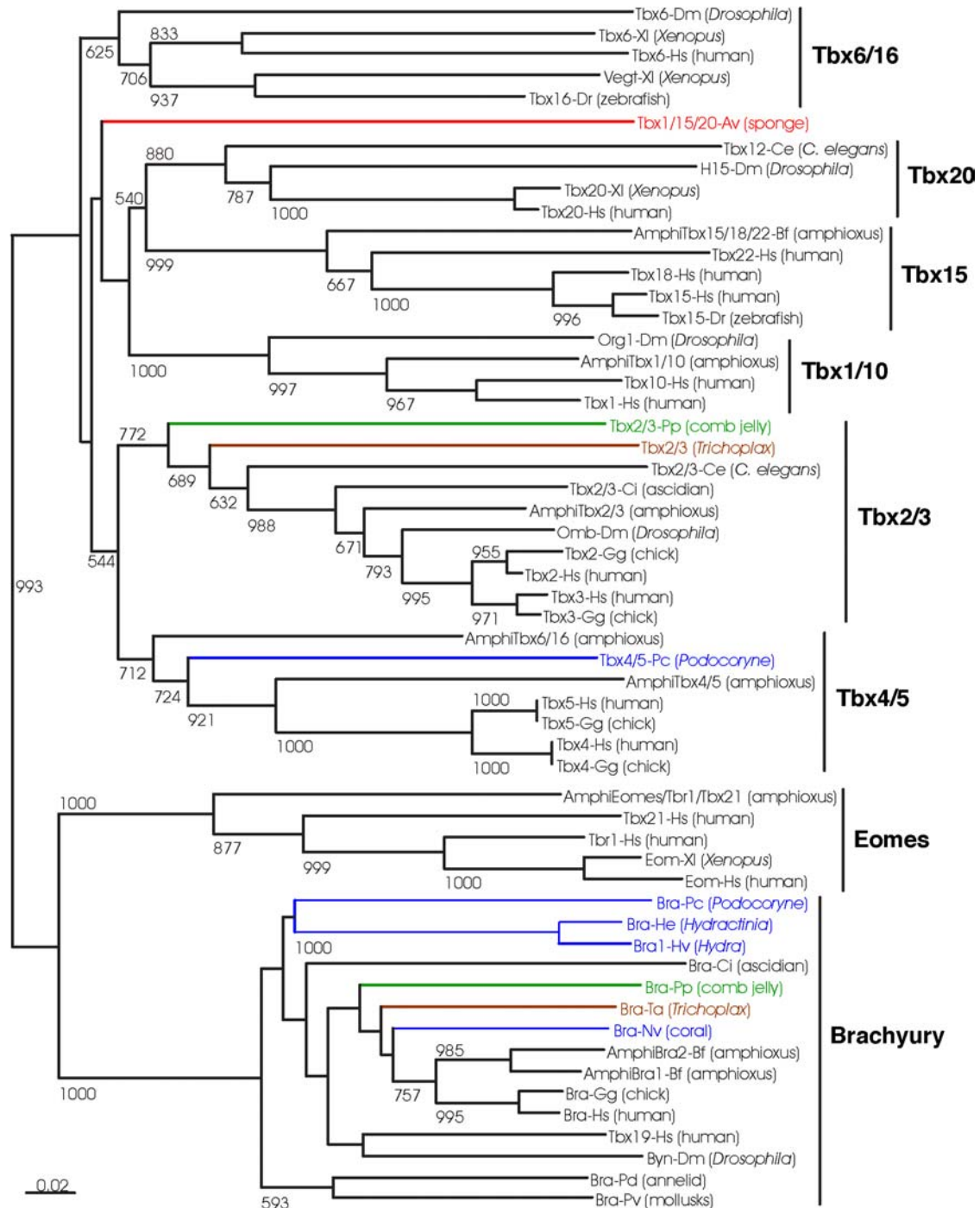
Two distinct T-box family members were isolated from the cnidarian *Podocoryne carnea* by homology PCR on genomic DNA in addition to the already cloned *Brachyury* homologue (Spring et al. 2002). One of the genomic fragments could not be extended on cDNA and the incomplete sequence gives a weak indication that it could belong to the Tbx1 subfamily (data not shown). The other gene is clearly a member of the Tbx4/5 subfamily, in fact in the T-box domain it is 70-75% identical to the Tbx4/5 subfamily, while it is 60-65% identical with the Tbx2/3 subfamily and less than 55% identical with other members of the T-box family.

### Distinct T-box gene subfamilies in non-bilaterians

Phylogenetic analyses for T-box genes were performed by the neighbour-joining methods (N-J) with Clustal X (Fig. 1) and by the maximum likelihood method (M-L) with TREE-PUZZLE (Fig. 2). T-box genes were selected from representative phyla, then the trees were calculated using only the T-box domain of each protein.

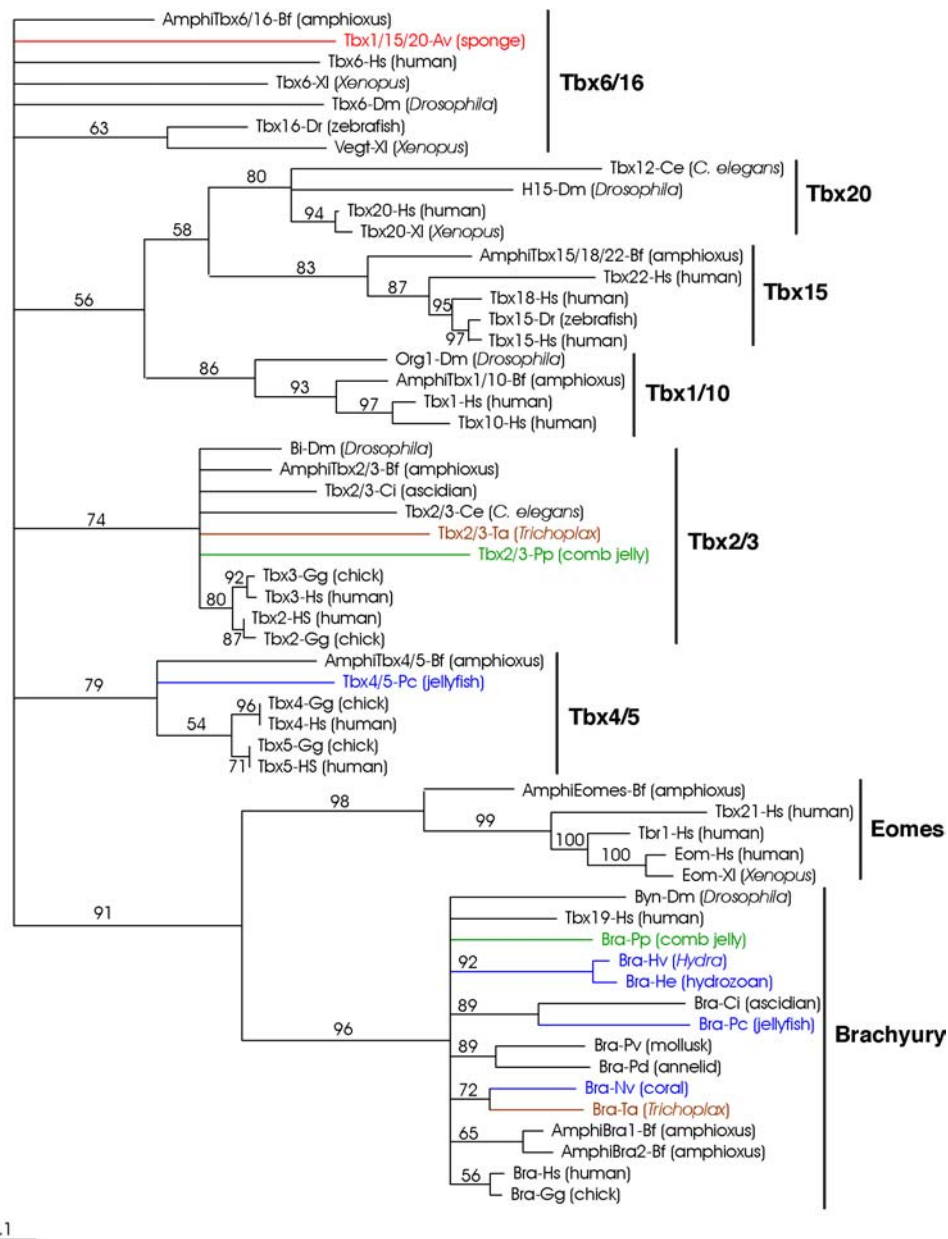
In accordance with previous analyses, the classification made on sequence homology allowed to recognize from 5 to 8 T-box gene subfamilies (Papaioannou 2001; Ruvinsky et al. 2000). Our phylogenetic analysis shows clearly the presence of 8 subfamilies of T-box genes,

this division is consistent in both N-J and M-L analyses (Fig. 1 and 2). Moreover, these subfamilies form two larger groups; in one group are the Brachyury and Eomes subfamilies, in the other group are the remaining subfamilies.



**Fig. 1.** Phylogenetic trees of T-box family based on the neighbour-joining method generated with Cluster X. Numbers on the branches are the values relative to 1000 bootstrap replicates that support the topology shown, the values below 500 are omitted; the bar indicates the number of substitutions per site. The non-bilateral phyla are highlighted in colour: the sponge sequences are red, the *Trichoplax* sequences are brown, the cnidarian sequences are blue and the ctenophore sequences are green. Abbreviations: Dm, *Drosophila melanogaster*; XI, *Xenopus laevis*; Hs, *Homo sapiens*; Dr, *Danio rerio*; Av, *Axinella verrucosa*; Ce, *Caenorhabditis elegans*; Bf, *Branchiostoma floridae*; Pp, *Pleurobrachia pileus*; Ta, *Trichoplax adhaerens*; Ci, *Ciona intestinalis*; Pc, *Podocoryne carnea*; He, *Hydractinia echinata*; Hv, *Hydra vulgaris*; Nv, *Nematostella vectensis*; Pv, *Patella vulgata*; Pd, *Platynereis dumerilii*.

The members that compose each subfamily are roughly the same in both N-J and M-L analyses, with the exception of *AmphiTbx6/16* from amphioxus and *Tbx1/15/20* from *Axinella*. The former belongs with high bootstrap value to the Tbx4/5 subfamily in the tree calculated by N-J (Fig. 1), while it belongs to the Tbx6/16 subfamily in the tree calculated by M-L (Fig. 2); the latter is basal to Tbx20, Tbx15 and Tbx1/10 subfamilies in the tree calculated by N-J with low bootstrap value (Fig.1), while it is with the Tbx6/16 in the tree calculated by M-L (Fig. 2). Although, the T-box gene isolated from sponge does not belong clearly to any subfamily, it seems more related to the Tbx20, Tbx15, Tbx1/10 or the Tbx6/16 subfamilies than to one of the other subfamilies.



**Fig. 2.** Phylogenetic tree of the T-box family based on the maximum likelihood method generated with TREE-PUZZLE. Bootstrap (percentage on 1000 replicates) values are shown at the nodes, the values below 50 are omitted. Explanations are as in Fig. 1.

The two T-box genes isolated from *Pleurobrachia* belong clearly to the Brachyury and Tbx2/3 subfamilies. In the Tbx2/3 subtree calculated by the N-J method, *Pleurobrachia Tbx2/3* is basal to the other members of the subfamily and placed near to the *Trichoplax* homologue, suggesting an early branching of ctenophores in the evolutionary history. On the contrary, *Pleurobrachia Brachyury* is not basal in the *Brachyury* subtree, but since phylogenetic analysis of *Brachyury* does not conform to any reasonable expectation, it is unlikely that this subtree would be informative in relation to the evolutionary history of the phyla represented.

The phylogenetic analyses confirm that the T-box gene isolated from *Podocoryne* belongs to the Tbx4/5 subfamily. *Podocoryne Tbx4/5* is the first gene of the Tbx4/5 subfamily that has been isolated from invertebrates. This subfamily was not expected in invertebrates, where only members of the Tbx2/3 subfamily have been isolated. In fact, a model had been developed for the evolution of the Tbx2/3 and Tbx4/5 subfamilies (Agonic et al. 1996), which proposed that a single primordial *Tbx2/3/4/5* gene underwent an unequal crossing-over event to produce *Tbx2/3* and *Tbx4/5* progenitor genes around the time of divergence of the vertebrate and invertebrate lineages when complete genome duplications are thought to have occurred (Holland et al. 1994). Later during vertebrate evolution, before the origin of tetrapods, a whole cluster duplication or transposition event gave rise to the *Tbx2-Tbx4* cluster on one chromosome and the *Tbx3-Tbx5* cluster on a second chromosome, generating two cognate gene pairs. The presence of a Tbx4/5 gene in cnidarians does not fit with the above-mentioned theory, at least in the timing of the first gene duplication. Anyway, we did not find in *Podocoryne* a member of the Tbx2/3 subfamily so it could be that in *Podocoryne* the primordial *Tbx2/3/4/5* gene evolved toward the Tbx4/5 subfamily instead of the Tbx2/3 subfamily.

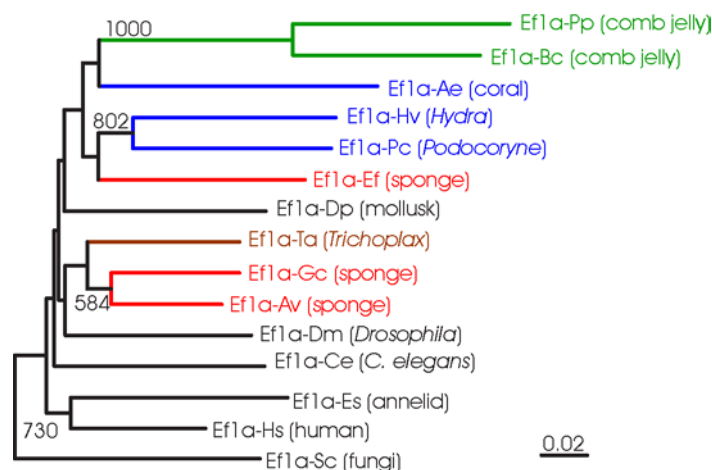
Phylogenetic relationship of non-bilaterians from amino acid sequences of elongation factor-1-alpha (EF1a)

*Elongation factor-1-alpha (EF1a)* genes from the ctenophore *Pleurobrachia pileus* and from the sponge *Axinella verrucosa* were isolated in addition to the previously described *Trichoplax* and *Podocoryne* homologues. *EF1a* encodes for a highly conserved ubiquitous protein that is involved in aminoacyl-tRNA binding in the step of translation of mRNA. EF1a amino acid sequences were already used to infer the evolutionary history of early metazoans

(Kobayashy et al. 1996), supporting the monophyletic origin of metazoans, but it did not resolve the phylogenetic relationship among non-bilaterians.

Phylogenetic analyses for EF1 $\alpha$  were performed by the neighbour-joining (N-J) method using Clustal X (Fig. 3) as well as by the maximum-likelihood method, but in that case the tree obtained does not have enough resolution to allow any phylogenetic consideration (data not shown). The tree in figure 3 was calculated with the EF1 $\alpha$  sequences from the most representative metazoans, and *Saccharomices cerevisiae* was chosen as outgroup.

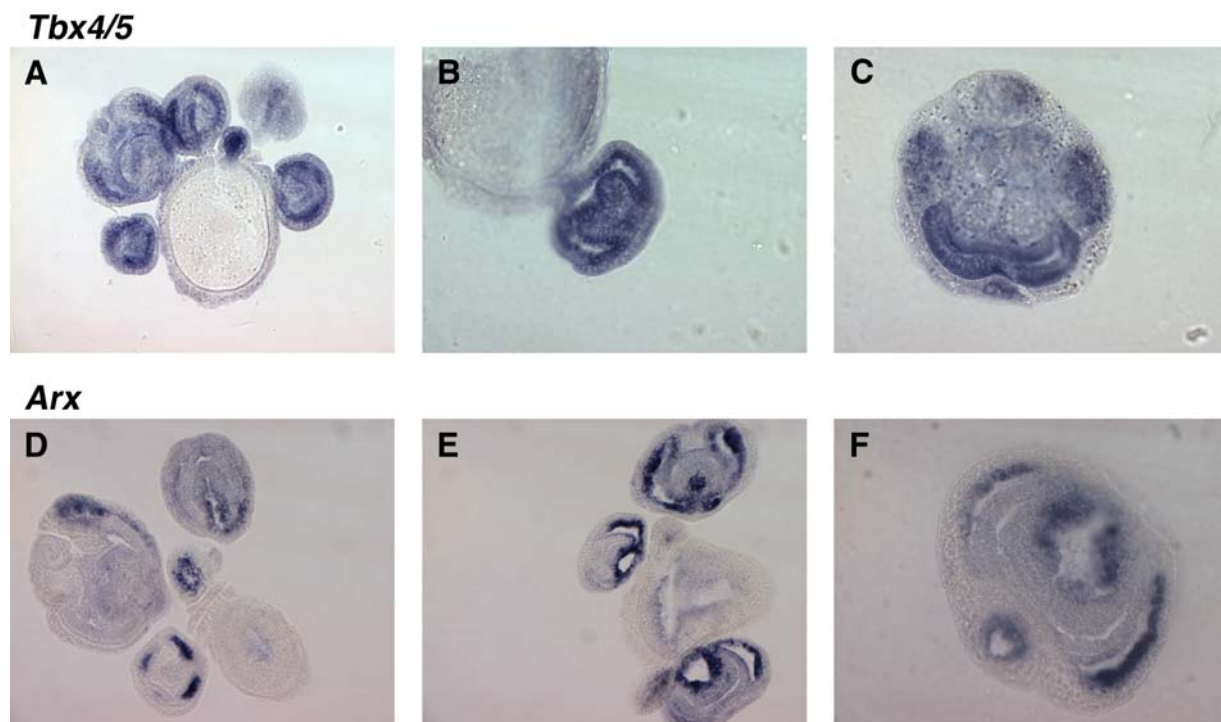
In such a tree the monophyletic origin of metazoans could be confirmed, but unfortunately it is also clear that the EF1 $\alpha$  sequences do not help in resolving the relationship among early metazoans. Nevertheless, this tree supplies some information on non-bilaterian phyla. About ctenophores, it confirms the hypotheses obtained by *18S rRNA* gene sequences (Podar et al. 2001), that they are monophyletic and highly derived. In fact the two species analysed *Pleurobrachia pileus* (class Tentaculata) and *Beroe cucumis* (class Nuda) form a clade with high bootstrap value. About sponges, it weakly suggests that the Calcarea and the Demospongiae classes are sister groups, since *Geodia cydonium* (class Calcarea) and *Axinella verrucosa* (class Demospongiae) form a clade, however, the freshwater demosponge *Ephydatia fluviatilis* does not group with the other sponges. Thus, these data have some value only if compared with similar conclusions reached by a previous phylogenetic analysis combining *18S* and *28S rRNA* genes (Medina et al. 2001). Like the sponges, the cnidarians do not form a clade; but at least sequences belonging to members of the same class are in the same group. EF1 $\alpha$  sequences for the class Cubozoa and Schyphozoa or Hexactinellida are still missing, so the analyses cannot be completed regarding cnidarians and sponges



**Fig. 3** Elongation factor 1 alpha tree. The phylogenetic tree is based on the neighbour-joining method generated with Clustal X. Bootstrap (1000 replicates) values are shown at the nodes, the values below 500 are omitted; the bar indicates the number of substitution per site. Explanation are as in Fig. 1 except for: Bc, *Beroe cucumis*; Ae, *Anemonia erythraea*; Ef, *Ephydatia fluviatilis*; Dp, *Dreissena polymorpha*; Gc, *Geodia cydonium*; Es, *Enchytraeus sp.*; Sc, *Saccharomices cerevisiae*.

Expression of *Tbx4/5* during the medusa bud development of *Podocoryne*

In situ hybridisations were performed in parallel for two genes of the hydrozoan *Podocoryne carnea*, *Tbx4/5* and the homeobox gene *Arx*; the latter was used to compare the expression of *Tbx4/5* with a second gene. In situ hybridisations revealed that both *Tbx4/5* and *Arx* genes had a detectable signal only in the medusa bud, while no signal was present either in polyps or in medusae. In order to investigate which tissues were labelled at the different medusa bud stages, some samples were processed for sectioning.



**Fig. 4** Preliminary expression analysis by in situ hybridisation of *Podocoryne carnea*. *Tbx4/5* is mainly expressed in the endoderm; (A) section of a colony shows the expression pattern from early medusa bud stage to a late medusa bud stage and in each of them the endoderm and its derivatives are labelled, (B) and (C) in higher magnification of medusa buds the signal is present in the ectoderm and in other regions of medusa bud still in the endoderm it is stronger than in the entocodon or ectoderm. *Arx* is expressed only in late medusa bud stages and the signal is sharp and localized in the radial canals, (D) and (E) sections of colonies show the signal is mostly localized in the endoderm of the radial canals, (F) higher magnification of medusa bud.

The sections show distinct expressions for *Tbx4/5* and *Arx*. *Tbx4/5* is expressed in all medusa bud stages and the signal is higher in the endoderm than in the rest of the bud (Fig. 4A). The signal is strong in the endoderm from beginning of bud formation up to late stages of bud development, where tissues deriving from endoderm appear labelled, such as the epithelium of ring canals (Fig. 4B and C). Expression of *Tbx4/5* appears as well in others tissues, but it is several times weaker than the endoderm signal. *Brachyury*, another T-box gene was described



in *Podocoryne* (Spring et al. 2002). *Brachyury* is involved in the process of mesoderm formation, a role that seems to be conserved during evolution (Papaioannou 2001). *Podocoryne Brachyury* is expressed only in the entocodon, that is the third germ layer of the medusa bud (Spring et al. 2002). The expression of *Tbx4/5* is completely different from that of *Brachyury* and this could be evolutionary meaningful. Members of the *Tbx4/5* subfamily characterized so far have not been found to be involved in mesoderm formation. If we consider the entocodon of hydrozoans to be the homologue of mesoderm, *Podocoryne Tbx4/5* should not be expressed there, as is confirmed by the in situ analyses. In addition the expression of *Podocoryne Tbx4/5* confirms that already early in the evolution the T-box genes there were already well-established and well separated roles for early duplicated members of the T-box family.

*Podocoryne Arx* is not expressed in all medusa bud stages, the signal appears only in the late stages (Fig. 4D and E). *Arx* has a sharper signal than *Tbx4/5* and is localized in the muscle of ring canals and manubrium (Fig. 4F). Further studies would be necessary to elaborate on the dynamics of *Arx* expression during medusa bud development of *Podocoryne*.

## CHAPTER 3

### **Role of the homeobox gene *Not* in body remodelling of the basal metazoan *Trichoplax adhaerens***

**Abstract** The new phylogeny, built on molecular data, assumes that protostomes and deuterostomes form a clade sharing a common ancestor. This theory gained popularity after the discovering in *Drosophila* and mammals of a class of *homeobox* genes organized in a cluster and expressed in a colinear pattern, the *Hox* genes. Although the conservation of the *Hox* genes was confirmed in several bilaterians, it is still not clear how the *Hox* cluster originated and whether it is a unique character of bilaterians. To address this issue *Hox* genes were explored in basal metazoans. Special efforts were made in cnidarians where so far only a partial proto-*Hox* cluster has been found. From others non-bilaterian phyla such as Porifera and Placozoa some *homeobox* genes were known, but no true *Hox* or *Hox* cluster were found. We investigated the existence of *homeobox* genes in the placozoan *Trichoplax adhaerens* and in the ctenophore *Pleurobrachia pileus*. We isolated a *Not* homologue from *Trichoplax* and from *Pleurobrachia* a *homeobox* gene that shows low similarity with members of the *Tlx* class. *Not* genes have been studied in lower vertebrates, where they have a main role in notochord formation, but are absent or derived in mammals. Surprisingly we found that *Not* has a unique expression profile in *Trichoplax*. It is highly expressed in the folds of intact animals and in the wounds of regenerating animals, suggesting *Not* involvement in the general *Trichoplax* body remodelling. Recruitment of genes for different processes is a typical feature of evolution and here we show a clear example of such a phenomenon. True *Hox* genes seem to be rare in low metazoans, perhaps they were lost in evolution or these taxa branched before this class originated.

**Keywords** *Homeobox*, *Hox* cluster, *Not*, *Trox-2*, *Trichoplax*, *Pleurobrachia*.

## Introduction

*Homeobox* genes are present in animals, plants and yeast (Duboule 1994). In bilaterians a class of them is grouped in clusters of genes called the *Hox* clusters (Brooke et al. 1998). *Hox* genes code for related homeodomain transcription regulators and generally respect the colinearity rule. Although the *Hox* cluster arose before the divergence of bilaterians, it is still unclear when and how the cluster originated. To address this issue it was already discussed that we should look at the most basal taxa to understand the origins of *Hox* genes (Ferrier and Holland 2001). Although *Hox*-like genes occur in cnidarians (Gauchat et al. 2000), up to now the only evidence for a cnidarian equivalent of the *Hox* cluster is that a *Hox*-like gene and an *even-skipped* orthologue have been shown to be closely linked in the coral *Acropora formosa* (Miller and Miles 1993), and in the anemone *Nematostella vectensis* (Finnerty and Martindale 1999).

Besides *Hox* genes other “dispersed” *homeobox* genes, called non-*Hox* genes, have been discovered in metazoans. Non-*Hox* genes were believed that do not to form any cluster, but recently it was reported that in mammals some are part of 14 clusters that in the bilaterian ancestor would come from 4 clusters (Pollard and Holland 2000). Many of the non-*Hox* gene classes turned out to be important in several developmental processes and they could also be useful in understanding evolution (Gauchat et al. 2000). In order to collect further information on *homeobox* genes in the most basal taxa we investigated *homeobox* genes in two species belonging to non-bilaterians, the placozoan *Trichoplax adhaerens* and the ctenophore *Pleurobrachia pileus*.

*Trichoplax adhaerens* is a small marine organism that is the only well characterized member of the phylum Placozoa (Grell 1971; Grell and Ruthmann 1991; Ruppert and Barnes 1994). It has a flat body with roundish shape, no axes, a dorso-ventral orientation with three cell layers and on both the dorsal and ventral side is covered by cilia (Ruppert and Barnes 1994). Little is known about reproduction of *Trichoplax*, in aquaria it does it irregularly by fission, while under special circumstances it was observed to build hollow swarmer spheres, stolons and eggs developing up to the 64 cells blastula stage (Grell 1971; Ruthmann et al. 1981; Thiemann and Ruthmann 1991). A fragment of a homeobox gene called *Trox-2*, a *Gsx* homologue, was found in *Trichoplax* (Schierwater and Kuhn 1998). Besides sequencing of a partial cDNA fragment no other analyses has been carried out on *Trox-2*.

A higher degree of structural complexity than in *Trichoplax* is reached in the ctenophore *Pleurobrachia pileus*, which has structures and organs comparable to bilaterians.

*Pleurobrachia pileus* belongs to the order Cydippida of Ctenophora (Mills 2001). Cydippida present some typical features: the body has a biradial symmetry and ovoid shape, small dimension (*Pleurobrachia* has diameter of about 2 cm), a longitudinal oral-aboral axis orientation, two long tentacles with 15-20 times the body length, 8 rows of fused cilia arranged as meridians called combs responsible for the animal movement, an apical organ at the aboral end controlling the movement of the animal and a peculiar development through a cydippide-like larva (Hernandez-Nicaise 1991). Based on these morphological features some authors believe them to be an intermediate step toward bilaterians (Martindale and Henry 1998) they were even thought to be a degenerated bilaterians (Nielsen 1995), but the same author changed his mind later (Nielsen 2001). Differently from cnidarians and *Trichoplax*, no *homeobox* genes have been analysed in ctenophores except through short PCR fragments (Yamada and Martindale 2002).

We isolated a *Not* homologue from *Trichoplax* and a *homeobox* gene with low homology to the *Tlx* class from *Pleurobrachia*. No further analyses were possible with *Pleurobrachia*, but the *Trichoplax Not* gene was studied more carefully.

*Xnot* was discovered in *Xenopus* (von Dassow et al. 1993), successively *Not* genes were characterized in zebrafish (Talbot et al. 1995) and in chicken (Stein et al. 1996). Interestingly, no obvious *Not* homologue was found in mammals. In the course of the human genome sequencing project a *Not*-like gene with reduced sequence similarity was detected. *Not* is involved in mesoderm formation, notochord formation and somites formation in fish, frogs and chicken (Gont et al. 1996; Talbot et al. 1995; Yasuo and Lemaire 2001). In the sea urchin *Strongylocentrotus purpuratus* *Not* is involved in mesoderm formation and secondary structure formation in the late stage of larval development (Peterson et al. 1999).

We wondered about the developmental processes in which *Not* could be involved in *Trichoplax*. Being aware that no embryology is available in *Trichoplax*, we investigated whether *Not* has any role in *Trichoplax* regeneration by processing some intact and wounded animals for in situ hybridisation. *Not* has a unique expression pattern in intact and regenerating *Trichoplax*. These patterns indicate a possible role of *Not* during body remodelling and wound closure in this organism. For comparison we included the only other known gene *Trox-2* which is best considered as an anterior Para-*Hox* gene (Hayward et al. 2001; Li et al. 1996). *Trox-2* gives a weak signal in the periphery of the body in intact samples. Besides, we compared the *Not* and *Trox-2* expression patterns with the signal of a third gene called *Secp1*, coding for a highly charged protein. This gene is expressed in a clear ring-shaped pattern along the animal.

## Materials and methods

### Animals

*Trichoplax adhaerens* appear regularly on the glass walls of marine aquaria maintained to culture the hydrozoan jellyfish *Podocoryne carnea*. Animals are removed from the glass walls of the aquaria by a blow with a Pasteur pipette and transferred to a clean glass dish. They were left to adhere on the glass surface for about 2-3 hours and washed 4 times with Millipore filtered sea-water. Wounded animal were obtained by cutting with a little blade when they were sitting on the bottom of the dish. Individual specimens were selected for RNA and DNA extractions or *in situ* hybridization experiments.

The comb jelly *Pleurobrachia pileus* was fished at the Marine Biology Station of Roscoff during plankton collection. We were able to keep some of those specimens in our laboratories long enough to collect some larva that were washed with Millipore filtered sea water and then processed for RNA and DNA extraction. The adult animals survived about 10 days in our laboratory and no larvae were observed to reach the adult stage.

### Molecular cloning and sequence comparisons

Genomic DNA and total RNA were isolated from both *Trichoplax* and *Pleurobrachia* with the TriReagent (Molecular Research Center) according to the manufacturer's recommendations. First strand cDNA was synthesized with the anchored oligo (dT) primer XT20V (5'-GGC AGG TCC TCG TTG ACT CGA GAC GT<sub>(20)</sub>(AGC)-3') by using the SMART RACE cDNA Amplification Kit (Clontech). With the Smart cDNA of *Trichoplax* we were able to perform homology PCR, 3' and 5' RACE to obtain the full length of the *Not* homologue and 5' RACE to obtain the full length of *Trox-2*. With the Smart cDNA of *Pleurobrachia* we were able to conduct homology PCR, 3' and 5' RACE to obtain the full length of the *Tlx*-like gene. A 83 bp *Not* fragment was amplified with the set of degenerated primers HoxE and HoxF (Pendleton et al. 1993). For PCR round standard conditions were used, except that the annealing temperature was 37°C for 10 cycles and 50°C for 40 cycles. A 83 bp *Tlx*-like fragment was isolated from *Pleurobrachia* with the same procedure described previously for *Trichoplax Not*. PCR products of the expected size were gel purified with a Qiaquick column (Qiagen), subcloned in the pCRII-TOPO vector (TOPO TA cloning Dual

Promoter kit, Invitrogen) and sequenced on an ABI PRISM 310 genetic analyser (Applied Biosystems). Based on the sequences gene-specific primers were designed to carry out the 5' and 3' RACE as described elsewhere (Müller et al. 1999). Clones with the complete coding sequences were isolated for *Trichoplax Not* (1715 bp) and *Trox-2* (1273 bp), a clone with the complete coding sequence was isolated for *Pleurobrachia Tlx-like* (1355 bp). The sequences have been submitted to the DNA.

Nucleotide and deduced amino acid sequences were analysed by using the GCG software package. BLAST searches (Altschul et al. 1997) were performed on the blast network service at the NCBI (<http://www.ncbi.nlm.nih.gov>). Multiple sequence alignments and phylogenetic trees based on the neighbour-joining method were generated with Clustal X (Jeanmougin et al. 1998).

### *In situ* hybridisation

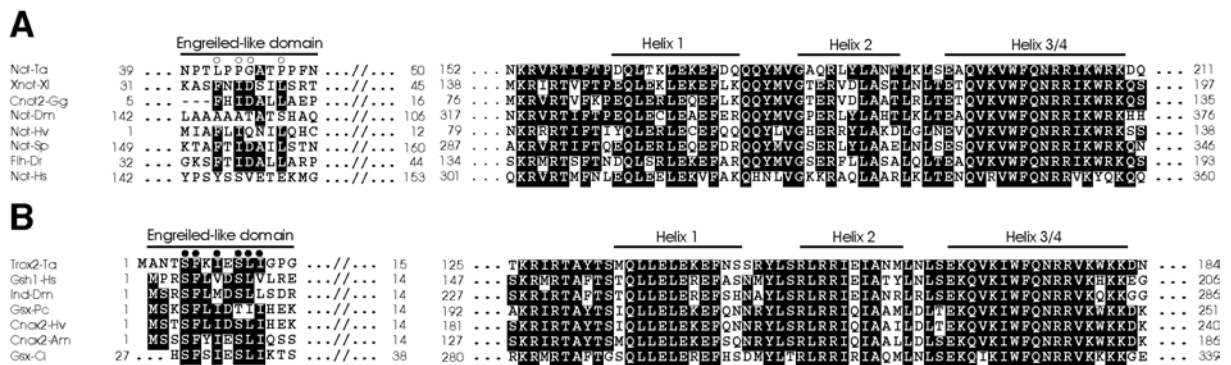
Whole mount *in situ* hybridisation experiments were carried out with adaptations to the protocol described for *Podocoryne carnea* (Spring et al. 2002; Yanze et al. 2001). Animals were fixed overnight at 4°C in Lavdowsky fixative supplemented with 0.2% glutaraldehyde; washed 3 times for 30 min in PBST (0.1% Tween20 in PBS); stepwise increasing the concentration from 15%, 30%, 45%, 60%, 75%, 90% and 100% of step-HB (5 x SSC; 50% formamide; 50 mg/l heparin; 0.1% Tween20) in 5 x SSC, shaking at each step 10 min at room temperature (RT); prehybridised in hybridisation buffer (HB) (5 x SSC; 50% formamide; 50 mg/l heparin; 100 mg/l tRNA from *E. coli* (Sigma type XXI R4251); 0.1% Tween20) for 1h at 50°C; hybridised with DIG-labelled antisense probe (approximately 20 ng/ml HB) over night at 50°C; washed 2 times for 20 min at 50°C respectively with WS1 (5 x SSC; 50% formamide; 0.1% Tween20), WS2 (2 x SSC; 50% formamide; 0.1% Tween20), WS3 (2 x SSC; 0.1% Tween20) and WS4 (0.2 x SSC; 0.1% Tween20); washed in PBST for 5 min at RT; washed in blocking solution (1% blocking reagent; 100 mM maleic acid pH 7.5; 150 mM NaCl) for 1h at RT; incubated with anti-DIG-AP Fab fragments 1:5000 in blocking solution, over night at 4°C; washed 3 times with PBST and equilibrated in TMNT (100 mM Tris-HCl, pH 8; 50 mM MgCl<sub>2</sub>; 100 mM NaCl; 0.1% Tween20) for 30 min at RT. Detection was done incubating the specimens in TMN (100 mM Tris-HCl, pH 8; 50 mM MgCl<sub>2</sub>; 100 mM NaCl) with 340 mg/l NBT and 175 mg/l BCIP. When the chromogenic reaction was ready (2 to 4 hours), was stopped by washing the specimens in PBST. DIG-RNA-labelled probes were synthesized according to the manufacture's recommendation (DIG-RNA-labelling, Roche).

The DNA fragments for *Not* and *Trox-2* probes were amplified with primers designed outside the homeodomain as following: TNot F3 (5'-TTG TCA GCC TAA TCG CTT ACC-3') and TNot R2 (5'-CCC ACC ATA TAC TGT TGT TGA TC-3'), Trox F2 (5'-ACA TTC CAT CAT TCA TCG CCT G-3') and Trox R3 (5'-AGT TCT GGA AGA ATC ACT GGT C-3'). These fragments were subcloned in the pCRII-TOPO vector (TOPO TA cloning Dual Promoter kit, Invitrogen) and used as template for the probe synthesis.

## Results and discussion

### Highly conserved *Not* homologue from *Trichoplax*

*Trichoplax Not* has 70-80% of sequence identity with members of the *Not* class in the homeodomain, while there is less than 60% identity with others *homeobox* genes. In *Xenopus* it was shown that *Xnot* codes for a transcriptional activator, confirmed to the presence of an engrailed-like activator domain at the amino terminus of the amino acid sequence (Yasuo and Lemaire 2001). In the alignment done of *Trichoplax Not* with other members of the family it appears that some of the species represented lack the engrailed-like activator domain (Fig. 1A). This lack does not have anything to do with the progress of evolution, since for instance among the non-bilaterians *Trichoplax* misses the domain while *Hydra* might have it, and in bilaterians *Drosophila* and humans lost it, while all the others species represented have the domain (Fig. 1A). Most probably the common ancestor of metazoans had the engrailed-like domain, but during evolution multiple species lost it independently. Perhaps in these organisms *Not* was recruited for other functions that do not need this activator domain.



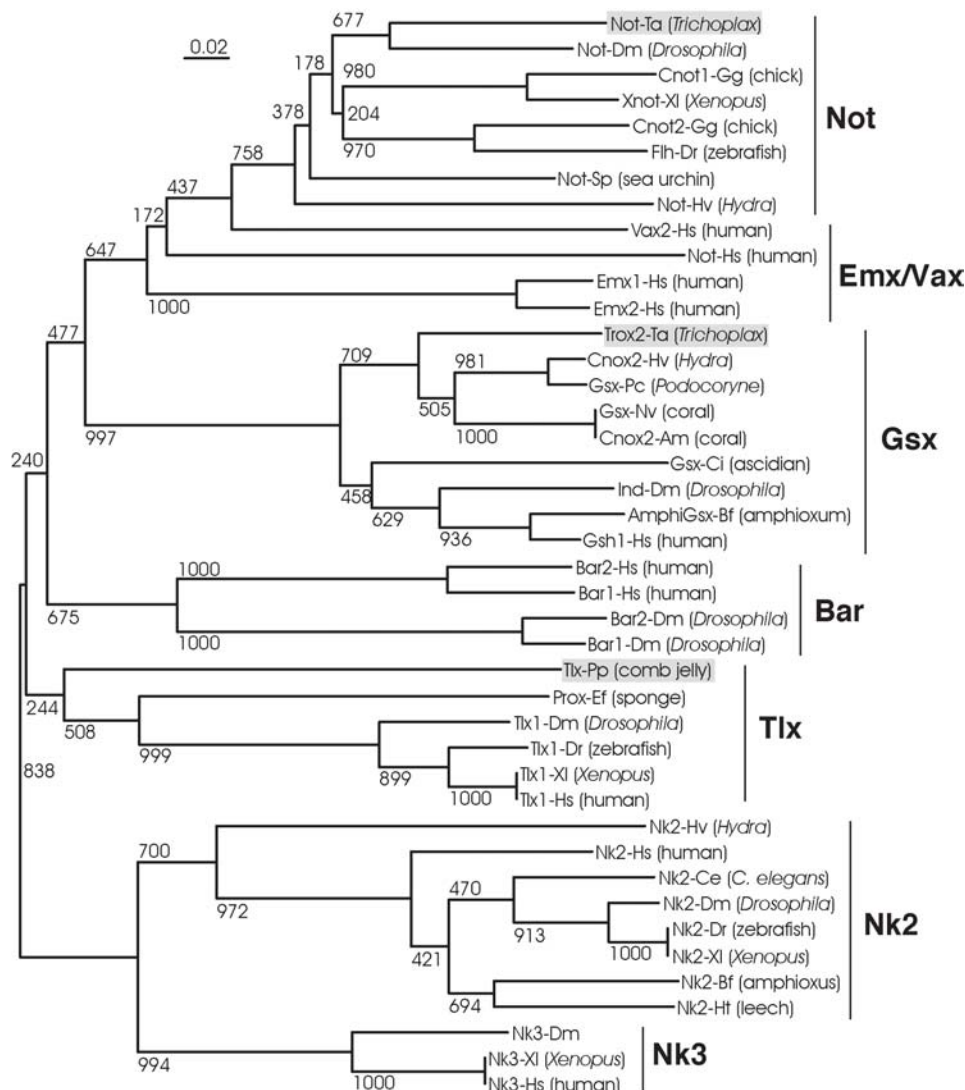
**Fig. 1.** Comparison of *Trichoplax Not* and *Trox-2* with class members. (A) The alignment shows the two regions most conserved in the *Not* class, the putative engrailed-like activator domain and the homeodomain. The engrailed-like activator domain is missing in *Trichoplax*, as well as in *Drosophila* and human. Conserved residues in other species are labelled with empty circles. In the *Not* homeodomain the four putative helices involved in DNA binding are highlighted with a bar (after Duboule 1994). (B) The *Gsx* class has the engrailed-like activator domain placed at the amino terminal end of the amino acid sequence. Full circles label the residues that are conserved in *Trichoplax*. The *Gsx* homeodomain is labelled like in A. Abbreviations: Ta, *Trichoplax adhaerens*; Xl, *Xenopus laevis*; Gg, *Gallus gallus*; Dm, *Drosophila melanogaster*; Hv, *Hydra vulgaris*; Sp, *Strongylocentrotus purpuratus*; Dr, *Danio rerio*; Hs, *Homo sapiens*; Pc, *Podocoryne carnea*; Am, *Acropora millepora*; Ci, *Ciona intestinalis*.

We sequenced the missing 5' end of *Trox-2* in order to have the full information of the coding sequence and compare it with the other members of the *Gsx* class (Fig. 1B). *Trox-2* has 83-86% of sequence identity with the members of the *Gsx* class in the homeodomain, while the identity is below 70% with other homeobox genes. Like the *Not* class, the *Gsx* class has an engrailed-like activator domain, which, in contrast to the *Not* class, is highly conserved



throughout the Gsx class and is placed at the amino terminus of the amino acid sequence characterizing this genes (Hayward et al. 2001).

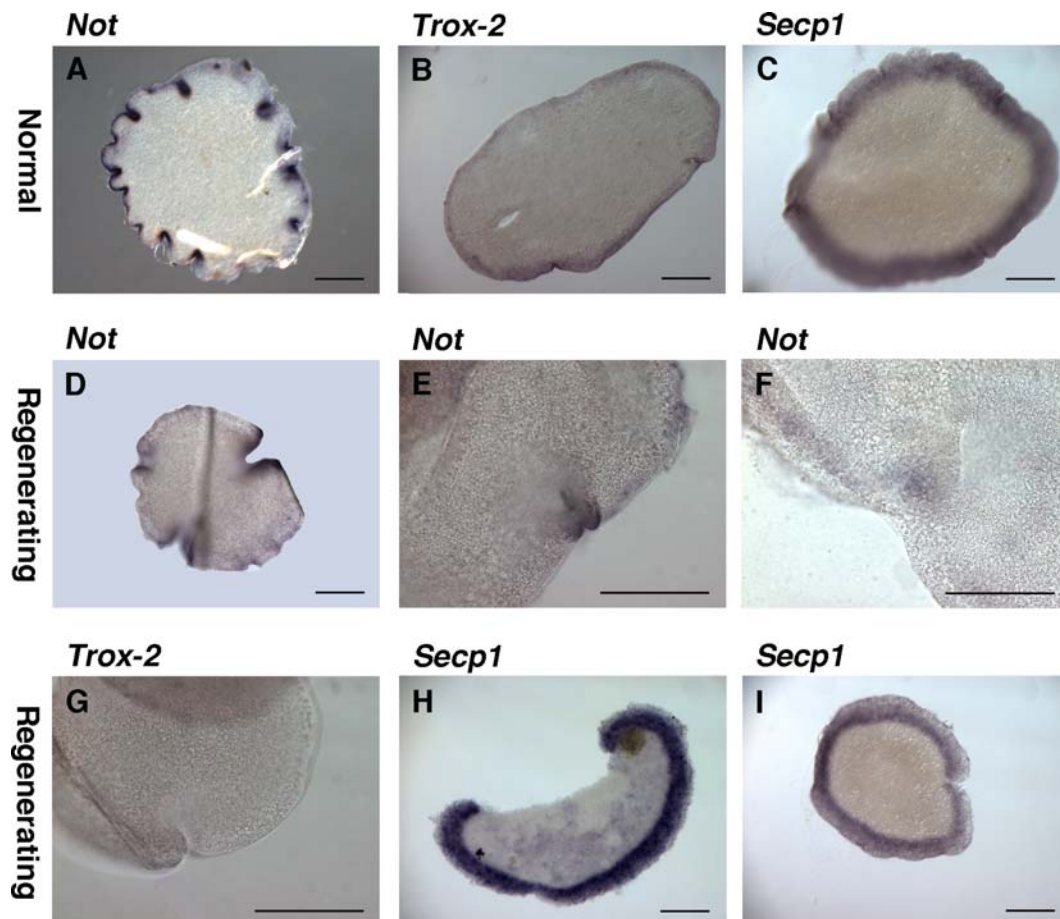
Phylogenetic studies were carried out with the neighbour-joining method based on the homeodomain of each sequence. The tree indicates that the two *Trichoplax* gene *Not* and *Trox-2* belong to well-defined classes (Fig. 2). The *homeobox* gene that we found in *Pleurobrachia* has a low sequence identity of about 50% with several members of non-*Hox* classes such as *Tlx*, *Bar* and *Nk2*. In the phylogenetic tree that we show the *Pleurobrachia homeobox* gene groups with the *Tlx* class, but the low similarity and bootstrap value does not allow to consider this gene as a clear *Tlx* orthologue.



**Fig. 2.** Phylogenetic tree based on the neighbour-joining method generated with Clustal X. Numbers on the branches are the values relative 1000 bootstrap replicates that support the topology shown, on the bar is the number of substitution per site. The tree shows clearly that *Trichoplax Not* belongs to the Not class related to the Emx and Vax classes. Equally clearly *Trichoplax Trox-2* belongs to the para-Hox group Gsx. *Pleurobrachia Tlx* belongs to the Tlx class with a low bootstrap value. Abbreviations are like in Fig. 1 except for: Nv, *Nematostella vectinis*; Ef, *Ephydatia fluviatilis*; Bf, *Branchiostoma floridae*; Ht, *Helobdella triserialis*; Pp, *Pleurobrachia pileus*.

Not involvement in remodelling and wound closure of *Trichoplax*

The powerful tools of functional assays in vertebrates allowed to understand the roles of *Not* during development. The gain of function assay in *Xenopus* shows an enlarged notochord and reduced muscle tissues (Gont et al. 1996), while the loss of function in zebrafish shows an opposite pattern with lack of notochord and fusion of somites (Talbot et al. 1995). *Xnot* is involved in a pathway where the expressions of *BMP* and *Wnt* are inhibited in the dorsal mesoderm by *ftz*, allowing the maintenance of *Xnot* expression (Yasuo and Lemaire 2001). The little knowledge on *Trichoplax* embryology does not allowed any developmental investigation. As an alternative we proceeded by studying regeneration, we treated by in situ hybridisation intact and regenerating *Trichoplax* with the following genes: *Not*, *Trox-2* and *Secp1*.



**Fig.3** Expression analysis by in situ hybridisation of *Trichoplax adhaerens*. (A) *Not* is expressed in the folds of intact *Trichoplax*. (B) *Trox-2* is weakly expressed at the edge of the intact animal, in some region more than in others. (C) In an intact animal, *Secp1* is expressed with a typical ring-shaped pattern along all along the edge. (D) A *Trichoplax* that was wounded in three places, each wound shows high expression of *Not*. (E) *Not* is still highly expressed during the closure of the wound. (F) After closure of the wound by edge fusion the *Not* signal became weak and is almost undetectable. (G) During wound closure *Trox-2* is not expressed in a regenerating animal. (H) *Secp1* does not change the expression pattern a few minutes after a cut, the wound edge has no signal at all. (I) *Secp1* expression displays an interruption of the ring-shaped pattern in proximity of the wound. The bar is 300  $\mu\text{m}$ .

*Not* in the intact animals is expressed in the body folds (Fig. 3A). This unique expression pattern is characterized by a strong signal at the bottom of each fold, while the unfolded body edges appear unlabelled as well as the central part of the body. In intact animals, *Trox-2* appears to have a weak signal at the edge of the body (Fig. 3B), differently than *Secp1* that has a strong signal covering a broader area along the animal in a typical ring-shaped expression pattern (Fig. 3C). Moreover, the presence of folds does not affect the expression pattern of *Trox-2* and *Secp1*.

As most of the lower metazoans, *Trichoplax* has a regeneration potential inducible by cutting one or more pieces of the body. No data are available regarding the dynamic of wound closure, so it is not clear whether in *Trichoplax* the regeneration follow a morphallactic process where no proliferation is required or an epimorphic process. Only the morphological changes during regeneration have been observed. The wound closure process is fast, just after the cut the central part of the body contracts, while the two edges of the wound get close by a sliding movement of the external regions on the central part of the body. In half an hour the two lips of the wound touch each other and need 10 minutes to fuse with each other, so less than one hour is required to close the wound (Schwartz 1984). Moreover, it has been shown that fragments containing both peripheral and central regions of the body are able to regenerate, while fragments containing only a central part of the body or a peripheral part are not able to regenerate (Schwartz 1984). The *Not* expression changes during regeneration with a dynamic pattern. 10 minutes after the cut a regenerating *Trichoplax* shows a strong expression at the wound level (Fig. 3D). During the earliest phases of regeneration the expression pattern is slightly different than in the intact animal, the regenerating animal has a high signal at the lip level of the wound while in the remodelling animal the signal is strongest at the bottom of the folds. *Not* expression is still high during wound closure, after about 30 minutes the two lips of the wound come closer and the *Not* signal is still clearly detectable (Fig. 3E). After the wound closed by fusion of the edges, the *Not* signal is weak and is almost undetectable (Fig. 3F).

Like for the intact specimens, we investigated in regenerating animals the expression patterns of *Trox-2* and *Secp1*. The *Trox-2* signal is absent at wound level during regeneration (Fig. 3G). Although the *Trox-2* signal looks weaker than in an intact animal it is not clear whether the *Trox-2* expression went down all over the animal during regeneration. *Secp1* is still expressed along the edge of the *Trichoplax* body except at the wound level (Fig. 3H) and the regeneration process does not have any effect on *Secp1* expression in the part of the *Trichoplax* body that is still intact. During the last phase of wound closure *Secp1* is expressed

in the typical ring shaped pattern in the regenerating animal, but at the wound level has still an interruption of the signal (Fig. 3I). The dynamic expression pattern of *Not* during wound closure suggests that *Not* is involved in regeneration, although it is not clear which is its role. In contrast to *Not*, *Trox-2* and *Secp1* do not change their expression patterns during regeneration, suggesting that they are not involved in the wound closure process.

## References

- Agulnik SI, Garvey N, Hancock S, Ruvinsky I, Chapman DL, Agulnik I, Bollag R, Papaioannou V, Silver LM (1996) Evolution of mouse T-box genes by tandem duplication and cluster dispersion. *Genetics* 144:249-54
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402
- Amacher SL, Draper BW, Summers BR, Kimmel CB (2002) The zebrafish T-box genes *no tail* and *spadetail* are required for development of trunk and tail mesoderm and medial floor plate. *Development* 129:3311-3323
- Amano S, Hori I (1996) Transdifferentiation of larval flagellated cells to choanocytes in the metamorphosis of the demosponge *Haliclona permollis*. *Biol Bull* 190:161–172
- Amano S, Hori I (2001) Metamorphosis of *Coeloblastula* Performed by Multipotential Larval Flagellated Cells in the Calcareous Sponge *Leucosolenia laxa*. *Biol Bull* 200:20-32
- Arendt D, Technau U, Wittbrodt J (2001) Evolution of the bilaterian larval foregut. *Nature* 409:81-85
- Ax P (1996) *Multicellular Animals. A New Approach to the Phylogenetic Order in Nature. Volume I.* Springer Verlag, Berlin, Heidelberg, New York
- Bagby R, M. (1966) The fine structure of myocytes in the sponges *Microciona porifera* and *Tedania ignis*. *J. Morphol.* 118:167-182
- Boero F, Gravili C, Pagliara P, Piraino S, Bouillon J, Schmid V (1998) The cnidarian premises of metazoan evolution: from triploblasty, to coelom formation, to metamerism. *Ital J Zool* 65:5-9
- Böger H (1988) Versuch über das phylogenetische System der Porifera. *Meyniana* 40:67-90
- Borchiellini C, Boury-Esnault N, Vacelet J, Le Parco Y (1998) Phylogenetic analysis of the Hsp70 sequences reveals the monophyly of Metazoa and specific phylogenetic relationships between animals and fungi. *Mol Biol Evol* 15:647-655
- Bridge D, Cunningham CW, DeSalle R, Buss LW (1995) Class-level relationships in the phylum Cnidaria: molecular and morphological evidence. *Mol Biol Evol* 12:679-689
- Brooke NM, Garcia-Fernandez J, Holland PHW (1998) The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* 392:920-922
- Brown DD, Binder O, Pagratis M, Parr BA, Conlon FL (2003) Developmental expression of the *Xenopus laevis* Tbx20 orthologue. *Dev Genes Evol* 212:604-7
- Cavalier-Smith TM, Allsopp TEP, Chao EE, Boury-Esnault Na, Vacelet J (1996) Sponge phylogeny, animal monophyly, and the origin of the nervous system: 18S rRNA evidence. *Can J Zool* 74:2031-2045
- Chapman DL, Agulnik I, Hancock S, Silver LM, Papaioannou VE (1996) Tbx6, a mouse T-Box gene implicated in paraxial mesoderm formation at gastrulation. *Dev Biol* 180:534-42

## References

---

- Chapman DL, Papaioannou VE (1998) Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature* 391:695-697
- Chombard C, Boury-Esnault N, Tillier A, Vacelet J (1997) Polyphyly of "Sclerosponges" (Porifera, Demospongiae) Supported by 28S Ribosomal Sequences. *Biol Bull* 193:359-367
- Christen R, Ratto A, Baroin A, Perasso R, Grell KG, Adoutte A (1991) An analysis of the origin of metazoans, using comparisons of partial sequences of the 28S RNA, reveals an early emergence of triploblasts. *Embo J* 10:499-503
- Coll M, Seidman JG, Müller CW (2002) Structure of the DNA-bound T-box domain of human TBX3, a transcription factor responsible for ulnar-mammary syndrome. *Structure* 10:343-356
- Collins AG (1998) Evaluating multiple alternative hypotheses for the origin of Bilateria: an analysis of 18S rRNA molecular evidence. *Proc. Natl. Acad. Sci. U S A* 95:15458-63
- Collins AG (2002) Phylogeny of Medusozoa and the evolution of cnidarian life cycles. *J Evol Biol* 15:418-432
- Cunliffe V, Smith JC (1992) Ectopic mesoderm formation in *Xenopus* embryos caused by widespread expression of a Brachyury homologue. *Nature* 358:427-430
- Dobrovolskaïa-Zavadskaïa N (1927) Sur la mortification spontanée de la queue chez la souris nouveau-née et sur l'existence d'un caractère héréditaire 'non-viable'. *C R Soc Biol* 97:114-116
- Duboule D, (Ed.) (1994) A guide book on Homeobox genes. Oxford university press, Oxford
- Ender A, Schierwater B (2003) Placozoa are not derived cnidarians: evidence from molecular morphology. *Mol Biol Evol* 20:130-4
- Ereskovsky A, V. (2000) Reproduction cycles and strategies of the cold-water sponges *Halisarca dujardini* (Demospongiae, Halisarcida), *Myxilla incrustans* and *Iophon piceus* (Demospongiae, Poecilosclerida) from the White Sea. *Biol Bull* 198: 77-87
- Ferrier DEK, Holland PHW (2001) Ancient origin of the Hox gene cluster. *Nat. Rev. Genet.* 2:33-38
- Field KG, Olsen GJ, Lane DJ, Giovannoni SJ, Ghiselin MT, Raff EC, Pace NR, Raff RA (1988) Molecular phylogeny of the animal kingdom. *Science*. *Science* 239:748-753
- Finnerty JR, Martindale MQ (1999) Ancient origin of axial patterning genes: Hox genes and ParaHox genes in the cnidaria. *Evol. Dev.* 1:16-23
- Freeman G (1967) Studies on regeneration in the creeping ctenophore *Vallicula multiformis*. *J Morph* 123:71-84
- Freeman G (1977) The establishment of the oral-aboral axis in the ctenophore embryo. *J Embryol Exp Morphol* 42:237-260
- Galliot B (2000) Conserved and divergent genes in apex and axis development of cnidarians. *Curr Opin Genet Dev* 10:629-37
- Gauchat D, Mazet F, Berney C, Schummer M, Kreger S, Pawlowski J, Galliot B (2000) Evolution of Antp-class genes and differential expression of Hydra Hox/paraHox genes in anterior patterning. *Proc Natl Acad Sci U S A* 97:4493-8

## References

---

- Gibson-Brown JJ, Agulnik SI, Chapman DL, Alexiou M, Garvey N, Silver LM, Papaioannou VE (1996) Evidence of a role for T-box genes in the evolution of limb morphogenesis and the specification of forelimb/hindlimb identity. *Mech Dev* 56:93-101
- Gibson-Brown JJ, Agulnik SI, Silver LM, Niswander L, Papaioannou VE (1998) Involvement of T-box genes Tbx2-Tbx5 in vertebrate limb specification and development. *Development* 125:2499-509
- Gluecksohn-Schoenheimer S (1938) The development of two tailless mutants in the house mouse. *Genetics* 23:573-584
- Gont LK, Fainsod A, Kim SH, De Robertis EM (1996) Overexpression of the homeobox gene Xnot-2 leads to notochord formation in *Xenopus*. *Dev Biol* 174:174-8
- Grell KG (1971) Embryonalentwicklung bei *Trichoplax adhaerens* F. E. Schulze. *Naturwissenschaften* 58:570
- Grell KG, Ruthmann A (1991) *Microscopic Anatomy of Invertebrates, Volume 2, Placozoa*. Wiley-Liss, New York
- Griffin K, Amacher SL, Charles CB, Kimmelman D (1998) Molecular identification of spadetail: regulation of zebrafish trunk and tail mesoderm formation by T-box genes. *Development* 125:3379-3388
- Griffin KJ, Stoller J, Gibson M, Chen S, Yelon D, Stainier DY, Kimmel D (2000) A conserved role for H15-related T-box transcription factors in zebrafish and *Drosophila* heart formation. *Dev Biol* 218:235-47
- Gröger H, Callaerts P, Gehring WJ, Schmid V (2000) Characterization and expression analysis of an ancestor-type Pax gene in the hydrozoan jellyfish *Podocoryne carnea*. *Mech Dev* 94:157-169
- Gross JM, Peterson RE, Wu SY, McClay DR (2003) LvTbx2/3: a T-box family transcription factor involved in formation of the oral/aboral axis of the sea urchin embryo. *Development* 130:1989-99
- Harbison GR (1986) On the classification and evolution of the Ctenophora. In: The origin and relationship of lower vertebrates. Morrison SC, George JD, Gobson R, Platt HM. Systematic Association by Clarendon Press, Oxford pp 78-100
- Harbison GR, Madin LP (1982) *Synopsis and Classification of Living Organisms*. McGraw-Hill Book Co, New York
- Harrison FW, De Vos L (1991) *Microscopic Anatomy of Invertebrates, Volume 2, Porifera*. Wiley-Liss, New York
- Hayward DC, Catmull J, Reece-Hoyes JS, Berghammer H, Dodd H, Hann SJ, Miller DJ, Ball EE (2001) Gene structure and larval expression of *cnox-2Am* from the coral *Acropora millepora*. *Dev Genes Evol* 211:10-9
- Hernandez-Nicaise ML (1991) *Microscopic Anatomy of Invertebrates, Volume 2, Ctenophora*. Wiley-Liss, New York
- Herrmann BG, Labeit S, Poutska A, King TR, Lehrach H (1990) Cloning of the T gene required in mesoderm formation in the mouse. *Nature* 343:617-622
- Hibberd DJ (1975) Observations on the ultrastructure of the choanoflagellate *Codosiga botrytis* (Ehr.) Saviile-Kent with special reference to the flagellar apparatus. *J Cell Sci* 17:191-219

## References

---

- Holland PW, Garcia-Fernandez J, Williams NA, Sidow A (1994) Gene duplications and the origins of vertebrate development. *Dev Suppl*:125-133
- Holland PWH (1999) The future of evolutionary developmental biology. *Nature* 402 (supplement):C41-44
- Horb ME, Thomsen GH (1997) A vegetally-localized T-box transcription factor in *Xenopus* eggs specifies mesoderm and endoderm and is essential for embryonic mesoderm formation. *Development* 124:1689-1698
- Hug B, Walter V, Grunwald DJ (1997) *tbx6*, a Brachyury-related gene expressed by ventral mesendodermal precursors in the zebrafish embryo. *Dev Biol* 183:61-73
- Hyman LH (1940) *The invertebrates: Protozoa through Ctenophora*. McGraw-Hill, New York
- Iio A, Koide M, Hidaka K, Morisaki T (2001) Expression pattern of novel chick T-box gene, *Tbx20*. *Dev Genes Evol* 211:559-62
- Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal X. *Trends Biochem Sci* 23:403-405
- Jerome LA, Papaioannou VE (2001) DiGeorge syndrome phenotype in mice mutant for the T-box gene, *Tbx1*. *Nat Genet* 27:286-291
- Kim J, Kim W, Cunningham CW (1999) A new perspective on lower metazoan relationships from 18S rDNA sequences. *Mol Biol Evol* 16:423-427
- King N, Carroll SB (2001) A receptor tyrosine kinase from choanoflagellates: molecular insights into early animal evolution. *Proc Natl Acad Sci U S A* 98:15032-7
- Kispert A, Herrmann BG, 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 Dev* 8:2137-50
- Kispert A, Koschorz B, Herrmann BG (1995) The T protein encoded by Brachyury is a tissue-specific transcription factor. *EMBO J* 14:4763-4772
- Kobayashi M, Wada H, Satoh N (1996) Early evolution of the Metazoa and phylogenetic status of diploblasts as inferred from amino acid sequence of elongation factor-1 alpha. *Mol Phylogenet Evol* 5:414-422
- Kopp A, Duncan I (1997) Control of cell fate and polarity in the adult abdominal segments of *Drosophila* by *optomotor-blind*. *Development* 124:3715-26
- Krumbach T (1907) *Trichoplax*, die umgewandelte Planula einer Hydromeduse. *Zool Anz* 31:450-454
- Kusch T, Reuter R (1999) Functions for *Drosophila* brachyenteron and forkhead in mesoderm specification and cell signalling. *Development* 126:3991-4003
- 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* 129:1411-21



## References

---

- Leys SP, Degnan BM (2001) The cytological basis of photoresponsive behaviour in a sponge larva. *Biol Bull* 201:323-338
- Leys SP, Degnan BM (2002) Embryogenesis and metamorphosis in a haplosclerid demosponge: gastrulation and transdifferentiation of larval ciliated cells to choanocytes. *Invertebrate Biology* 19:171-189
- Li CW, Chen JY, Hua TE (1998) Precambrian sponges with cellular structures. *Science* 279:879
- Li H, Zeithner PS, Valerius MT, Small K, Potter SS (1996) Gsh-1, an orphan Hox gene, is required for normal pituitary development. *Embo J* 15:714-724
- Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, Jurecic V, Ogunrinu G, Sutherland HF, Scambler PJ, Bradley A, Baldini A (2001) Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* 410:97-101
- Manuel M, Le Parco Y (2000) Homeobox gene diversification in the calcareous sponge, *Sycon raphanus*. *Mol Phylogenet Evol* 17:97-107
- Martindale MQ, Henry JQ (1995) Diagonal development: establishment of the anal axis in the ctenophore *Mnemiopsis leidyi*. *Biol Bull* 189:190-192
- Martindale MQ, Henry JQ (1998) The development of radial and biradial symmetry: The evolution of bilaterality. *Am. Zool.* 38:672-684
- Martindale MQ, Henry JQ (1999) Intracellular fate mapping in a basal metazoan, the ctenophore *Mnemiopsis leidyi*, reveals the origins of mesoderm and the existence of indeterminate cell lineages. *Dev Biol* 214:243-257
- Medina M, Collins AG, Silberman JD, Sogin ML (2001) Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. *Proc Natl Acad Sci U S A* 98:9707-9712
- Mehl D, Reiswig HM (1991) The presence of flagellar vanes in choanocytes of porifera and their possible phylogenetic implications. *Z. Zool Syst Evolutionsforsch* 29:312-319
- Miller DJ, Miles A (1993) Homeobox genes and the zootype. *Nature* 365:215-216
- Miller RL (1971) *Trichoplax adhaerens* Schulze 1883: return of an enigma. *Biol Bull* 141:374
- Mills C (2001) Phylum ctenophora: list of all valid species. <http://faculty.washington.edu/cemills/Ctenolist.html>.
- Minot CS (1883) An apparently new animal Type. *Science* 1:305
- Misevic GN, Schlup V, Burger MM (1990) Larval metamorphosis of *Microciona prolifera*: evidence against the reversal of layers: in *New Perspectives in Sponge Biology*. K. Rützler. Institution Press, Washington, DC
- Monticelli FS (1896) *Adelotacta zoologica*. 2. *Trepoplax reptans* Montic. *Mitt. Aus zool. Station Neaples*. 12:444-462
- Müller CW, Herrmann BG (1997) Crystallographic structure of the T domain-DNA complex of the Brachyury transcription factor. *Nature* 389:884-888

## References

---

- Müller P, Seipel K, Yanze N, Reber-Müller S, Streitwolf-Engel R, Stierwald M, Spring J, Schmid V (2003) Evolutionary aspects of developmentally regulated helix-loop-helix transcription factors in striated muscle of jellyfish. *Dev Biol* 255:216-29
- Müller P, Yanze N, Schmid V, Spring J (1999) The homeobox gene *Otx* of the jellyfish *Podocoryne carnea*: role of a head gene in striated muscle and evolution. *Dev. Biol.* 216:582-594
- Müller WE, Schroder HC, Skorokhod A, Bunz C, Müller IM, Grebenjuk VA (2001) Contribution of sponge genes to unravel the genome of the hypothetical ancestor of Metazoa (Urmetazoa). *Gene* 276:161-73
- Nielsen C (1995) *Animal Evolution: Interrelationships of the Living Phyla*. Oxford Univ. Press, Oxford
- Nielsen C (2001) *Animal Evolution: Interrelationships of the Living Phyla*. 2<sup>nd</sup> ed. Oxford Univ. Press., Oxford
- Odorico DM, Miller DJ (1997) Internal and external relationships of the Cnidaria: implications of primary and predicted secondary structure of the 5'-end of the 23S-like rDNA. *Proc R Soc Lond B Biol Sci* 264:77-82
- Papaioannou VE (2001) T-box genes in development: from hydra to humans. *Int Rev Cytol* 207:1-70
- Paxton C, Zhao H, Chin Y, Langner K, Reecy J (2002) Murine *Tbx2* contains domains that activate and repress gene transcription. *Gene* 283:117-224
- Pendleton JW, Nagai BK, Murtha MT, Ruddle FH (1993) Expansion of the Hox gene family and the evolution of chordates. *Proc. Natl. Acad. Sci. U S A* 90:6300-6304
- Peterson KJ, Eernisse DJ (2001) Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. *Evol Dev* 3:170-205
- Peterson KJ, Harada Y, Cameron RA, Davidson EH (1999) Expression pattern of *Brachyury* and *Not* in the sea urchin: comparative implications for the origins of mesoderm in the basal deuterostomes. *Dev. Biol.* 207:419-431
- Pflugfelder GO, Roth H, Poeck B, Kerscher S, Schwarz H, Jonschker B, Heisenberg M (1992) The lethal(1)optomotor-blind gene of *Drosophila melanogaster* is a major organizer of optic lobe development: isolation and characterization of the gene. *Proc Natl Acad Sci U S A* 89:1199-203
- Philippe H, Chenuil A, Adoutte A (1994) Can the cambrian explosion be inferred through molecular phylogeny? *Development* 120 (supplement):15-26
- Podar M, Haddock SH, Sogin ML, Harbison GR (2001) A molecular phylogenetic framework for the phylum Ctenophora using 18S rRNA genes. *Mol. Phylogenet. Evol.* 21:218-230
- Poock B, Hofbauer A, Pflugfelder GO (1993) Expression of the *Drosophila* optomotor-blind gene transcript in neuronal and glial cells of the developing nervous system. *Development* 117:1017-29
- Pollard SL, Holland PWH (2000) Evidence for 14 homeobox gene clusters in human genome ancestry. *Curr. Biol.* 10:1059-1062
- Porsch M, Hofmeyer K, Bausenwein BS, Grimm S, Weber BH, Miassod R, Pflugfelder GO (1998) Isolation of a *Drosophila* T-box gene closely related to human *TBX1*. *Gene* 212:237-48

## References

---

- Ruppert EE, Barnes RD (1994) *Invertebrate zoology* (Sixth Edition). Saunders College Publishing
- Russ AP, Watter S, Colledge WH, Aparicio SA, Carlton MB, Pearce JJ, Barton SC, Surani MA, Ryan K, Nehls MC, Wilson V, Evans MJ (2000) Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature* 404:95-99
- Ruthmann A, Behrendt G, Wahl R (1986) The ventral epithelium of *Trichoplax adhaerens* (Placozoa): cytoskeletal structures, cell contacts and endocytosis. *Zoomorphology* 106:115-122
- Ruthmann A, Grell KG, Benwitz G (1981) Dna content and fragmentation of egg nucleus of *trichoplax adhaerens*. *Z. Naturforsch.* 36:564-567
- Ruthmann A, Wenderoth H (1975) Der DNA-gehalt der Zellen bei dem primitiven Metazoon *Trichoplax adhaerens* F. E. Schulze. *Cytobiologie*:421-431
- Ruvinsky I, Silver LM, Gibson-Brown JJ (2000) Phylogenetic analysis of T-Box genes demonstrates the importance of amphioxus for understanding evolution of the vertebrate genome. *Genetics* 156:1249-1257
- Ruvinsky I, Silver LM, Ho RK (1998) Characterization of the zebrafish *tbx16* gene and evolution of the vertebrate T-box family. *Dev Genes Evol* 208:94-9
- Ryan K, Garrett N, Mitchell A, Gurdon JB (1996) Eomesodermin, a key early gene in *Xenopus* mesoderm differentiation. *Cell* 87:989-1000
- Schacke H, Schroder HC, Gamulin V, Rinkevich B, Müller IM, Müller WE (1994) Molecular cloning of a tyrosine kinase gene from the marine sponge *Geodia cydonium*: a new member belonging to the receptor tyrosine kinase class II family. *Mol Membr Biol* 11:101-107
- Schierwater B, Kuhn K (1998) Homology of Hox genes and the zootype concept in early metazoan evolution. *Mol Phylogenet Evol* 9:375-81
- Schlichter D (1975) The importance of dissolved organic compounds in sea water for the nutrition of *Anemonia sulcata* Pennant (Coelenterata). in H. Barnes (ed.). *Proceedings of the 9th European Marine Biological Symposium*:395-405
- Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18:502-504
- Scholz CB, Technau U (2003) The ancestral role of Brachyury: expression of *NemBra1* in the basal cnidarian *Nematostella vectensis* (Anthozoa). *Dev Genes Evol* 212:563-70
- Schubotz H (1912) Ist *Trichoplax* die umgewandelte Planula einer Hydromeduse? *Zool Anz* 39:582-585
- Schuchert P (1993a) *Trichoplax adhaerens* (Phylum Placozoa) has cells that react with antibodies against the neuropeptide RFamide. *Acta Zoologica (Stockholm)* 74:115-117
- Schuchert P (1993b) Phylogenetic analysis of the Cnidaria. *Zool Syst Evolut-forsch* 31:161-173
- Schuetze J, Krasko A, Custodio MR, Efremova SM, Mueller IM, Mueller WEG (1999) Evolutionary relationships of Metazoa within the eukaryotes based on molecular data from Porifera. *Proc R Soc Biol Sci Ser B* 266:63-73

## References

---

- Schulte-Merker S, van Eeden FM, Halpern ME, Kimmel CB, Nüsslein-Volhard C (1994) No tail (ntl) is the zebrafish homologue of the mouse T (Brachyury) gene. *Development* 120:1009-1015
- Schulze FE (1883) *Trichoplax adhaerens*, nov. gen., nov. spec. *Zool Anz* 6:92-97
- Schwartz V (1984) The radial polar pattern of differentiation in *Trichoplax adhaerens* F. E. Schulze. *Z Naturforsch* 39c:818-832
- Simon HG, Kittappa R, Khan PA, Tsilfidis C, Liversage RA, Oppenheimer S (1997) A novel family of T-box genes in urodele amphibian limb development and regeneration: candidate genes involved in vertebrate forelimb/hindlimb patterning. *Development* 124:1355-66
- Simpson T, L. (1984) *The Cell Biology of Sponges*. Springer Verlag, New York
- Snell EA, Furlong RF, Holland PW (2001) Hsp70 sequences indicate that choanoflagellates are closely related to animals. *Curr Biol* 11:967-70
- Spring J, Yanze N, Josch C, Middel AM, Wittinger B, Schmid V (2002) Conservation of Brachyury, Mef2, and Snail in the myogenic lineage of jellyfish: a connection to the mesoderm of bilateria. *Dev. Biol.* 244:372-384
- Spring J, Yanze N, Middel AM, Stierwald M, Gröger H, Schmid V (2000) The mesoderm specification factor twist in the life cycle of jellyfish. *Dev Biol* 228:363-75
- Stein S, Niss K, Kessel M (1996) Differential activation of the clustered homeobox genes CNOT2 and CNOT1 during notogenesis in the chick. *Dev Biol* 180:519-33
- Syed T, Schierwater B (2002) The evolution of the Placozoa: a new morphological model. *Senckenbergiana lethaea* 82:315-24
- Takabatake Y, Takabatake T, Takeshima K (2000) Conserved and divergent expression of T-box genes Tbx2-Tbx5 in *Xenopus*. *Mech Dev* 91:433-7
- Takada N, Goto T, Satoh N (2002) Expression pattern of the Brachyury gene in the arrow worm *Paraspadella gotoi* (Chaetognatha). *Genesis* 32:240-245
- Talbot WS, Trevarrow B, Halpern ME, Melby AE, Farr G, Postlethwait JH, Jowett T, Kimmel CB, Kimelman D (1995) A homeobox gene essential for zebrafish notochord development. *Nature* 378:150-7
- Tamura K, Yonei-Tamura S, Belmonte JC (1999) Differential expression of Tbx4 and Tbx5 in Zebrafish fin buds. *Mech Dev* 87:181-4
- Tardent P (1978) Coelenterata, Cnidaria. In *Morphogenese der Tiere, Erste Reihe, Deskriptive Morphogenese*. Lieferung 1, A-I. Edited by F. Seidel. Gustav Fischer Verlag, Jena
- Technau U (2001) Brachyury, the blastopore and the evolution of the mesoderm. *BioEssays* 23:788-794
- Technau U, Bode HR (1999) HyBra1, a Brachyury homologue, acts during head formation in *Hydra*. *Development* 126:999-1010
- The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282:2012-2018

## References

---

- Thiemann M, Ruthmann A (1989) Microfilaments and microtubules in isolated fiber cells of *Trichoplax adhaerens* (Placozoa). *Zoomorphology* 109:89-96
- Thiemann M, Ruthmann A (1991) Alternative modes of asexual reproduction in *Trichoplax adhaerens* (Placozoa). *Zoomorphology* 110:165-174
- von Dassow G, Schmidt JE, Kimelman D (1993) Induction of the *Xenopus* organizer: expression and regulation of *Xnot*, a novel FGF and activin-regulated homeo box gene. *Genes Dev* 7:355-66
- Wainright PO, Hinkle G, Sogin ML, Stickel SK (1993) Monophyletic origins of the metazoa: an evolutionary link with fungi. *Science* 260:340-2
- Walker C, Steisinger G (1983) Induction of mutations by gamma-rays in pregonial germ cells of zebrafish embryos. *Genetics* 103:125-136
- Wattler S, Russ A, Evans M, Nehls M (1998) A combined analysis of genomic and primary protein structure defines the phylogenetic relationship of new members of the T-box family. *Genomics* 48:24-33
- Wimmer W, Blumbach B, Diehl-Seifert B, Koziol C, Batel R, Steffen R, Müller IM, Müller WE (1999) Increased expression of integrin and receptor tyrosine kinase genes during autograft fusion in the sponge *Geodia cydonium*. *Cell Adhes Commun* 7:111-24
- Woollard A, Hodgkin J (2000) The *Caenorhabditis elegans* fate-determining gene *mab-9* encodes a T-box protein required to pattern the posterior hindgut. *Genes Dev* 14:596-603
- Yamada A, Martindale MQ (2002) Expression of the ctenophore Brain Factor 1 forkhead gene ortholog (*ctenoBF-1*) mRNA is restricted to the presumptive mouth and feeding apparatus: implications for axial organization in the Metazoa. *Dev Genes Evol* 212:338-348
- Yamamoto A, Amacher SL, Kim SH, Geisbert D, Kimmel CB, De Robertis ERM (1998) Zebrafish paraxial protocadherin is a downstream target of spadetail involved in morphogenesis of gastrula mesoderm. *Development* 125:3389-3397
- Yanze N, Gröger H, Müller P, Schmid V (1999) Reversible inactivation of cell-type-specific regulatory and structural genes in migrating isolated striated muscle cells of jellyfish. *Dev Biol* 213:194-201
- Yanze N, Spring J, Schmidli C, Schmid V (2001) Conservation of Hox/ParaHox-related genes in the early development of a cnidarian. *Dev. Biol.* 236:89-98
- Yasuo H, Lemaire P (2001) Role of Goosecoid, *Xnot* and Wnt antagonists in the maintenance of the notochord genetic programme in *Xenopus* gastrulae. *Development* 128:3783-93
- Yi CH, Russ A, Brook JD (2000) Virtual cloning and physical mapping of a human T-box gene, *TBX4*. *Genomics* 67:92-5
- Zhang J, King ML (1996) *Xenopus* VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. *Development* 122:4119-29

*Axinella verrucosa* Tbx1/15/20

The T-box domain is underlined, the arrow represent the specic primer.

AxTbx F2 →

1 ACTAAAGCTGCGGACGCTCTCGTTGTAAACACAAAGACAATTATACTGCACCTTGCAGCTTGGAGGACTCGTCTCGATTGAGCAAGTTTGTGAATTTGCAGG 100

101 TCCTGAGGCAAGAAGTTAGCAAATTCAGCCCTCAAGTCAAACCTTGCAAACCTACAAGACAAGGAAAAATCAGCCACAATGACAGAAGACTACGACTTGA 200  
1 M T E D Y D L T 8

201 CTATCATGGATACTATGGAGGCAAACCTCTCCCGAATCATCCACAATGAACCTAGATTACCCACCTTCTCTGGAAATGAGATCGATGGACAGCGAGTACCT 300  
9 I M D T M E A N S P E S S T N E L D S P P S L E M R S M D S E Y L 41

AxTbx F3 →

301 ACCTGATGAGAAGTCCACTGCAGCAATCACCAGATAGCAAGATGGCAGTGGCACTGCAGCACCAGAGCAGAAAACTGTGGAACAGTTTGACTCAGTT 400  
42 P D E K V H C S N H P D S K M A V A L Q H Q S R K L W K O F D S V 74

401 GGCCAGGAGATGATAGTACCAGGAGAGAAACGGATGTTCCCTCCAATCAAAGTTGAGGTATTTGGCCTGGACCAATGCCTACTACGTTCTACTGA 500  
75 G T E M I V T R R G R R M F P P I K V E V F G L D P N A Y Y V L L M 108

501 TGGACATGGCCCTGTTGACAAGTACCGCTACAATAACAAAACCTCATCCTGGGTGAAATGCTTTGAAGAAGAGTGTCTCCGACACGGTTGTACGTACA 600  
109 D M A P V D K Y R Y K Y Q N S S W V K C F E E C S P T R L Y V H 141

AxTbx F1 →

601 CCCCATTCCCGGCTCTTGGCTCGCACTGGATGCACATGATCATCAACTTCTACAAGCTCAAGCTCAGCAACAACAGCTGGACAAAAGGGCCATATC 700  
142 P D S P A L G S H W M H M I I N F Y K L K L T N N O L D K O G H I 174

701 ATTGTAAACAGTATGCACCGCTACCAGCCACGGTTGCACATTGTGCGAGTCTCAAGACTGCAACAACCTGAACCTGGGAGAACTCCACACATTCTCTCC 800  
175 I V N S M H R Y Q P R L H I V E S O D C N N L N W E N F H T F I F P 208

AxTbx R1 ←

801 CCAGGACACAGTTCCACCACCGTCACTGCTTACCAGAATGACAAGATCACTCAGTTGAAGATTGAGAACAACCCCTTTGCCAAAGGATTCGCTCTCCGTT 900  
209 R T O F T T V T A Y O N D K I T O L K I E N N P F A K G F R S P F 241

901 TTCACGCTCAGAACACGATCCTCCCAAGCAGGTTTCCCTGGGGTCCAACAGCGACTCAGGGGCCGATGATAGTAGAAGCCCATCTATCATCCTTC 1000  
242 S R C R T R I L P K A G F P W G P T A T Q G P M I V E A H L S S F 274

1001 CTCCTCACTTCTCTTGCCTCTAACTCAAACAAGATCACATGCCAGTGGCCAAAACCGCAAGTCCGATGATGACAACAATGCTACTAGTCCAAGAGC 1100  
275 L L T S S C L L T S N K I T C Q S A K N A S P M M T T M L L V Q E Q 308

1101 AGATTGCTTGACAGCTGCAGCTCCAGAACAGGGGTCATCAGCATGGCCACCCCACTCATCCCCAACCAACTTCCACTGCAGTAGTCACTGCCCC 1200  
309 I C L T A A A P E Q G V I S I A T P Q L I P N P T S T A V V T A P 341

1201 CAGTTCACCCACTGTCCGGAATAGTGTCTGTGTACATGTCTACAAGCCAGTACTTCACTCAGCCGAGAGCAGTCTCCGAAACGTGTAGAGCCACC 1300  
342 S S P T V A N S A P V Y I V Y K P V L Q S A E K Q S P K R V E P T 374

1301 ACAACCCTGTGATCCATCCCGGGATGAATGGCACTACGTTTATCCAGGGAAGCAATGGGCACATCCTCAACGGTGCCACCATCATTCAAGCCCAGCCAC 1400  
375 T T T V I H P G M N G T T F I Q G S N G H I L N G A T I I Q A Q P P 408

1401 CCACCAGCCCGAGTACAATCTGCACAGCTGCTGTGCCCCAGTGGTTCGACCACAAGTTCTTTCTCCGAGTGTGGAGCTGCAGTTCCAGAATCTTAATGG 1500  
409 T S P S T I C T A A V A P V V R P Q V L S P S V G A A V Q N L N G 441

1501 TAGTACTGTTAGTACCGCCATTGTCCAGCCACTCCAACCTTCTCAATGGGAGCAGCAGATTGTTCAATGCAACAGGAGGAGCAGATTATCCATCCT 1600  
442 S T V S T A I V Q A T P T F L N G T T A L F N A N R Q A A F I H P 474

1601 CATAGCATGGCAAATGGACTTAGCTTTGCTGGCACATCCTCTCAGCTTGTGACAACCAACAGCGGCTCCGAGAATTACAGAGCAGTTTCTGAAGCCAAAA 1700  
475 H S I A N G L S F A G T S S Q L V T T T A A P R I T E A V S E A K R 508

AxTbx R4 ←

1701 GGAGCCATGGTCTTGCACATATTGACATCAGAAGCCTTCAAAGAGGGCAATGTGCAGCTCGCGAAATCCCGGCTCCACCTCTTGTGAAAGTCAAGTGA 1800  
509 S H G P C T I D I R S L T K E G N V H V A K S R P P P L V E V S E 541

1801 AGTTAAGAGAAAATCATCCGAGAAAAGTTTCCAAACCCACAGTACGACAGCATGAACTGGAAGGCTTGACTTCCATCTCAGCTAGCAACACTGCAGCA 1900  
542 V K R K S S R E K F S N P T V R Q H E L E G L T S I S A S N T A A 574

1901 ACTCAGAGCATACTGGTCCAGCACCTACAATGGAATGCAGAAGGTGTTTATTGGTGGGAAAAATCAGACTGTCTGTTACCATTTTCATCCCGGCAC 2000  
575 T Q S I L V P A P T N G M Q K V F I V A E K S D C P G T I F I P A P 608

2001 CAACAGCACCTGCCGCTGTGATGCTGCAGCTATGAACAACACAGACGGTATTCTTATGGCATATGATCAGCAAGTAGTTAATACCATGCCAATATATC 2100  
609 T A P A A V D A A A M N N T D G Y S Y G I \* 629

AxTbx R3 ←

2101 GATTGGTGCCCTCAACACACTACAGCCAATCAAATCCTTGCAACACCGGTTGCAGCTGGTCCGAATGCTTCCCTGAGTTTATTTAAGGCTCCAGCTAC 2200

2201 ATTTTATGTTGAAAATATCAGCTCTTTAACTTAGTGGAGGAGCTGGTCTGTATAGAGTTCAAACGCATTCTTGGCATTCTTCTTCAAAGTGCTGTGAA 2300

2301 CCTTCTGTGTAACACCCCTGGTTCAGACTGTAGGTCTATAAACTGCAGTGTATGCCCATTTTTTGCAACCCAGATTTTGGGAGTGTGTATCTT 2400

2401 TCTTATTTGCATCACTCATAACTGTACACTTGTTCCTTCCATGCTTAATATATTTTTTGCAGTTCAATAAGCAAGAGGAATGGTCCAGAATTTCCATTT 2500

2501 GTCTGTATTAGTTTCGTAACCTCCAGCTTATAGCTAACATACATGTGTGTGTTGCATAGCTGTACTGTATCTCTTCCAACACAATGACATTTTCT 2600

2601 ACTGGCA<sub>(20)</sub> 2607 ← AxTbx R2

*Axinella verrucosa* Elongation factor-1-alpha

The arrow represent the specific primer

	AxEf F3	→	
1	GATGCGCAACCCAAGTTCCTCTTTCTCTGACGTGCGAAGTTCCTGGCAAGTCGGGACACTGTAACCTGCGTAGTGAATC		80
81	TACCTCAATCGTTCTACCCATTCAATCCAGAGCCTCTGAAACGAGACAGTCACCACCATGCCGAAGGAGAAGATTCA		160
1		M P K E K I H I	8
161	TCAATATCGTGGTTATCGGCCACGTAGACTCCGGGAAATCCACCACCACCGGCCATCTCATCTACAAGTCCGGCGGTATC		240
9	N I V V I G H V D S G K S T T T G H L I Y K C G G I		34
	AxEf F2	→	
241	GACAAGAGAACGATCGAGAAGTTCGAGAAGGAAGCGCAGGAGATGGGCAAGGGTTCGTCAAGTATGCCTGGGTGTGGAA		320
35	D K R T I E K F E K E A Q E M G K G S F K Y A W V L D		61
321	CAAATAAAAGCAGAGCGTGAACGTGGAATCACCATCGACATTGCACATGGAAGTTGAGACCACCAATACTACATCA		400
62	K L K A E R E R G I T I D I A L W K F E T T K Y Y I T		88
401	CTGTAATTGACCCCCAGGACATCGTGACTTCATCAAGAACATGATTACAGGAACATCACAAGCTGACTGTGCTGTATTG		480
89	V I D A P G H R D F I K N M I T G T S Q A D C A V L		114
481	ATTGTGGCTGTGGCAGAGTGAATTTGAAGCAGGCATCTCCAAGAACGGGCAAACAAGAGAGCATGCGTGTGGCGTA		560
115	I V A A G T G E F E A G I S K N G Q T R E H A L L A Y		141
561	CACCCTCGGCGTGAAGCAGCTCATCGTAGGTGTCACAAGATGGACAGCAGAGCCACCATAAGTGTGAGCCTCGTTTCA		640
142	T L G V K Q L I V G V N K M D S T E P P Y S E P R F N		168
641	ACGAGATCATCAAGAAGTGGGGCCATCAAGAAGATTGGCTACAACCCGAAAGCTGTTGCCTTTGTTCCCATCTCT		720
169	E I I K E V G A Y I K K I G Y N P K A V A F V P I S		194
721	GGTTGGCATGGTGACAACATGCTTGAGCCTAGTGACAATATGTCATGGTTCAAGGGATGGAATGTTGAACGGAAGGAGGG		800
195	G W H G D N M L E P S D N M S W F K G W N V E R K E G		221
801	GAACGCAACCGTAAGACTTTGTTACATGCTCTCGATTCCATCCTCCCTCCAAGAGAGCTACTGAGAAGCCCCTCAGAC		880
222	N A T G K T L F T C L D S I L P P K R P T E K P L R L		248
	AxEf R2	←	
881	TTCCTCTCCAGGACGCTCTACAAGATTGGAGGTATTGGCACGGTTCAGTGGGCGGTGTTGAGACCGGTATCCTGAAACCT		960
249	P L Q D V Y K I G G I G T V P V G R V E T G I L K P		274
	AxEf F1	→	
961	GGAATGGTGGTACCATCTCTCTGCCAACATCACCACGGAGGTGAAGTCTGTGAGATGCATCAGGATCGCTTACCGA		1040
275	G M V V T I S P A N I T T E V K S V E M H H E S L T E		301
1041	GGCTTCCCGGAGACAACGTGGGATTCACATCAAGAACGTGTCTGTGAAGGAGTTCGTCGCGGTATGGTTCCCGGGG		1120
302	A L P G D N V G F N I K N V S V K E V R R G M V A G D		328
1121	ACTCCAAGAATGACCCCCGAAGGCTGCCAAGAGCTTCTTCTGCTCAGGTTATCATCCTAAATCACCTGCGCAGATCCAT		1200
329	S K N D P P K A A K S F L A Q V I I L N H P G Q I H		354
	AxEf R1	←	
1201	GCTGGCTACTCCCCTGTGCTTGATTGCCACACTGCCCATATTGCATGCAAGTTCGCTGAACTGAAGGAAAAGTGTGACCG		1280
355	A G Y S P V L D C H T A H I A C K F A E L K E K C D R		381
1281	GAGAAGTGGGAAGAAGCTGGAAGACAATCCCAAGTCCGTAAAGAGCGGGACGCTGCCATCGTCGAGCTTGTCCCAAGCA		1360
382	R S G K K L E D N P K S V K S G D A A I V E L V P S K		408
1361	AGCCAATGTGTGGAGGCTTTTTCAGAATATCCCCCTCTCGGCAGATTGTCAGTTCGTGACATGAAGCAAACGGTTGCC		1440
409	P M C V E A F S E Y P P L G R F A V R D M K Q T V A		434
1441	GTTGGCGTTATCAAGTCTGTGAGAAGCAGGAGGGCAGGGTGGCAAAGTAACAAGGAGCTCAGAAGCGCAGAAGAA		1520
435	V G V I K S V E K Q E G K G G K V T K A A Q K A Q K K		461
1521	GAAGTGATGAAGGGTATATTATAGGACGCTAATTGTTAGGAACATTATATTATCTTGTTCACATTCATTATTGTAT		1600
462	K *		462
1601	GTATATAATGGTTTCCTTTCAATAATGGCCATCAATTCCATTCTGA <sub>(20)</sub>		1648
	AxEf R3	←	

*Podocoryne carnea* Aristaless

The arrows indicate the specific primers, the homocedomain is underlined.

1	AGTTGAACTTACCAGACATGTTCCAGTGGTCAAGAAATTCATCGTACGCTTCCTGTTGTCGAGGAAGTTGCTGCACTGAA	80
1	M F Q W S R N S S Y A S C C R G S C C T E	21
	ArxF1 →	
81	AACATATCAAACGAAATGGAAGAAGAAATATTACTCAATGTGCAGATCAAAAACAAGAAGAAATCGAACAACATTCACAAA	160
22	N I S N E M E E E Y Y S M C R S K Q R R N R T T F T K	48
	← ArxR2	
161	ACAACAACCTCAAGAACTTGAGATCACTTCCAAAAGAAACTATCCTGATATAAATACAAGAGAAGCATTAGCTGAAA	240
49	Q Q L Q E L E I T F Q K K H Y P D I N T R E A L A E K	75
241	AGATTGGAATTACTGAAGCAAGAATACAGGTATGGTTTCAAACCGTCGTGCCAAATGGAGAAAGTTATCGTCCGCAAAA	320
76	I G I T E A R I Q V W F Q N R R A K W R K L S S A K	101
	ArxF2 →	← ArxR1
321	CAACGCATCCTACAAAGAAATGCCCAAATACGTTCAAGTCACGATTATTTACTTGAACATCCATGGTCACCCGTTCCAAG	400
102	Q R I L Q R N A Q I R S S H D Y L L E H P W S P V P S	128
401	CAGTGCATCTCAGTTAGTTCTCACAGATAGAATGAAGCATGAACGTCCAGATCATTGTGCGCCGCTTTAGCAACTTCAA	480
129	S A S Q L V L T D R M K H E R P D H L S P P L A T S T	155
481	CACATAGAAGTCTGTCAAACCTGGAAGACCAGAAGAAGATGGTGTGCAAGAGGTCAATTATACAATGAATAGGTTGAAA	560
156	H R S P V K L E R P E E D G V E E V N Y T M N R L K	181
561	CCAGCAGAAAAAATGTTTTAGAAAGAATCCCAATCAGACAAAATCTGAATTTGATCGGATATATCCCGAAAGACATCG	640
182	P A E K N V L E R I P I R Q N S E F D R I Y P E R H R	208
641	GAAGTATCCCGAATATCGCCATGAAGTTCATCGTACAGAACTGATTCTCCTTCAAGAATAGAATCTAGACATCATTGGA	720
209	K Y P E Y R H E V H R T E T D S P S R I E S R H H S K	235
721	AACAATTTCCACCAATGATGTCAATTCATATCACCTGCCAATACCGTGTCCATGTAAGGTTGCTATATGAATCCCAT	800
236	Q F P P M M S F P Y H L P I P C P C K S C Y M N P I	261
	← ArxR3	
801	GAAATGATGTATAAAAATCATCACTCAAGACACCCTTCTTTTCGTCAAATCGTCTCATAATCGCTCAAAAACATTGAAAG	880
262	E M M Y K N H H S R H P S F V K S S H T S L K N I E R	288
881	AAAAACATACGGAGAGGAGAGAGACACGCGATGACATGCTCGAGGATTA AACATAAAAAACAACGCGTTTCGaGTAA	960
289	K T Y G E E R E T R D D M L E D *	304
961	CCAAACAAAACAATAACAACCTACGAACTCTTTTATTGTACAAAATCTTGCTTGCAATAAAGCGTAACTTCTTGCGAT	1040
1041	GTAATTTTAAATTATTTCGCAATACATTGGCATTAAAAGTTA	1081



*Podocoryne carnea* Elongation factor-1-alpha

The arrows represent the specific primers.

	PcEfF5 →	
1	ATCTAATTCAATTCAATATGGCCCCAAAAGAAAAGCCAAGATACACATCAACATTGTTGTAATTGGTCATGTCGATT	80
1	M A P K E K A K I H I N I V V I G H V D S	21
81	TGGAAAATCAACAAGTACTGGACATTGATTACAAATGTGGTGGTATTGATAAACGAACCATCGAAAAATTCGAGAAGG	160
22	G K S T S T G H L I Y K C G G I D K R T I E K F E K E	48
	PcEfF7 →	
161	AAGCCCAGGAAATGGGTAAAGTTCTTTCAAATACGCCTGGGTCTTGGACAAATTGAAGGCAGAACGTGAACGTGGTATT	240
49	A Q E M G K G S F K Y A W V L D K L K A E R E R G I	74
241	ACTATTGATATTGCTTTATGGAAATTCGAAAACAACCAAAATTTACGTGACAATTATTGATGCTCCTGGACATCGTGATT	320
75	T I D I A L W K F E T T K F Y V T I I D A P G H R D F	101
321	CATCAAGAACATGATCACTGGTACATCACAAAGCTGATTGTGCTGTGTTGATTGTTGCTTCGAGTACTGGTGAATTCGAGG	400
102	I K N M I T G T S Q A D C A V L I V A S S T G E F E A	128
401	CTGGTATTCTAAAAATGGACAGACTCGTGAACATGCCTTGTGGCTACACTTTGGGTGTGAAAACAACTTATCATTGGT	480
129	G I S K N G Q T R E H A L L A Y T L G V K Q L I I G	154
481	GTAATAAAAATCGATAATACTGAGCCCCATACAGTGAGGCTCGATTCAATGAAATCAAGAAAGAAGTCGAAGGATATGT	560
155	V N K I D N T E P P Y S E A R F N E I K K E V E G Y V	181
561	AAAAAAGTGGTTACAACCCAAAAGCAGTTGCTATCCTCCCAATTTCTGGTTGGCATGGAGATAACATGTTGGAAGAAT	640
182	K K V G Y N P K A V A I L P I S G W H G D N M L E E S	208
	PcEfR5 ←	
641	CGACCAAAATCCAATGGTTCAAAGTTGGGCAACTGAACGAGTAGACGAGGACAAAAAAGTGATAAACAGCTCAGGCAAA	720
209	T K S K W F K G W A T E R V D E D K K V I N S S G K	234
721	ACATTGTTGGAAGCATTGGATGCTATTGTGCCACCTTACGACCATCGAATAAACCTCTCCGTCTGCCACTTCAAGACGT	800
235	T L F E A L D A I V P P S R P S N K P L R L P L Q D V	261
801	ATACAAAATGGAGGTATTGGTACAGTGCCAGTTGGCAGAGTTGAAACAGGTAATAATCATGCCTGGTATGGTTGTTACAT	880
262	Y K I G G I G T V P V G R V E T G K I M P G M V V T F	288
881	TTGCTCCATGTGGCATCTCCACTGAAGTCAAATCTGTTGAAATGCATCATACTGCTATGCCAGAAGCTTTACCTGGTGAC	960
289	A P C G I S T E V K S V E M H H T A M P E A L P G D	314
	EF1aF →	
961	AATGTTGGCTTCAATGTCAAGAATGTGTCAATCAAAGAAATTAACGTTGGTATGGTTGGCTCTGACAGTAAAAATGATCC	1040
315	N V G F N V K N V S I K E I K R G M V A S D S K N D P	341
	PcEf For6 →	
1041	TGCCAAGGAAGCCAAGACCTTTTACGCCAGGTCAATCCTTGAACCATCCCGGTGAGATCCACGCTGGTTACCAACCTG	1120
342	A K E A K T F Y A Q V I I L N H P G E I H A G Y Q P V	368
	PcEf Rev4 ←	
1121	TGCTCGATTGTCACACCGCTCATGTTGCTTGCAAGTTCACCTGAATTGAAACAGAAATGTGACAGACGTAGTGGCAAAGTT	1200
369	L D C H T A H V A C K F T E L K Q K C D R R S G K V	394
	EF1c ←	
1201	TTGGAAGAAAATCCTAAACTTGTAAATCAGGTGATGCTGCCATGATCACTTTGACACCATCAAAACCAATGTGTGTTGA	1280
395	L E E N P K L V K S G D A A M I T L T P S K P M C V E	421
	EF1aR ←	
1281	AGCTTTTTCAGACTATGCACCATTAGGACGTTTCGCTGTAAGAGATATGAGACAAACCGTAGCCGTTGGCGTTATCAAAT	1360
422	A F S D Y A P L G R F A V R D M R Q T V A V G V I K S	448
1361	CAGTAGAGAAGGCCAAGCTGGTGGTAAAATGACAAAATCTGCTCAAAAAGCAACTGGTGGCAAGAAGAAGTGACATCCT	1440
449	V E K A E A G G K M T K S A Q K A T G G K K K *	471
	PcEfR3 ←	
1441	TCATTGAAATATTGTGCTATGTTATTCCGAACTAGAATTTTTCATGAACACTTATGAAGGAACTGTAATCTGTCAATTG	1520
1521	ATTCGATGTGA 1531	

*Podocoryne carnea* Tbx4/5

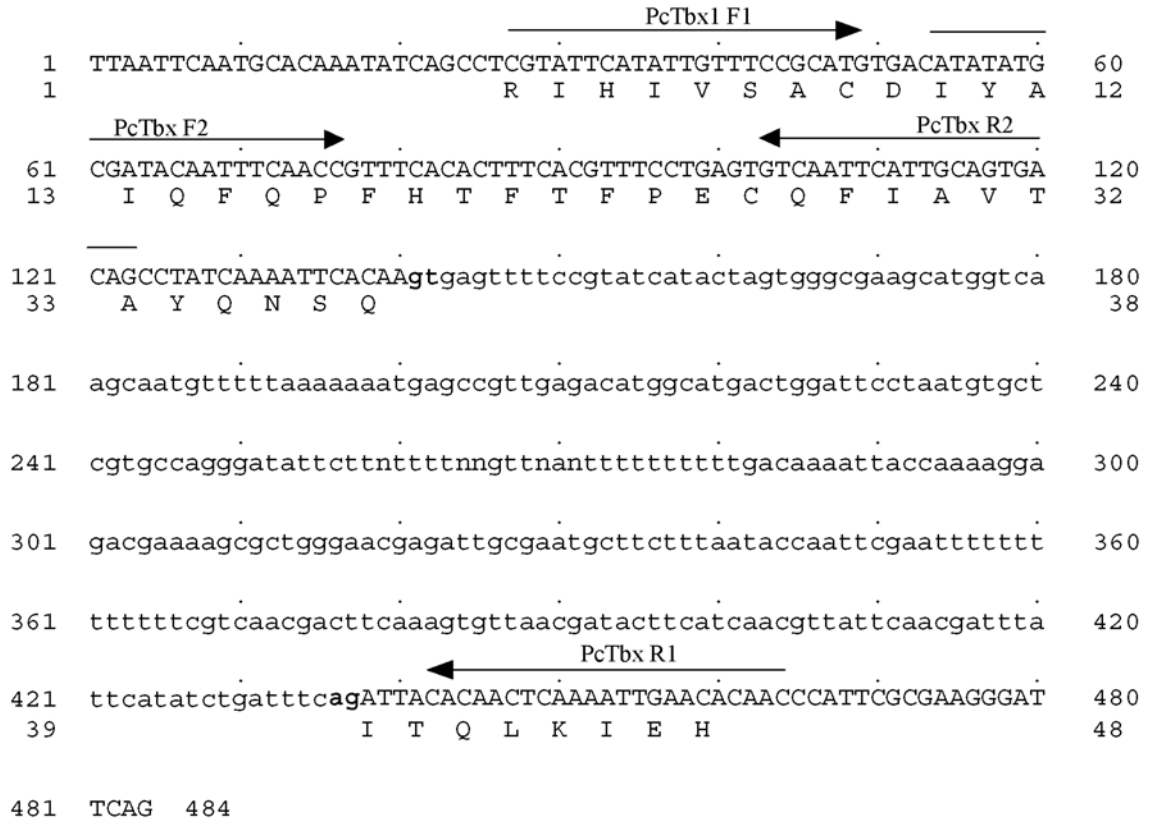
The arrows represent the specific primers, the T-box domain is underlined.

	PcTbx2 F3	→	
1	TATTTACTTTTAAAAATAGTTTTTAATAGTGATATATATATTGCAAAAAGTTAATCTCTCGTCAAACAGTACTCGACCCAG		80
81	GAGGCTGGATTAAAATGGCAGACAAGAAAAAGAAACAAACATCTTCTTTATGTATACAAGGTACTCCTGGAGCCCGTCAG		160
1	M A D K K T E Q T S S L C I Q G T P G A R Q		22
161	TATCGGCGGAATTTGCGGCCGCTCATTGCTGCATCAAACAAGTTCAAACAAGAACCAGAATATACTGATTGAAACCCAC		240
23	Y R R N L R P L I A A S N K F K Q E P E Y T D L K P T		49
		← PcTbx2 R3	
241	AACGGATTCTGAAGAACGACATGGTTCAAAACCTCATAATCATATCAACGAAGCAGCGCAGTATCTACAACAACCATCGC		320
50	T D S K N D M V Q N L T N H I N E A A Q Y L Q Q P S Q		76
321	AAGTGACCGGTGATGGCGAAGTTACGCATTATGGTGCAGCAAAAACGCGAAGACGCGAACGTTCTGTGCATCTACAAAAC		400
77	V T G D G E V T H Y G A Q K R E D A N V R V I L Q N		102
401	AAAGAGTTGTGGAGCAAATTCATAGTGTGGGTACGGAGATGATCATTACGAAAGCCGGAAGGAGAATGTTTCCAGTGAT		480
103	K E L W S K F H S V G T E M I I T K A G R R M F P V I		129
481	CAAAGTGAATATTTCTGGATTGAATCCAAAATTGAAATATATCTTGTGTGGATGTTGTACCGGTAGACGACAATCGCT		560
130	K V N I S G L N P K L K Y I L V M D V V P V D D N R Y		156
561	ACAAATATCATAATTCAGAATGGACTGTAGCGGGGAAAGCTGAGCCACACCTACCGGGCAGATTATATGTCCACCCGGAT		640
157	K Y H N S E W T V A G K A E P H L P G R L Y V H P D		182
641	GGTCCTTCCACCGGTGCACAGTGGATGCGACAAACTGTTAGTTTTTCAGAAAAGTAAACTTACAAATAATCATTGGATCA		720
183	G P S T G A Q W M R Q T V S F O K V K L T N N H L D O		209
		PcTbx2 F1	
721	GTTTGAGCATGTGATTTGAATTCATGCACAAGTACCAGCCTCGTATTCATATGTACAAGCCGATGATAATCCCGAG		800
210	F E H V I L N S M H K Y O P R I H I V O A D D N S A E		236
	PcTbx2 F2	→	
801	AAGCATTACGTAAGAGTACTTTCACGACACACGTGTTTACAGAAACCGAGCTGATAGCAGTCACTGCGTATCAAAGCCCA		880
237	A L R K S T F T T H V F T E T E L I A V T A Y O S P		262
		← PcTbx2 R2	
881	CGTATCACACAACCTGAAAATGAAGACAATCCATTGCGAAAAGGATTTCTGTGGCGCGTGAACACGGATTATCATGGTAT		960
263	R I T Q L K I E D N P F A K G F R G A C N T D Y H G M		289
		← PcTbx2 R1	
961	GAAAAGATATCAGACCAAGACTTACTATACGCTTCAAAGAGAGCATATTTCTCCATCATACAATACCGTGGGTTTACCTA		1040
290	K R Y H D Q D L L Y A S K R A Y S P S Y N T V G L P S		316
1041	GTCTACAGTGTACATGGGTGGAGGAATGTTTCATGCAGAGGCCCGGTTTATTTCCAGATATTTTATGGAAATGCAGCACGC		1120
317	L Q C H M G G G M F M Q R P G Y I P D I Y G N A A R		342
1121	TACAGTTCAAGCGAGTATCCATATGCTATCTCACCATACATGTGCAAAAACGTTAATCAAATACAACCGTGGCATGTGAC		1200
343	Y S S S E Y P Y A I S P Y M S K T C N Q I Q P W H V T		369
1201	AGATGACCAGAGACCTGGCAGTCTTCGCCAAAACCAAACGACATATGTAATCTCCAAGTTTAGCAGTTCCTCCTCTAT		1280
370	D D Q R P G S S S P N S N D I C N P P S L A V P P L S		396
1281	CGCATCAGCCGGTCTTCTCCTCATATCTCTGTCACTGGGTACGGCGTTACTCCATCTCAGATGCCATTTTATTTCGTCC		1360
397	H Q P V L P S Y S L S P G Y G V T P S Q M P F Y S S		422
1361	ACCAGTCCAACCTCAACACGGAACAAGTTATACAACCTCAGGAACCAAAGTACCTTATAACGCATGGCAACCGTCTGAGAC		1440
423	T S P T Q H G T S Y T T Q E P K S P Y N A W Q P S E T		449
1441	TATATACCCATACCATTAAGACAGCACAGGTTGAAGCTTTCGAGTTCGACAGAATTCGTTGATACATTTTAGATAATAT		1520
450	I Y P Y H *		454
1521	GACAAAAGAGATATTTATTATATTATTCATGCGTCGTTTGTAAATTTCTGTGCATTGTGTACATAAATATAATATAGATGC		1600
		← PcTbx2 R4	
1601	GGTATATAAAGGATTTTATGTAATATGAAGAAGCTATTTCAGACTTA <sub>(20)</sub>		1647

*Podocoryne carnea* *Tbx1* (fragment)

The intron-*exon* junction are bold typed, the intron is with small letters.

The arrows represent the specific primers.



*Pleurobrachia pileus* Brachyury

The arrows represent the specific primers, the T-box domain is underlined.

	PpBra F4 →	
1	CGGCGCAGTGTGTGGATTTCGCCTTTTCAACGAGAGTCGCGGGGACTCTCGGATCGAGCACTCTAACCCCTCGTGGTGT	80
81	TCCCACTGATCACCAGTTATTATATACTTGTTCCTCACCCTTAAATACAGTTCAACAACCTCTCAATGAGTACTGC	160
1	M S T A	4
161	CAACTCGTTTTGTTCCAGTTTCTCAAGCAACCAGCAGCAGCAGCAGCAGCTTCATTCTCCGTTAACTCGCTCCTTGCCT	240
5	N S F C S Q F L K Q P A A A A A S F S V N S L L A S	31
241	CGGGAATGGAAGAGCAACCACAGTACAATACGGCTGATATGACAACAGCAGGAGATCCCACAACCTGCTCAGATCCCTG	320
32	G M E E Q P Q Y N T A D M T T A G D P T T A H S I L	57
321	TTGAGCAGAAACCTGACCAGCCACATGCTGAGCTAGATGAGAAGGAGCTGTGGGTGCGGTTCAATCGTCTTACAAATGA	400
58	L E Q K P D Q P H A E L D E K E L W V R F N R L T N E	84
401	GATGATAGTCACCAAGAATGGAAGGCGAATGTTTCCAGTATGAAGGTGAATTTGAGTGGATTAGACCCCCACGGAATGT	480
85	M I V T K N G R R M F P V L K V N L S G L D P H G M Y	111
481	ACTCCCTCTGCCTGGGGTTCACCCCATGACAATCAGGTGGAATATGTGAATGGAGAGTGGGTTCCAGGTGGTAA	560
112	S L C L G F T P I D N H R W K Y V N G E W V P G G K	137
	PpBra F3 →	PpBra F1
561	CCGGAGCCCCAGCTCCCGTTGTGTCTACGTCCACCCCGATTCCCCGAACTTTGGAGCACACTGGATGAAGCAAACAGT	640
138	P E P P A P G C V Y V H P D S P N F G A H W M K O T V	164
	PpBra F2 →	
641	GAGCTTCTCAAAGTGAAACTCACCAATAAACTTAACGGCGGGACAGATAATGCTTAACTCGTTGCACAAGTACGAGC	720
165	S F S K V K L T N K L N G G G O I M L N S L H K Y E P	191
	PpBra R2 ←	PpBra R1 ←
721	CACGCCCTCACATTATTCGAGTGGGTGCCCTGATGTCTCAAAGAGTGTCTTCCCTACTCTTTCGCGGAGAGTGCCTTC	800
192	R L H I I R V G A P D V S K S V F T Y S F A E S R F	217
801	ATAGCCGTGACCGCCTACCAGAATGAAGAGATTACCGGGCTGAAGATAAATAATCCGTTTCGCTAAGCGTTCCTAGA	880
218	I A V T A Y Q N E E I T G L K I K Y N P F A K A F L D	244
881	TGCAAAGGAAAGCAAGAGAAAGATATCCATGGACAGAAGCGGGAGAGAGAGGATTGTGGCTCGACTAGTAAATCCCTTT	960
245	A K E R Q E K D I H G Q K R E R E D C G S T S K S L S	271
961	CATCGTGTAGTTTCTATAATCGAGTGGCAGCTGGAGGACACCATCCAGATCATGAAAGGAGGCCAGCTAAACGGGTGTGC	1040
272	S C S F Y N R V A A G G H H P D H E R R P A K R V C	297
1041	CTGAACACCCCTAACTATGCTTTCCTCGCCATCCTGCCTTGACCTCGCGGAGAGTAAGCAACAGCAGGAGGGTT	1120
298	L N T P N Y A F P R H P A L D L A G R V S N T T G G L	324
1121	ACACGGCTTCAAGGTGAAACAGTGAACCAACAGCAGCAGTAGCTCCACCTACTCGGATGTATGGACATCACCTTCA	1200
325	H G F Q G E T V Q P T A A V A P T Y S D V W T S P F T	351
1201	CACCACATACTGAATCCTACTGGACTTCATCCTACAATACATACAGTGCAGCACTGCAACCTTATCTTCCTGTTaCCTCA	1280
352	P H T E S Y W T S S Y N T Y S A A L Q P Y L P V T S	377
1281	ACCTATACCTCCCCATCCTCCTTGTTCATCCGCGGGTCAAGGAGAACCAACAGAAGTCCGACGATTGGCACTACTGTT	1360
378	T Y T S P S S L V H P R V K E E P T E V R R L A L L L	404
1361	GCTATGGATTTATTCTCAGCTGTGGGGACGGCACCAGCAGATCCAGTTTCTATCTGCCTTGGAGCGCAAACCTCGA	1440
405	L W I Y S Q L L G T A P S T I P V P I C L G A Q T S T	431
1441	CAGCAGACGGTCAGTGGTGAACCTAGCCACAGTTCCCTGACAGAATGGAACAACAATGTACGATATATGAGAGTGGC	1520
432	A D G Q W W N L A H S S P D R M E Q Q C T I Y E S G	457
1521	ATGCGATGTCAGATAACAGATTATGGACATTAACCAGCTTGTGATATATACAGGATATATTGTTATTTCAGCACTTTATT	1600
458	M R C Q I T D Y G H *	468
1601	TCCTACGTCGGAGATAGCTGTTCTCGATTGGATAGGCTCGAATACATAGAAGCCAGTGTATTGAGTATGAACAATTTTG	1680
1681	AATATCAAGCCTCATTTTTATACGATCCGTTTGGATTGCAAATCCTTAGTTTA <sub>(20)</sub>	1731
	← PpBra R3	

*Pleurobrachia pileus* Elongation factor-1-alpha

The arrows represent the specific primers.

	PpEfF3 →	
1	ATTTTAGTAGCCAGAAGCCATGCCTCCCAAGCGAAATAAGGAGAAGCCCCATCTGAACATTGTTGTCATTGGTCATGTGC	80
1	M P P K R N K E K P H L N I V V I G H V D	21
81	ACTCTGGTAAATCTACCACCACCGGACATTGATCTTTAAGTGCCTGGTGTGCGATGACCGAACCATTGAGAAGTTCAAG	160
22	S G K S T T T G H L I F K C G G V D D R T I E K F K	47
161	AAGGAGCCGAGGAGATGGGCAAGGGATCCTTCTGTTATGCTTGGGTTCTGGACAAGCTGAAGTCCGAGCGTGAGAGGGG	240
48	K E A E E M G K G S F C Y A W V L D K L K S E R E R G	74
241	AATCACAATCGATATTGCTCTTATGCGAGTTCGAGACCCCCAAGTTCGATGTCACCATTATTGATGCCCCCGGACACAGGG	320
75	I T I D I A L M Q F E T P K F D V T I I D A P G H R D	101
	PpEfF2 →	
321	ATTTTCATCAAGAACATGATCACTGGAACCTCTCAGGCTGATGCTGTGCTCATTGTGCTGCCGGAACCGGAGAATTC	400
102	F I K N M I T G T S Q A D A A V L I V A A G T G E F	127
401	GAGGCTGGAATTTCTCCAACGGACAGACCAGGGAACATCTCTCTGGCTTTTACCCTCGGAGTTCGAGAGGTTATCAT	480
128	E A G I S S N G Q T R E H L L L A F T L G V R E V I I	154
481	TGCCATCAACAAGTGGACACCACATCTCCCCCTTATAACGAGGCACGATACAATGAGATCAAGAAGGAGGTTGGTGCCT	560
155	A I N K M D T T S P P Y N E A R Y N E I K K E V G A Y	181
561	ACCTGAAGAAGGTTGGTTTCCAGGTTCCCGAGTGCCTTCATCCCATCTCTGGGTTCAAGGGAGATAACATGGATACC	640
182	L K K V G F Q V P R V P F I P I S G F K G D N M D T	207
641	CTGACTACTAACATGAGCTGGTGGAAAGGGAGCCAAAGGTGCTGACGGGCGAGAAAGGAAAGGAAAAGGAAGCCACCGTGA	720
208	L T T N M S W W K G A K V L T G E K G K E K E A T V H	234
721	CTTCTGACTGAGGCCCTCGACAATGTCATCCCCCACCCGACCCTTCAACAAGCCCTCAGGCTGCCCTTCAGGATG	800
235	F L T E A L D N V T P P T R P F T K P L R L P L Q D V	261
801	TCTACAAGATTGGTGGCATTGGAACGGTTCCCGTGGGCGAGTTCGAGACCGGAATCATCAAGCTGGCATGGTTGTTACC	880
262	Y K I G G I G T V P V G R V E T G I I K P G M V V T	287
	PpEfF1 →	
881	TTCGGACCGTCCGACTCACCCTGAGGTTAAGTCTGTTGAGATGCATCAAGAATCCCTCCCTGAGGCCGTCCCGGAGA	960
288	F G P V G L T T E V K S V E M H H E S L P E A A P G D	314
961	CAACGTTGTTTCAACATCAAGAAGCTCGCCGTCAGGAGATCAAGCGAGGATTGTTGCTTCTGACAGCAAGAACGACC	1040
315	N V G F N I K N V A V K E I K R G F V A S D S K N D P	341
	PpEfR1 ←	
1041	CAGCCAAGGAGGCCGTCGAATTCGAGGCCAGGTTATTATCATGAACCACCCTGGCTCCATCAGTAACGGATATACCCCC	1120
342	A K E A V E F E A Q V I I M N H P G S I S N G Y T P	367
	PpEfR2 ←	
1121	GTCCTCGATTGCCACACATCTCACATCGCCTGTAAGTTCACAGAGATTAATCGAAGATCGACAAGCGATCAGGAAAGGT	1200
368	V L D C H T S H I A C K F T E I K S K I D K R S G K V	394
1201	CATGGAGGAGAATCCTAAGCCATCAAGAACGGTGTGATGCTTCCATGTGTATCTTGAAGCCACCAAGCCCATGGTTGGG	1280
395	M E E N P K A I K N G D A S M C I L K P T K P M V V E	421
1281	AAACCTTCAAGGAGTATGCTCCTCTTGACGATTCGCTGTCCGAGACATGAAGCAGACCGTCCGCTGTCGGCGTTCGTAAG	1360
422	T F K E Y A P L G R F A V R D M K Q T V A V G V V K	447
1361	ACCGTCACCAAGGAGAAATCATCAGGCAAGACCACCAAGTCCGCCGTCAAGGCCGAGGCAAGAAATAAGCAGTCTGGAC	1440
448	T V T K G E S S G K T T K S A V K A G G K K *	470
1441	TACCGCCCGTGGCATTTTTAGAGCTCTGGTTATTTAACTTGAATTGATGAGTTTTATTTTTTACAATGACA <sub>(20)</sub> 1512	
	← PpEfR3	

*Pleurobrachia pileus* Tlx-like

The arrows represent the specific primers, the homoedomain is underlined.

	PpTlx F1 →	
1	CTACTTTAAACAAATTAATTTGAGCTAAACGATACACCTCATGGCAAAACGTCTCAAGTT	60
1	M A K R L K L	7
61	AAGCTTCTCCATCGACCAAATCCTCGGTTTGTCCCTGGACAAGAGTTCTGGAGAGGCGAG	120
8	S F S I D Q I L G L S L D K S S G E A R	27
121	GAATCACCATGGTGACGCGGCTAGCCAGGTGATGGCAATGATCAGCAGCGCCGGAGCGGG	180
28	N H H G D A A S Q V M A M I S S A G A G	47
181	ATCTCTGGGTGGTACTGAGGAACATACCGCAGAGTACAGGCTACACTCTAGCCCTGGATC	240
48	S L G G T E E H T A E Y R L H S S P G S	67
241	TGATGGACTCAATGACTACCACAGCGATGAAGAACTCAAGAACAGGAAACGAAAACGAAC	300
68	D G L N D Y H S D E E L K N R K R K R T	87
301	GAGAACAACCTTCTCGTCAGCACAGGTGTACGAGCTGGAGAAGAAGTTTCAGAGGTCACA	360
88	R T T F S S A Q V Y E L E K K F Q R S Q	107
	PpBar F1 →      PpBar F2 →	
361	GTATCTGAGTGCAGGTTGACAGGTTGAACCTTGCAGCTGCCCTCTCCATGAGTGATGTCCA	420
108	<u>Y L S A V D R L N L A A A L S M S D V Q</u>	127
	← PpBar R2      ← PpBar R1	
421	GGTAAAAAGATGGTTTCAAACCGACGATGCAAAGAGAGACACCGAGCGGAGGGAGAAGA	480
128	<u>V K R W F Q N R R C K E R H R A E G E D</u>	147
481	CACATGGTATCCGATCGCCACTCATGCTATACCAGCAATATCTGGGAACGACTGCCGACA	540
148	<u>T W Y P I A T H A I P A I S G N D C R H</u>	167
541	CGATCCTCCAAACACATCTCCGAAATCTCGTCCAGCTACAACTTTCCAGCGCGGCATGGC	600
168	D P P N T S P K S R P A T T F Q R G M A	187
601	CAAGCACCTTGGCATCACCCACTCCCACCGTCCCATCCCAACTCTCTTCTTCGAGTTGAG	660
188	K H L G I T H S H R P I P T L F F E L R	207
661	ATAAAGATAGACCAACCTCTTACGATGATAGAAGTAGAGACAGTCTGGCGAGAGAGAGG	720
	* ← PpTlx R1	
721	GAAGAGCGTGAGACACTGAAGTGTGCATGTGTTATTATATGTAGCGACTGGATAACAAGT	780
781	ACTATGTGAAAGAGAAAAGTGAGAGAGAGATGTAGCATGACGACACGTGCAGTTACCGT	840
841	GCTCGGACGAGTGTTCTACATGTTCTGTGAGAGGCGACCACCAGGGCGTGTGAAAAAA	900
901	GGCCCCCTCGCTCTCCCCCTCAGATGTGGGTCTAACCGCACCTCCAGCCACATCGGAGCC	960
961	ATGGTTGTTGGGAGGATTCTAATTAAGTATAATATGTATTGGTTTCCAAAAGACAATTT	1020
1021	ACACGTGATTGTGATGATTACACGTGCCTGCTTGTGCTTCTGTTACGTGTGTTATTGAG	1080
1081	TCAATGTTTTTTTCTATTCTCCTTAATAAGTATGGATTAAGATATCATGAGCCTCGCTAA	1140
1141	TTGCCTGATCTCTAAACTTTTTTTGATGTTAGGAGACTCTAATGAGAACTTTTAAGCTCA	1200
1201	CATTTGTTCTCTCCTAAGAAGCTACTATATATTCTTTATCTTTTTCGTGTGCTTTTTTGA	1260
1261	CCAGTTTACGTTTTTTTTCTGATTTTTTTGTCAAACAACCCACTATATGACACAAGTGT	1320
1321	CTTGACAGATTAATGACTTTAATTAATGATCTGA	1355

*Pleurobrachia pileus* Tbx2/3

The arrows represent the specific primers, the T-box domain is underlined

← PpTbx F5 →  
 1 TCTCCACACCAAAAGTTCAACAATATGGCTGAAAGTCGTCTCATGTCCTCCCGAGCCCGTGGTTTGGCCGTTAACTCCCTCGTAGGAGCTGATCACAGC 100  
 1 M A E S R L M S S R A R G F A V N S L V G A D H S 25  
 101 CCTGGCGGAGGCGGTGGAGGAATAGCTTCATATGACGAACCTTTGATAAAGAACGAATACAAACCCTGACTATCAGACAATCAAACTGAGGCTTATGGAT 200  
 26 P G G G G G G I A S Y D E P L I K N E Y N P D Y Q T I K T E A Y G S 59  
 201 CACCAGGCACACAGTTCACAGACGCTGGCTCGGACCGAGACCAGGCCCCACCCACCCTGAACGAAGAAAGTGGGTGATCTTAGAGCTGAAAAA 300  
 60 P G T Q F T D A G S D R D Q A P P P P P L N E E S E V I L E L E K 92  
 301 GAAGGACCTCTGGCAGGCTTCTGTAACATGGCACCAGAAATGGTTCATACAAAAGCTGGAAGGCAGCTATTCCAGCCCTAAAATGCAAAAGTATCGGGA 400  
 93 K D L W O A F C K H G T E M V I T K A G R R L F P A L K C K V S G 125  
← PpTbx R3 →  
 401 CTAGATCCAAATGCAAAATACGCTTTTCTAGTGGACATTGTTCCAGCCGATGACTGCAGGTACAAGTTCAGTAATTGTGAGTGGGTTGTAGCCGGCAAA 500  
 126 L D P N A K Y A F L V D I V P A D D C R Y K F S N C E W V V A G K A 159  
 501 CGGACCCGGAACCCCGAAGCGAATGTACGTTCCACCCGAAATCACAAACACTGGCGCCACTGGATGAAGAAAATAGTCAAGCTTCCATAAACTAAAAAT 600  
 160 D P E P P K R M Y V H P E S P N T G A H W M K K I V S P H K L K M 192  
← PpTbx F1 →  
 601 GACGAACAACGCTCTCGGACACCGGTGGTTATGCGATCCTAAATTCATGCATCGGTACCGCCCGGTGTACACATCGTCCAGTGTGATGATGATCTACAA 700  
 T N N V S D T G G Y A I L N S M H R Y Q P R V H I V O C D D V Y K 225  
← PpTbx F2 →
← PpTbx R2 →
← PpTbx R1 →  
 701 ATCCCGTGGTGTGCTTTCCGAACCATGATCTTCCAGAAACAGAGTTTTATGCGGTTACAGCCTACCAGAGTGAAAAATAACCCAAATTAAGATCGAAC 800  
 226 I P W C A F R T M I F P E T E F Y A V T A Y O S E K I T O L K I E H 259  
 801 ACAACCCCTTCGCGAAAGGTTTCCGAGAGCCCGGACTGCCACCCAGCAGCAGAGGAGCAGCAGAGTCTAAGATGTGAATCATGTGATAATAGCTCT 900  
 260 N P F A K G F R E P G T A T Q A A E E R H E S K D V N H V I I A L 292  
 901 GGAACACTACGGTGCACCCCTTCATTGACGCTGGGGATCTATCCACATACCCAGGCACCCCGGAAAAATCAAGCTGCAAAATCAACAGAGTCAG 1000  
 293 E H Y G A P P F I A A G G S I Q H T Q A P A G K S S C K Y Q Q S Q 325  
 1001 TTTGACATGAATCTCCAGTCCAGAACGGTGTTCCTCCATCACATCGCTCCCTACCAGAGGACCTCCCTACCCACAGGCGCCAGTTCGCGAGTACCC 1100  
 326 F D M N S P V Q N G V P L H I A P Y P E D L P T P Q A T S S A V P L 359  
← PpTbx F3 →  
 1101 TTCCCATGTTATTGGGATCAAGTACCAGCCCGCAACTCAGCAAGGATGCTATGACATTGACACCTACCCCTTACTCGGCAACCGCTGGAGCGATGG 1200  
 360 P M L L G S S T S P S Q L S K A C Y D I D T Y P Y S A T A W S D G 392  
 1201 GCTAGCTAACGCTCAGATACCATACAGCGCTGCTCTGGCAGATTTTCTCCCGCTGCTCGAGAGAGTGTGGATATCGTGGCTCGCCTCTGAACGCTGCC 1300  
 393 L A N A Q I P Y S A A L A D F P P A A R E S V G Y R G S P L N A A 425  
 1301 ACGGCATCGATTACCCGGGATTGTAGGGTCACCACTTCTATAACTCCAGTTACATACCTCTGATGTATGGTCATGCTCTGGGGTACTGGTGGGT 1400  
 426 T A I D Y P G D L R V T N F Y N S S Y I P L M Y G H A P G D T G A W 459  
 1401 GGAGTTATCCAGGTATCCAGGCTCTTGAACGGGGATCTGCTACCCCTCCCGGCTCAGAATCAACCAAACTCTGATACCCCTCGCAGACCACTCCAA 1500  
 460 S Y P G I Q A L E R G S A T P P A V S E S T K P L I P L A D H S N 492  
 1501 TCCTGGTAGTCTGCTCCCGTAGTGAATACTCCACCCCACTCCTGTTGTGTCTACCCGAACTCTGGCCTAATCGGAGCTATTCTCTGTTCTGATAGT 1600  
 493 P G S P A P V V N T P T P P P V V S T P N S G L I G A I P A S D S 525  
← PpTbx F4 →  
 1601 GACTTTAGTGCCGTGAATAACTGTCTCTACAGCAAGGAACACCCCGATCTCGGTTCTCAAACCGTGGTGATCACCCCTATTGGGCCTGGCAACCGCCA 1700  
 526 D F S A V N N C L Y S K E H P R S R F S N R G D H P Y W G L A T A N 559  
 1701 ACCATTAACAGTGGCATGTTACATGTGTAGTAGCATCAGTTACCATGGTTACAATCAGTTACATTGGTGACCACCATGCTAGACACACTTCAT 1800  
 560 H \* ← PpTbx R4 → 560  
 1801 AAAGGGCAATAACATACTTTAGAACTCATTAGTCGATTAAATATTACAATTAATGCACCGTTTTGAAAGCAATATTTTCATTACTACACCCGCGAT 1900  
 1901 ATTTTCCACAAAAGCTGACTAAATGAGTGTGATGTGATTTTGTCCAATGAATGATTTTGTAGAAGCAATATTACTGCAGTGATAGAACGTTAGAACTACT 2000  
 2001 TTTTGTGACGCCACTTTTTGTACCCACTATAATTATCCCAATTATTATTAATAGAGTCCACTATTTGAACAACAGCCAAATCCCAACAGATGATTAGT 2100  
 2101 TCAATTACTACTACAGTATACATTGACACTATTTTACAGATGAGATATTCATACGGAACCTGGGATGAAATAATATTTGACTAATACTCTGTATTATATT 2200  
 2201 GTGAATTAATATACTCAGGTACACAATACACATGATGCTTGTGATGTTTTTCAAACCACTAGCAGTAGCAATTGATGTTTGTGATGATAATATATAA 2300  
 2301 ATAGTATGATGATCATGATGATGAAGCGATGGATTAATAACGTAATAATAGTGTGCATATTTTAATTGAGTGTGATGCTGATGATGATGAGACAAGT 2400  
 2401 TTGAGAAGTATTAACCCGATTACCTGGTTCTAATTTATTTTCTACCATCATAGAAGCGGTTTTTACACATTAATTTACCCGATCTAATAAAGTGT 2500  
 2501 AGCTTTCTATAATTTTATCAGCA<sub>(2,0)</sub> 2523

*Trichoplax adhaerens* Actin

The arrows represent the specific primers

1	GCTCCACTTTTCGTCGACGTTGCAACTCGAAGGTACCGTGCGAAACTTCCACTTGACCGTTACAAGGATTTTTTCCTCTT	80
81	GGCATTCTGATTTTTTTTTTCTCTCCATAAAACCTAATTTCAATTCACCATGTCAGAAGAAGATGTTCAAGCTTTGGT	160
1		10
	M S E E D V Q A L V	
	TAct F1 →	
161	AATCGACAATGGCTCCGGTATGTGTAAAGCCGGATTTGCTGGCGATGATGCTCCTCGTGCTGTTTTCCCTTCAATCGTCG	240
11	I D N G S G M C K A G F A G D D A P R A V F P S I V G	37
241	GCCGACCTCGCCATCAAGGTATCATGGTTGGAATGGGTCAAAAAGACTCGTATGTTGGTGACGAAGCACAAAGTAAGAGA	320
38	R P R H Q G I M V G M G Q K D S Y V G D E A Q S K R	63
	TAct R2 ←	
321	GGTATCCTTACCTTGAATATCCTATCGAGCACGGAATTGTCACTAACTGGGATGATATGGAAAAGATCTGGCATCACAC	400
64	G I L T L K Y P I E H G I V T N W D D M E K I W H H T	90
401	TTTTTACAACGAATTGCGTGTAGCTCCAGAAGAACATCCCGTCTTACTAACTGAAGCTCCCATGAATCCCAAAGCTAACC	480
91	F Y N E L R V A P E E H P V L L T E A P M N P K A N R	117
481	GTGAAAAATGACTCAAATCATGTTTCGAAACATTCAACACCCAGCTATGTATGTCAACATCCAAGCTGTACTATCCTTG	560
118	E K M T Q I M F E T F N T P A M Y V N I Q A V L S L	143
561	TATGCTTCTGGTCGTACTACTGGTATCGTTCTCGATTCTGGCGATGGTGTAAAGCCATACCGTTCCAATCTACGAAGGTTA	640
144	Y A S G R T T G I V L D S G D G V S H T V P I Y E G Y	170
641	CGTTTGCCTCATGCTATGATGCGTGTGGACTTGGCTGGCCGTGACTTAACTGACTACATGATGAAAACTTGACTGAAC	720
171	A L P H A M M R L D L A G R D L T D Y M M K I L T E R	197
721	GTGGCTATACTTTCACTACTACCGCTGAACGTGAAATCGTTCGTGATATCAAGGAAAAATTAGCCTATGTTGCTTTAGAC	800
198	G Y T F T T T A E R E I V R D I K E K L A Y V A L D	223
801	TTTGAGCAAGAAATGCAAACCGCTGCTTCTAGTTCAAGCCTTGAAAAGAGCTATGAATTACCTGACGGTCAAGTCATCAC	880
224	F E Q E M Q T A A S S S S L E K S Y E L P D G Q V I T	250
881	CATTGGTAATGAACGATTCCGCTGCCAGAAAGCTCTCTTCCAACCCGCTTTCTTAGGTATGGAATCCGCTGGTATCCAG	960
251	I G N E R F R C P E A L F Q P A F L G M E S A G I H E	277
961	AAACTACCTACAACCTCTGTAATGAAATGCGACGTGATATTCGAAAAGACTTATACAGTAATATTGTCTTGTCTGGCGGT	1040
278	T T Y N S V M K C D V D I R K D L Y S N I V L S G G	303
1041	ACTACTATGTTCCCTGGTATTGCTGACCGTATGCAAAAAGGAAATGTTAGCTTGGCACCTCCAACATGAAAATTAAGAT	1120
304	T T M F P G I A D R M Q K E I V S L A P P T M K I K I	330
	TAct R2 ←	
1121	CGTTGCTCCACCAGAAAGGAAATACTCTGTCTGGATCGGAGGATCTATTTGGCTAGTTTATCCACCTTCCAACAGATGT	1200
331	V A P P E R K Y S V W I G G S I L A S L S T F Q Q M W	357
1201	GGATCTCCAACAAGAATAACGATGAATCTGGCCCCGTATCGTTCACCGTAAATGCTTCTAAGTTCGTGTTAAGTTTAAT	1280
358	I S K Q E Y D E S G P A I V H R K C F *	376
1281	TTATACAATATCCAAAATCGAATAACTATAGCAATATCAAAAATCTGATAATAAACTGATTTAAATCCAGA <sub>(20)</sub>	1352



*Trichoplax adhaerens* Brachyury

The arrows represent the specific primers, the T-box domain is underlined

	→ TBra F5		
1	ACCACACAGCGCTTGATAAGTCTTAACCAAGTCGGTGGCTTGTGAATGTCAAATCAAATAATGAAATCAA	100	
1	<u>M S N Q I M K S T S S F S V T N L L S</u>	19	
101	GCTAGCGAAAATGATGTCAGTAATCAATCGTCTGATGATGAAAAACAACCTAATTTGATGGCGAATCATAGTCACCAATGGAGCAACAACCAATGGCGA	200	
20	A S E N D V S N Q S S D D E K Q P N L M A N H S H T N G A T T M A T	53	
201	CAGCTGCTGCAGTGTGTCAGTGTGCTGAGCTGCAAAATGCGAATGCGTGGTAGTAATCATCGTCATCATGTTAACGCTACAAACAAGCCATAAAAAGGG	300	
54	A A A V A A V A A A A N A N G G S N H R H H V N A T N T S H K K G	86	
	← TBra R5		
301	TGCTGCTACTTCCACCACACCAGCATCCAATAATCGTAATCAAATAATCGATTCCGATGCTAATGTTAAAGTCAATTTGTTAGATATCGATTTATGGAAA	400	
87	A A T S T T P A S N N R N Q N I D S D A N V K V N L L D I D L W K	119	
	→ TBra F7		
401	CGATTTGAAAATTAACCAACGAAATGATCGTGACCAATCGGGCAGACGAATGTTTCTATACTCAAAGTTCAAATAAGTGGTCTCGATCCCCATGCAA	500	
120	R F R K L T N E M I V T K S G R R M F P I L K V O I S G L D P H A M	153	
	← TBra R3		
501	TGTATTCACTTGTAGACTTTGAGCCGCTGATAATCATCGTTGGAATATGTTAATGGTGAATGGATTCTCGTGGTAAAGCCGAACTGCACCACC	600	
154	Y S F L L D F V A A D N H R W K Y V N G E W I P G G K P E P A P P	186	
	→ TBra F1		
601	AAGTTTCGTTTACATCCATCCAGATTCACTAATTTGTCATCGCATTGGATGAAACAACAGTTCATTTTCGAAAGTAAAATTAACCAATAAATCAAA	700	
187	S F V Y I H P A S N N R N F A S H W M K O P V S F S K V K L T N K S N	219	
	← TBra R1		
701	GGCGCGCCAGATTATGCTAAACTCATTGCATAAATATGAGCCTCGGCTGCACATTGTTAAAGTGGGTCACTCCAGATTCCGATAAAAACAATTACTG	800	
220	G G G O I M L N S L H K Y E P R L H I V K V G S S P D S D K T I T A	253	
	← TBra R2		
	→ TBra F4		
801	CTCATTCTTCCCTGAACTCAATTTATGCGGTGACAGCCTACCAAAAACGAAGAGATTACCGCCTTAAAAATCAAGTATAATCCTTTCGCTAAAGCGTT	900	
254	H S F P E T O F I A V T A Y O N E E I T A L K I K Y N P P A K A F	286	
	→ TBra F3		
901	CTGGATGCCAAGAAAGAGCTGAGCAAAAAGAGCTGAAGAAAAGAGCCATGAATTGAGAGCACCTGCATATCCACCTCACTTAAATGGCGTTGGTAC	1000	
287	L D A K E R A E Q K E A E K S H E L R A P A Y P P H F N G G W Y	319	
	← TBra R6		
1001	TTACCAAATCTAATCCAATGTGTCGACACCTCGATCCATTCAACCTCCGTATTGCAAAACAGGTGTTCCATCCACATCATCGCATGCCGCTAGAGC	1100	
320	L P N S N P M C R T P R S I Q P P S L Q T G V H P H H R M P A V E R	353	
1101	GATATTGCGCATATCTCGCCCTGCAGGCTATCCTCAACAATATCATCAGGCATCAACCTATGCGAGCGGCCACCTTACTGATCATGACGCCAATGCGAT	1200	
354	Y S P Y S R P A G Y P Q Q Y H Q A S P Y A G G H L T D H D A N A M	386	
1201	GTTATTACAAAGTGCAGCAGCTGCTAATGTTGGCACCCCGACAGATCCCTGGGGCGATCTCGTCTTAATCCACCATGTTGCGCTGCGCAACTGCCTTG	1300	
387	L L Q S A A A A N V G T P T D P W G R S R P N P P M L P S P T A L	419	
	→ TBra F6		
1301	CCAGGGCCACTCTAGTTTATGGTTCGTCGACATCAACGCCATTAGGAGCAGCAGAGCCCAATTTAATTCACATCCATCATTGCAACGGCAGCTACGG	1400	
420	P G P P P S L W S S T S T P L G A A R A Q F N S H P S F A T A A T V	453	
1401	TGGCCGCTGCAGCAGCTTCTAGTCTAGCGCGGACATGTTTCTAGCTCCAATAGTCTTATCAATGCTAGCCATATTGCAACGATCAGGGATGTACGGTAAT	1500	
454	A A A A A S S A S A D M F L A P I V L S M L A I L H D Q G C T V M	486	
1501	GCCCATATACAAGATATCAATGACGAAAGTGGTTCGGTGAACCCGGACATCTTTATACACATCCACCGCTATGCCTGTTTACTACTATGGCATCTTGA	1600	
487	P I Y K I S M T K V V P V K P D I F I H I H R L C L F I L W H L G	519	
1601	ATTCCATGGTCCCAAGTTAATAAGTGTAAAAATTTTCACTGTAATATCAAGAAAGTAAGCCGATGAGCAAAATACTCCATGCGCATACTTTCTACA	1700	
520	I P W C H K F N K C K K F S L *	534	
	← TBra R4		
1701	TCCTATCAAACAAGTAGCAAAGTACAAACATCTACTATTTTCATTTTTCCTTAAATGACCTAAGCTGTAACATTTACTATTTTA	1784	

*Trichoplax adhaerens* Elongation factor-1-alpha

The arrows represent the specific primers.

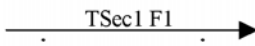
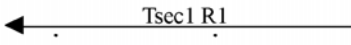
		TEF2 →	
1	GCTCTTCTCTTCGTCGTTTCAGTTGCAGCGTGATAGCTTGTACAGTAGGCTGACATCAGACAAATCTTACAATGGGTAAGGAAAAGATACATATTAACAT		100
1		M G K E K I H I N I	10
101	TGTCGTCATTGGTCATGTCGACTCTGGTAAATCTACCAGTACCGGCCACTTGATCTACAATGTGGTGGTATCGATAAGCGAGCCATTGAAAAATTGAA		200
11	V V I G H V D S G K S T S T G H L I Y K C G G I D K R A I E K F E		43
201	AAAGAAGCCCAGAATGGGTAAGGTTCAATCAAAATACGCTTGGGTCTTAGACAACTAAAAGCTGAACGCGAACGTGGTATCACGATTGATATTGCC		300
44	K E A Q E L G K G S F K Y A W V L D K L K A E R E R G I T I D I A L		77
301	TATGGAAATTTGAAACTACCAAGTATTACGTCACCGTTATTGATGCCCTGGTTCACAGAGATTTACCAAAAATATGATCACTGGTACATCTCAGGCCGA		400
78	W K F E T T K Y Y V T V I D A P G H R D F T K N M I T G T S Q A D		110
401	TTGTCCGTTACTGATCGTTGCTGCCGTTGGTGAATTCGAAGCCGGTATTTCCAAGAAGCGTCAAACCTCGTGAACATGCTTTATTAGCTTTTACTCTT		500
111	C A V L I V A A G V G E F E A G I S K N G Q T R E H A L L A F T L		143
501	GGTGTAAAGCAAATGATCGTTGGCATTATAAAATGGACAGCACTGAGCCACCTATAGTGAAGCCGTTATAATGAAATCGTAAAAGAGTGAGCACTT		600
144	G V K Q M I V G I N K M D S T E P P Y S E A R Y N E I V K E V S T Y		177
601	ATATCAAAAAGTCGGTTATAATCCCAAATCAGTAGCCCTACGTTCCCATCTCCGGTGGCATGGCGATAACATGATAGAAGAAAGTACAAACATGAAATG		700
178	I K K V G Y N P K S V A Y V P I S G W H G D N M I E E S T N M K W		210
701	GTTTAAGGGTTGGTCAAGTAAGGAAAGCAACGCCAGTGGTAAAACCTTATTTGAAGCCCTCGATGCTATCTTGCCTCCAAAAGGCCTTCAGAC		800
211	F K G W S V E R K E G N A S G K T L F E A L D A I L P P K R P S D		243
801	AAGCCATTGCGCTTACCTTTACAGGATGTATACAAGATGGAGGATTTGGTACCGTCTCGTGGACGTGTCGAACTGGCACTATAAAACCTGGTATGG		900
244	K P L R L P L Q D V Y K I G G I G T V P V G R V E T G T I K P G M V		277
901	TTGTCACATTGCCCCAGTCAACGTTACTACTGAAGTAAATCCGTTGAAATGCATCAGCAATCCTTGACTGAAGCTTTCCCGGTGATAATGTCGGCTT	← TEF1	1000
278	V T F A P V N V T T E V K S V E M H H E S L T E A F P P G D N V G F		310
1001	TAATGTCAAAAACGTCCTCTGTAAAGGATGTCCTCGTGGCAATGTAGCTGGTATTCTAAGAACGATCCTCCAAAAGAACCAAGTGTTCGATGCTCAG		1100
311	N V K N V S V K D V R R G N V A G D S K N D P P K E A K C F D A Q		343
1101	GTTATTGTTTGAATCATCCTGGTCAGATCAACCCCTGGCTATGCTCCAGTATTGGAAGTCCATGCTCATATTGCATGCAAATTCACCGCCTTAAAGG	← TEF1	1200
344	V I V L N H P G Q I N P G Y A P V L D C H T A H I A C K F T A L K E		377
1201	AGAAGATCGATCGTCAAGTGGTAAAAGTTAGAAGAAAACCCCAAATTCGTCAGAGTGGTGTGCTACCATTGCTACCCTAACTCCAAGTAAGCCTAT		1300
378	K I D R R S G K K L E E N P K F V K S G D A T I A T L T P S K P M		410
1301	GTGCGTTGAATCTTATCCGAATATCCTCCATTGGGTCGTTTCGCGCTCCGTCGATGAAACAAACTGTAGCAGTTGGTGTATCAAGAAAGTTGAAAAA		1400
411	C V E S Y S E Y P P L G R F A V R D M K Q T V A V G V I K K V E K		443
1401	GTAGAGAAGAGTGGTAAAGTCAAAAAGTCTGCGCAAAGGCTACTAAGGGCAAGAAATAATCAATCCTGATGCCCTGATATATTTAGATTAATAAAAAG		1500
444	V E K S G K V T K S A Q K A T K G K K *		462
1501	TGTAATACAATATCGTAGATGATCATGTGAATTTGTTACAATTTTATTATGCTGTACTGCTGTTGCTGGCCTAAGCGTAAATTGACTTATCTTTCACTCT	← TEF2	1600
1601	GACAGATTATCAATAAAAAGTACTGCGACA <sub>(20)</sub>		1632



*Trichoplax adhaerens* Secp1

*Secp1* codes for a small, secreted, highly charged protein.

The signal peptide and a CXC cysteine motif are underlined, the arrows represent the specific primers.

		
1	CGATCCCACGGTAACTGTAAGGACTGAAAAATGAATTTCAAATTGGTTGTCATCGCCGCT	60
1	<u>M N F K L V V I A A</u>	10
61	CTATTGTTTGTAGCCTGTGTTGCTGCCGCCGACACAGCTAAGAAAGCTGAAACCCCTGTC	120
11	<u>L L F V A C V A A</u> A D T A K K A E T P V	30
121	AAACAGGAAGCTAAAACCTCCAGAAAAGGAGGAACCAGCAGAGAAGGAAGAACAAGAAGAA	180
31	K Q E A K T P E K E E P A E K E E Q E E	50
181	GAAGAAGAAGAGGAATCTCCAGTTGAAGAACGTGAAGAAGATGACGAGGATGAAAACGCA	240
51	E E E E E S P V E E R E E D D E D E N A	70
241	CTTCAAGAAAAGAAAATGATGATGGTGAAGTAAAGGAAGAAGAATCCTCCGATAAGAGC	300
71	L Q E K E N D D G E V K E E E S S D K S	90
301	CTCTGGAGCCGCCGTGCGCGTCGTCGTCGCTTCTATATCCGTCGTCGTCGTTATATCCGT	360
91	L W S R R R R R R R F Y I R R R R Y I R	110
361	CGTCGTATCTTTCTCTACAGGCGTCGCTGCACCTGCTATTATAGACGTCGATTCTTATAC	420
111	R R I F L Y R R R <u>C T C</u> Y Y R R R F L Y	130
421	AGACGTAGAGTCTTATTAAGGCGTAGATTCTCTACAGACGTCGCCGCTTTGGCAAGAAG	480
131	R R R V L L R R R F L Y R R R R F G K K	150
		
481	TAAAAAGTTTCACGATAAGGATTTTCGAAAAATTCAACAGCATTTCAGTGAAGATGATGCCG	540
	*	
541	TTAAACTTGATTAATGAGATATTTTTTAGCTGTCTTTGAAACTACCTTGTAGTATTTCA	600
601	TTGTTAGAAATCAATTTAGGCTAAAAAGGGTAGCACTACTTAATTTTCAGTTTCGATTT	660
661	TATATATTGTAAAATCTGTAGATCCCTTTATATGCCTGCAATTTAAGCATCAATAAAGAG	720
721	ACATTAGATTTGCATTA <sub>(20)</sub> 737	

*Trichoplax adhaerens* Not

The arrows represent the specific primers, the homeodomain is underlined.

	TNot F3 →	
1	GCAGTCCTGTTTTTTGTCAGCCTAATCGCTTACCCTTACTCTTCATTTTGCAACTAGGCTGTTTATACCCCATTCTATCC	80
81	TATCTAGACTATTTTCAGCCTATTTTCGATTCATCCATTTTACTTCATAGAAGAAAATAGCCATTTGATCCTGTTTTGATC	160
161	CTGTCTAGATTATTTCCAGCCCATTTCGAGCCAGCCATTTTGGCTTCATTGGAGAAGATAGTCTATTGGCCCTCTTTTGT	240
241	TGCTTTCTAGAAGACTGTTTCCAGCCTATTTTCGAGCTCACCATTTTGGCTCCATTGAAGAAGATAGTCTACTGATCTTATT	320
321	TGACTGTGGATTTCGCCTTACTCAGAAAATTGATTGAATTATCAGAGCTATTTTTCCTTCATCTGTTTTTGGTAGTCGCTCA	400
401	AAACAACAGCÁTGC AAAATCCAGTTTATGTTTACGGTGAAATCTACCAACAGGACCACCTAGTACTGTGCAGCTCCTGT	480
1	M Q N P V Y V Y G G N L P T G P P S T V A A P V	24
481	TCCAATTTATTTGACTGCTCGGCCATGGCAAGCTGCCTATATGAATCCTACATTACCACCAGGGGCAACGCCACCTTTCA	560
25	P I Y L T A R P W Q A A Y M N P T L P P G A T P P F N	51
561	ATTTTAATACCTTAATTCCAACAACGAAAAGTGATGACAACAGCAATCCACCGATGCACAATGTTCAACATCCTATGGCT	640
52	F N T L I P T T K S D D N S N P P M H N V Q H P M A	77
641	GCTGCTGTTCAAGGTGGCTATAGGCCTGCCAATCCAGTGCTATCCGATTTTACGATCCTAATATTAATCCATTAGCCGC	720
78	A A V Q G G Y R P A N P V P I R F Y D P N I N P L A A	104
721	TCAAGCCAATCCAATGAATAAATCAGTTCTAGCGAGAAAAGAAAAGTAACTCTACAGCAATTAGCAGTGGCCATC	800
105	Q A N P M N K S R S S E K E S K S N S T A I S S G H H	131
801	ACCAGCATGTTGAAAAGTTAAGATCATCAGCATCTAATCAACAACATTGGAAATCAGGCAAGAATAAACGTGTTTCAACT	880
132	Q H V E K L R S S A S N Q Q H W K S G K <u>N K R V R T</u>	157
	TNot F1 →	
881	ATCTTTACGCCGATCAACTGACCAAGTTAGAGAAAAGAAATTTGATCAACAACAGTATATGGTGGGCGCTCAACGTTTATA	960
158	<u>I F T P D Q L T K L E K E F D Q Q O Y M V G A O R L Y</u>	184
	TNot F2 →      TNot R2 ←	
961	CTTAGCGAATACATTAATACTGAGTGAGGCACAAGTTAAGGTATGGTTTCAAATCGACGAATTAATGGCGAAAAGATC	1040
185	<u>L A N T L K L S E A Q V K V W F O N R R I K W R K D Q</u>	211
	TNot R1 ←	
1041	AAATTAAGCGTAGACATACAACAGCAAGTAATGCTAATCATCAGAATACTTCTTAACTCTGTATCTGCTTCTATTGAT	1120
212	I K R R H T T A S N A N H Q N T S L T S V S A S I D	237
	TNot R3 ←	
1121	GCTCAATCTACAAGTAGCAATGGCAGTAGTAATCCTCAAAGCAGCAGTCGTAGTAGTACCAGTAGTAATTCATAAATGTA	1200
238	A Q S T S S N G S S N P Q S S S R S S T S S N S *	261
1201	AAATATTCTACCACGGTTGATGATGGACATCGTCAATCTATCATTACCATAGAGTAGAAGCCTTGAAAAAGCTATC	1280
1281	ATGAATGATATCCTGTGCTTGAAAAATAATATCATCGACTCATACTTTCGTCCTTAAATTGCAAAACTTCTAGCGATTCA	1360
1361	AATCTACAGTTGATTGATGTTGTGCGATATTTTACATAAATCGAACATGAATATTAGCATTTTGTGCGCTAATTCATT	1440
1441	ATGTAATATATTGATATTCTTGAACACTCAATTAATTAATGAAATATCGTCCATTAAAGTAAGAATATAAATTTGATTA	1520
1521	CAACATTGAGTCACGTCGGTTCCTAATAAATCGTAGGAAATGATCATTTTTAAATTTAGATCAGTTGCATGGTATTACTAA	1600
1601	GCGAACTTAATTCGTGCGTTTTATTATTGATACTCCGATTTAATGATGTATAAGCATAAAAAACGCAATATAACTATTT	1680
	TNot R4 ←	
1681	TATGTGTAGAGCGCTTTATGCTTACAAAACGAGGCAAATAAACGCTTCAGACGCACATAGTCCACACCTCA <sub>(20)</sub>	1754

*Trichoplax adhaerens* Trox-2

With capital letters are indicated the nucleotides found by 5'RACE, with the small letters the nucleotides already present in the gene bank.

The homeodomain is underlined, the residues of the engrailed-like domain are bold and the arrows represent the specific primers.

Trox F2 →

1 ACATTCCATCATTTCATCGCCTGAGTTAACCAACATTGATAGTTAAACTTAATTACTGCTAACATCAATTTAGACTTTAAAG 80

81 TTTTAACTACTATCCATATCTGTTTAAACACTAATCATCATTTTCAATTATTGCCACAATTGATTAATGATAAATTAGT 160

161 TTATGCTAGAATTTTATAACAATATTTAACAGGTGAAACATGGCCAATACATCTTTTAAATGAAAGTTTAAATCGGTCC 240

1 M A N T S F K I E S L I G P 14

241 AGGTAATCCAACCAATGCCATCATTAACGGTTGTCCATGATGGCACCATTTTATCGCCGTCGAATTACTACTCTTGGC 320

15 G N P T N A I I N G C S M M A P F Y P P S N Y H S W P 41

← Trox R2

321 CTCATGATAGCAAcccccaattattcatcatgtcaagtctgtataccaattcaactgtcacttggccagtactagtaaag 400

42 H D S N P N Y S S C Q V C I P N S T V T W P V L V K 67

Trox F →

401 ccatcgctatagtcgatcctgcatataccatcagcaacaatatcgattggataaacatcatcaccaatcattaccagc 480

68 P S P I V D P A I Y H Q Q Q Y R L D N H H H Q S L P A 94

481 cttccaaggaatccaaaataccgacactcaacaacatccagatttagcatctgataaaaactagcatgaaatcgaccagtc 560

95 F Q G I Q N T D T Q Q H P D L A S D K T S M K S T S H 121

← Trox R3

561 attcttccagaactaagcgtatacgaacagcttataccagtatgcagttattagaattagaaaaggagttcaacagtagc 640

122 S S R T K R I R T A Y T S M O L L E L E K E F N S S 147

641 cgatatctatcgcgactgaggaggattgagattgctaataatgtaaaattatcggagaagcaagttaaaataggttcca 720

148 R Y L S R L R R I E I A N M L N L S E K Q V K I W F O 174

← Trox R

721 gaatcgtcgcgtaaagtggaaaaggacaataagtttagatggacaagatcatgatgaaaacgaaagtaaaggatctcctg 800

175 N R R V K W K K D N K L D G Q D H D E N E S K G S P D 201

801 atcaatatccactatctgatagtgagaagaattatctaactgataataatactcaggcctaattttgtgattgacacgat 880

202 Q Y P L S D S E E E L S N R \* 215

881 cataataatagttgctgattatctacttattctaatacaattataataatcatccaatattcaaatttagcaaatTTTTG 960


961 cagaacgatttcttacagccgcaatgaattacattacaataaactttcattgtaagttaagcgatgataaattaatta 1040


1041 gaatacagtaagtaccgaaataagtacaattagtatacaaatgtttctcaataatcgagtatttaaatccttaaggatg 1120

1121 atatcagccgaaaaaattttcgcatctgacaataccttcaatcgctcatgatttgtgaaataaaaaataaattatagcaa 1200

1201 aaaatagtattctattcaatccataagtcctgttggcttgatgaaaccataacttaataatcaataaaaaaa 1272

*Trichoplax adhaerens* Ribosomal protein S25

TRp25S F1 

1	CTAGTTCGGTGGTTCTGCGAGGAGATCTGTAGTTATGCCTCCAAAAAAGCCGATAAAGAG	60
1	M P P K K P I K S	9
61	TGAGGATCAAAAGGCCAAAGCTGCGTCTGGTGGTGGAAAAGCCAAAAAGAAGAAATGGTC	120
10	E D Q K A K A A S G G G K A K K K K W S	29
121	AAAAGTTAAAGCTCGCGATAAGCTGAACAACCTTGGTATTATTTCGACAAGGCTACCTACGA	180
30	K V K A R D K L N N L V L F D K A T Y D	49
181	CAAGCTTTACAAGGAAATTCCTAGTTACAAACTTATTACTCCATCTGTCTGATCAGAAAG	240
50	K L Y K E I P S Y K L I T P S V V S E R	69
241	ACTGAAAATACGTATTTTCATTGGCTAAGCGTGCCTTGCAAGGAATTACAACAGAAAGGACT	300
70	L K I R I S L A K R A L Q E L Q Q K G L	89
301	AATCAAAGTTATAGACAAAACACCACGCACAAGTCATTTACTCGTGCAACTAAGGCATC	360
90	I K V I D K H H A Q V I Y T R A T K A S	109
361	TTAGAAATTATGTTGTATGCGAAACCAATCTTGTATAATTTTAAGGAATCGATGTTGTA	420
110	*	110
	TRp25S R1 	
421	GCTaTCGCTAAAGGAACTGAATAAAAAATAGAAGTTCTTATAAAGTA <sub>(20)</sub>	466

*Trichoplax adhaerens* Metallothionin

1 TCCGATCCTTCACTGTTGCTGTTACGTCGTCGGCGTTCCGTTAAAATTTCTACCGATCA 60

61 ATTTCCATTCTTTGCGTTACAATTTCTAACTTCAATGGATCCTCCTTGTAATTGTGCTG 120  
1 M D P P C N C A E 9

121 AAAGTGGTACTGCCAATGTCCTTCTAACTGTGCCTGCACCAATTGTAAATGCGCACCTA 180  
10 T G D C Q C P S N C A C T N C K C A P N 29

181 ACTGCAGACTCTGTGGTGGGAAAACATGCAAATGCGCTGAAAATATCAATAGCTGCGTCT 240  
30 C R L C G G K T C K C A E N I N S C V C 49

241 GCACAAGTTGCATATGTAAGTCAAGTGTCTAAAGGATGCAAGAAGTAAATATTA 300  
50 T T C I C T D C K C P K G C K N \* 64

301 CACCTGAAGATGTAAATGTACAACCTTTTTCTGGATTTGATTAGTTATTCTCGATGTTGTT 360

361 GATGTTTATCTGAAAATTTCTTTAACTTTGTTGTAGATGATAGTTCATGTTGTGAACCGT 420

421 ATTAATTAATGGTCCGGTTAATTTCTAGTGGtAAAAATCAAAACTTTCAATATACCTTTA 480

481 TTGATTAGCATGCACAATGCTGTTTGCTATCTGCTTGTAGTGATGTTTGTCTATCGGTCG 540

541 ATAGTGTTCAATTACAATAATAACTTGAGTCAGtTCTCAATATATGCGCGTGTGATCAAT 600

601 CaCTTAAGTAACCAGTTAAGGATAGTAGAATGTAATAAAAACCTAAAATCTACCCGTATTT 660

661 TATGCCTTCGTGCCTATCAGCTAACTAGTATCGTACGTAGCAAAGTAACGTTATAAAGTG 720

721 AATTAGAGTTATGGcTTGTCCTTCAATCAAATGTAAATCTGATTTCTGGTTTCATTGTAT 780

781 GAGTTCAGAATAAATTTCTTGCAATAAATTAATGA<sub>(20)</sub> 813



### Acknowledgments

First of all, I would like to thank Jürg Spring. He patiently led me during the preparation of this work, giving me a better insight into molecular biology techniques and introducing me to the biocomputing field. In addition, I would like to thank him for the stimulating discussions and for helping me with corrections of this thesis. My special thanks go to Professor Volker Schmid for his support. He followed my progress with attention and gave me many precious advices. He provided me many of the organisms that I studied and helped me in extracting DNA and RNA from them. Moreover, he was helpful establishing an in situ hybridisation technology for *Trichoplax*. I would also like to thank Professor Markus Affolter for his spontaneous agreement to referee my thesis. Furthermore, I feel grateful to Professor Renata Batistoni for her advices and the continued interest in my work. I would also like to thank the Swiss National Science Foundation for their financial support of this work.

I would like to thank the members of the Schmid lab: Peter Müller, Tsuyoshi Momose, Natalie Yanze, Michael Stierwald, Susanne Reber-Müller and Katja Seipel, who all helped me during many experiments with precious advice and with whom I had fruitful discussions. Moreover, I would like to thank Brigitte Aeschbach and Ruth Streitwolf-Engel, the former helped me in the daily lab life, and the latter introduced me to the microsectioning technology. Many others I would like to acknowledge for their friendship and the nice time that we spent together: Tamas Mengesha, Robert Lichtneckert, Michael Eberhardt, Lukas Jegge, Elena Pescia, Sabina Galle, Claudia List and Thomas Ortega.

This work is dedicated to my family and my friends, without them sometimes efforts and hard work would seem to be without aim. Those who dealt with Italian families know how big they can be, and it is almost impossible to mention all the members, however, the love for my parents is unique. The warm friendship with Nicola Miglino, Lorenzo Lenci, Nicola Solcá helped me a lot and others that I thank to be like they are: Roberta, Daniela, Michele, David, Renato, Forese and Dino, my brother Giovanni, Michelle, Roberto, Aurelian, Stefanie, Nakka, Ugo, Kiatto, Paolone, Salvatore, Adriano, Daniele and many many others!

The warmest thanks I owe to Arianna who gave birth to our Matilde while I was working on this thesis. There are no words I can easily find to thank you for that. Special thanks to the parents and relatives of Arianna, who took care of her and Matilde during this time.

## Curriculum vitae

### Personal information

Name and surname: Cosimo Martinelli  
Place and date of birth: Pisa, August 2<sup>nd</sup> 1971  
Nationality: Italian

### Education

June 1991: Technical certificate in Industrial Physics.  
Valuation: 54/60

July 1999: Degree in Biology at the University of Pisa with a thesis about “*DjMCM2* expression shows proliferating cells both in intact and regenerating planarians”, supervisor Prof. Dr. Renata Batistoni.  
Valuation: 110 /110 cum laude

May 2001: Attended international course: TMR Course in Evolutionary Developmental Biology (5 weeks), Roscoff, France.

### Professional and research activities

Oct. 1997-July 1999: Preparation of diploma thesis at the laboratory of Cell Biology and Development Biology of the University of Pisa, Italy.

Jan. 2000-June 2003: Preparation of Ph.D. dissertation at the Zoological Institute of the University of Basel, Switzerland.

### Scientific publications

Martinelli C and 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*, in revision.

Martinelli C and Spring J (2003). Role of the homeobox gene *Not* in the body remodelling of the basal metazoan *Trichoplax adhaerens* (Placozoa). In preparation.

Martinelli C and Spring J (2003). T-box genes in non-bilaterians: further indications for the monophyletic origin of metazoans. In preparation.

During my studies at the University of Basel I have attended lectures of the the following lecturers: M. Affolter, T. Blickle, D. Ebert, W. J. Gehring, T. Momose, S. Reber-Müller, H. Reichert, V. Schmid, K. Siepel, J. Spring, S. C. Stearns, E. T. Stöckli, N. Yanze.