

RESEARCH ARTICLE

The nervous system of Xenacoelomorpha: a genomic perspective

Elena Perea-Atienza¹, Brenda Gavilán¹, Marta Chiodin¹, Josep F. Abril^{1,2}, Katharina J. Hoff³, Albert J. Poustka⁴ and Pedro Martínez^{1,5,*}

ABSTRACT

Xenacoelomorpha is, most probably, a monophyletic group that includes three clades: Acoela, Nemertodermatida and Xenoturbellida. The group still has contentious phylogenetic affinities; though most authors place it as the sister group of the remaining bilaterians, some would include it as a fourth phylum within the Deuterostomia. Over the past few years, our group, along with others, has undertaken a systematic study of the microscopic anatomy of these worms; our main aim is to understand the structure and development of the nervous system. This research plan has been aided by the use of molecular/developmental tools, the most important of which has been the sequencing of the complete genomes and transcriptomes of different members of the three clades. The data obtained has been used to analyse the evolutionary history of gene families and to study their expression patterns during development, in both space and time. A major focus of our research is the origin of ‘cephalized’ (centralized) nervous systems. How complex brains are assembled from simpler neuronal arrays has been a matter of intense debate for at least 100 years. We are now tackling this issue using Xenacoelomorpha models. These represent an ideal system for this work because the members of the three clades have nervous systems with different degrees of cephalization; from the relatively simple sub-epithelial net of *Xenoturbella* to the compact brain of acoels. How this process of ‘progressive’ cephalization is reflected in the genomes or transcriptomes of these three groups of animals is the subject of this paper.

KEY WORDS: Acoelomorpha, Acoela, *Symsagittifera*, *Xenoturbella*, Neural system, Evolution, bHLH, GPCR, cephalization

INTRODUCTION**Phylogenetic affinities of the members of Xenacoelomorpha**

Acoel flatworms together with the nemertodermatids constitute the taxon Acoelomorpha: a group of mostly marine worms with a simple body plan. With the xenoturbellids they constitute the phylum Xenacoelomorpha. Acoelomorphs are the best-characterized clade within this phylum and for that reason will be discussed in detail in the following paragraphs. The history of the classification of acoelomorphs has been complex and their assumed affinities within the metazoans has changed. Traditionally (in precladistic

analysis) the Acoelomorpha were included within the phylum Platyhelminthes, which was considered the most basal bilaterian taxon (Hyman, 1951). The introduction of molecular characters (18S rDNA) to build phylogenetic trees showed, for the first time, that Platyhelminthes were polyphyletic, with the Acoela and the Catenulida branching as the first two clades of bilaterians and the Rhabdithophora nesting within the protostomes (Carranza et al., 1997). A deeper analysis using scores of acoels, some slow-evolving, recovered the three major bilaterian superphyla currently agreed upon (Deuterostomia, Ecdysozoa and Lophotrochozoa) and showed that the Acoela were the sister group of the remaining Bilateria (Ruiz-Trillo et al., 1999). The position of acoels seemed to be clearly established, turning them into good proxies for an ancestral bilateral animal (Baguña and Riutort, 2004). This position was confirmed in some exhaustive phylogenomic analyses (Hejnol et al., 2009). However, using alternative methodologies and datasets, Philippe and collaborators proposed (Philippe et al., 2007; Philippe et al., 2011) that Acoela could instead be a deuterostomian group, with affinities to the Ambulacraria (Echinoderms and Hemichordates). Those authors suggested that Acoelomorpha sit within the Ambulacraria, grouped with another taxon, the Xenoturbellida, and thus erected a new monophyletic group, the Xenacoelomorpha, a classification also supported by some morphological data (Tyler and Rieger, 1975; Tyler and Rieger, 1977; Ehlers, 1985; Smith et al., 1986; Lundin and Hendelberg, 1996; Lundin, 1997; Lundin, 2000; Lundin and Sterrer, 2001). Now, these alternative views, one placing the Acoelomorpha as basal bilaterians and the other placing them within the Deuterostomia, continue to provoke debate. However, and since Xenacoelomorpha seems clearly to be a monophyletic taxon, the inference of specific trends within the group remains immune to the specific problems posed by its metazoan affinities.

The acoel nervous system in context

For many years, the structure of the nervous system has been considered to be important for ascertaining phylogenetic information (Haszprunar, 1996). The Acoelomorph nervous system is a pertinent case here because it shows a high degree of variability within the group. It is characterized, according to some authors, by a low degree of centralization in the anterior domain (Raikova et al., 2004a; Hejnol and Martindale, 2008; Raikova, 2008). In the past 10 years, several studies using electron microscopy and immunocytochemistry have added to our knowledge of acoel neuroanatomy. The use of a single set of immunochemical markers, such as 5HT, GYIRF-amide and FMRF-amide has made analysis of the nervous system in this group possible (Raikova et al., 2004a; Achatz and Martínez, 2012). The systematic use of these immunochemical markers, in particular antibodies against 5-HT, led some authors to propose that acoels have a simple ‘commissural brain’ (Raikova, 2004c). However, since it is well known that these

¹Departament de Genètica, Universitat de Barcelona, Av. Diagonal, 643, 08028-Barcelona, Catalonia, Spain. ²Institut de Biomedicina de la Universitat de Barcelona (IBUB), Av. Diagonal, 643, 08028 Barcelona, Catalonia, Spain. ³Ernst Moritz Arndt University of Greifswald, Institute for Mathematics and Computer Science, Waltherr-Rathenau-Str. 47, 17487 Greifswald, Germany. ⁴Dahlem Centre for Genome Research and Medical Systems Biology, Evolutionary and Environmental Genomics Group, Fabeckstraße 60-62, 14195 Berlin, Germany. ⁵Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys, 23 08010 Barcelona, Catalonia, Spain.

*Author for correspondence (Pedro.martinez@ub.edu)

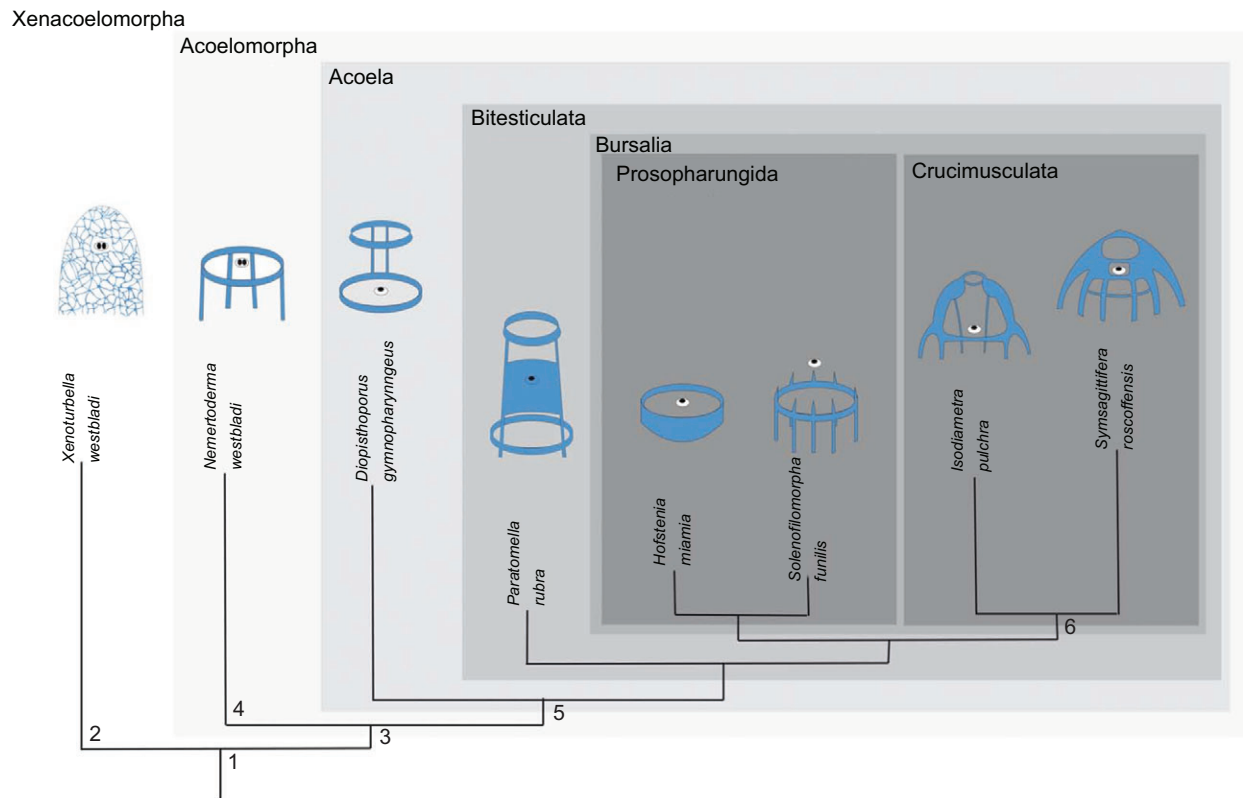


Fig. 1. Schematic organization of the nervous system in different members of the Xenacoelomorpha. (1) Putative xenacoelomorph ancestor. (2) *Xenoturbella* nerve net located basiepidermally. Presence of a statocyst. (3) Presence of a ring-commissure plus a low number of neurite bundles, extending along the anterior–posterior body axis. (4) Presence of a statocyst with two statoliths, typical of the Nemertodermatida. (5) Progressive loss of the basiepidermal location of the nervous system and tendency to possess a second, anterior, ring commissure. The statocyst has one statolith. (6) Replacement of the anterior ring-commissure by bilobed brain, with a dense neuropil, plus one to three anterior commissures. Nervous systems information from: *X. westbladi* (Raikova et al., 2000b; Israelsson, 2007); *N. westbladi* (Raikova et al., 2004b); *D. gymnoharyngeus* (Smith and Tyler, 1985); *P. rubra* (Crezée, 1978); *H. miamia* (Correa, 1960); *S. fumilis* (Crezée, 1975); *I. pulchra* (Achatz and Martínez, 2012); *S. roscoffensis* (Bery et al., 2010; Bery and Martínez, 2011; Semmler et al., 2010) (current study). Adapted from a diagram in Achatz and Martínez (Achatz and Martínez, 2012).

antibodies only label particular groups of cells, leaving the rest of the nervous system unstained (Grimmelikhuijzen et al., 2002; Cebrià, 2008), it is necessary to use complementary methodologies to characterize the detailed ultrastructure. The information provided by all these different studies has shown that the nervous system of acoels has a high degree of structural variability, although all share a general pattern: the possession of an intraepidermal plexus and three to five pairs of similar neurite bundles that are regularly distributed and spaced around the anterior–posterior body axis.

There are a few morphological features shared by acoels, Nemertodermatids and Xenoturbellids; for instance, the lack of a stomatogastric system and the presence of a pervasive peripheral plexus (Raikova et al., 1998, Raikova et al., 2000a; Raikova et al., 2001; Reuter et al., 2001a; Reuter et al., 2001b). The structural characteristics of the nervous system in the different taxa allow us to infer correlations (trends) between their neuroanatomy and their relative taxonomic positions within the group. Xenoturbellida is the most basal clade of Xenacoelomorpha. Its neuroanatomy is the simplest, consisting of a well-developed basal intraepidermal nerve net without any submuscular nervous structures (Fig. 1) (Raikova et al., 2000b). No areas of obvious condensation have been detected in *Xenoturbella*, yet they possess a statocyst (Israelsson, 2007). The Nemertodermatida nervous system comprises, in addition to a subepidermal nervous plexus, a ring of processes located at the statocyst level, which is suggestive of what could be the initial

stages of the evolutionary process leading to an anterior concentration of neurons (Fig. 2) (Raikova et al., 2004b). A quick comparison of the general neuroanatomy within the different members of Acoela – we follow the detailed description given by Achatz and Martínez (Achatz and Martínez, 2012) – reveals an increase in the complexity of neural structures, particularly in the most recent groups. In the most basal family of acoels, the Diopisthoporidae, the nervous system is organized as a ring-shaped commissure with two ganglionic lobes posterior to the statocyst, and a smaller ring located more anteriorly. These two rings are linked by ventral tracks (Fig. 1) (Tyler, 2001; Achatz and Martínez, 2012). The data obtained for *Paratomella rubra*, which is a member of the Paratomellidae, a group that branched later and is considered the sister group of Bursalia, indicate that the nervous system consists of a dense plexus around the statocyst with two neural extensions that connect, dorsolaterally, two rings and a pair of dorsolateral neurite bundles that extend from the posterior to the end of the animal (Fig. 1) (Crezée, 1978). The clade Bursalia is divided into two major groups, the Crucimusculata and the Prosopharungida. In the latter group, the nervous system is organized in one to three ring commissures plus eight neurite bundles (Fig. 1) (Crezée, 1975). The sister group of Prosopharungida is Crucimusculata (Jondelius et al., 2011), a clade that includes all the named ‘higher acoels’. This clade, which includes some of the most studied acoels (e.g. *Isodiametra pulchra* or *Symsagittifera roscoffensis*), shows a more-

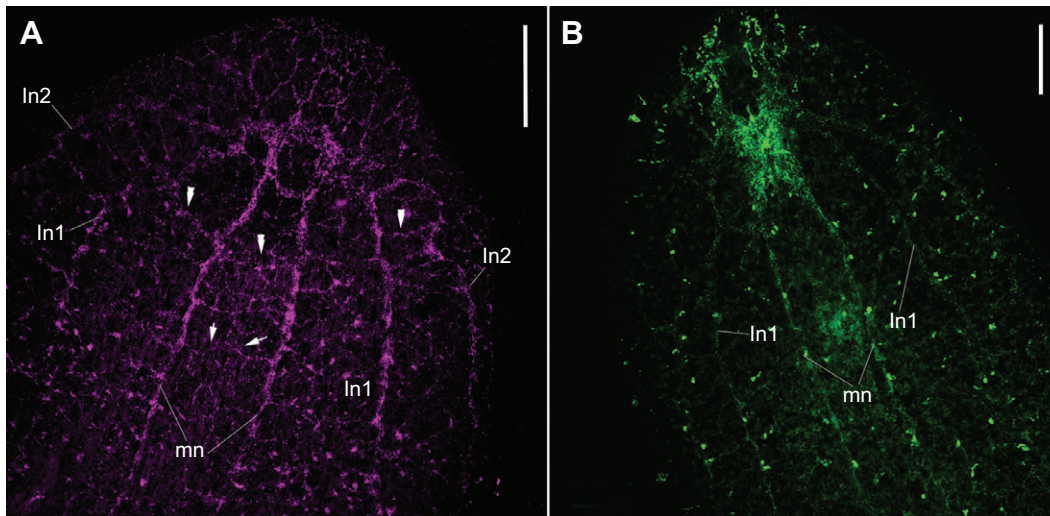


Fig. 2. Synaptotagmin- and Elav-positive structures in the anterior part of the adult *Symsagittifera roscoffensis* nervous system. (A) *S. roscoffensis* stained with synaptotagmin antibody. Three pairs of neurite bundles are shown: the median (mn) the medio-lateral (ln1) and the lateral pairs (ln2). In an irregular pattern, thin commissures connect the neurite bundles all along the anterior–posterior body axis, forming a cellular net (arrowheads). A few, intensely stained, commissures link the neurite bundles in the vicinity of the brain (double arrowheads). Neural processes extend clearly to the most frontal region of the animal. (B) *S. roscoffensis* stained with Elav polyclonal antibody. Two strongly stained medio-lateral neurite bundles (mn) are shown in the most central region, all along the anterior–posterior body axis. A lateral pair (ln1) is lightly stained. A high concentration of neurites is evident in the anterior domain. Some nervous processes emerge from the ‘brain’ and extend to the periphery. Scale bars: 75 μ m.

complex nervous system, with anterior concentrations, a bilobed brain with a cellular cortex and a dense internal neuropile (Fig. 1) (Achatz and Martinez, 2012; Achatz et al., 2013). Thus, the condensation seen in some acoels, particularly those belonging to the Crucimusculata, is clearly an evolutionary derived character within the Xenacoelomorpha (Hejnlol et al., 2009; Philippe et al., 2011).

Since the xenoturbellids represent the most basal clade within the phylum Xenacoelomorpha, the presence of an intraepithelial nerve net of neurones (Raikova et al., 2000b) is now regarded as ancestral – it is also found in the sister group of the Bilateria: the Cnidaria. The nemertodermatids, instead, have an anterior concentration of neurons and processes in the form of a ring commissure; a condition that could be interpreted as intermediate, in terms of structural complexity, between those found in xenoturbellids and the acoels (mostly those belonging to Crucimusculata). Thus we could visualize the evolution of xenacoelomorphs as following a trend towards the anterior concentration of the nervous system. Unfortunately, very little is known about the embryological origin of the nervous system in Xenacoelomorpha. In the following paragraphs we summarize what is known about the embryology of the acoel nervous system and its molecular regulatory mechanisms.

The acoel nervous system development and its molecular control

The nervous system of the acoels originates from the micromeres located at the animal pole, which give rise to the ectodermal lineages, specifically the epidermis and all the neural cells (Henry et al., 2000). Some of the progeny of these micromeres are internalized at later stages of embryonic development and differentiate into neurons. The differentiation of neurons from these internalized cells is readily accounted for by the expression of a *SoxB* orthologue, a gene that is early expressed in pro-neural cells in a great variety of taxa, including non-bilaterian taxa such as the cnidarians (Magie et al., 2005; Hejnlol and Martindale, 2009). Later on in development, the Hox patterning system plays a clearly

important role in providing positional cues to the differentiating neurons. Only three *Hox* genes (one anterior, one central and one posterior) have been identified in the acoel species studied to date (Hejnlol and Martindale, 2009; Moreno et al., 2009; Sikes and Bely, 2010).

The two most anterior *Hox* genes seem to be restricted to ectodermalneural expression, although further corroboration is needed, but the posterior *Hox* gene is expressed in cells encompassing the three germ layers. Besides these genes, few other ‘classical’ neural genes have been studied in the acoel species. The orthologue of the posterior *ParaHox* gene caudal (*Cdx*) is also expressed in anterior neural structures in the acoel *Convolutriloba longifissura* (Hejnlol and Martindale, 2009), probably in the same position as, or in close vicinity to, the cells expressing the neural genes *Nk2.1* and *Otp* (Hejnlol and Martindale, 2008).

Although scarce, most of the molecular data relates to genes involved in antero–posterior patterning in the acoel nervous system, whereas no information is available regarding the regulation of its dorsal–ventral patterning. The identification of such a network should lead to exciting discoveries because, despite having differential dorsal–ventral expression of BMPs (bone morphogenetic proteins) and their antagonists, ADMPs (anti-dorsalizing morphogenetic proteins) (Srivastava et al., 2014), all acoel species have nerve cords equally distributed along their dorso–ventral axis, showing no preference for dorsal or ventral locations. Nothing is known about the control of neurogenesis in both the Xenoturbellids and the Nemertodermatids.

Neuronal regulatory and effector gene products (BHLHs and GPCRs)

In order to understand the process of neurogenesis in the different members of the Xenacoelomorpha, we undertook a thorough study of the members of gene families that are well known for their relationship with the formation of neural tissue. Our starting point was two gene families closely related with both the development and the functionality of the nervous system: the so-called basic

helix–loop–helix (bHLH) and the G-protein-coupled receptor (GPCR) families. Changes in numbers and subfamilies, and eventually the study of expression patterns for all these genes and within Xenacoelomorpha, should provide us with key insights into the process of progressive ‘cephalization’ that occurs in this clade. A summary description of these families follows.

A key, well-known group of transcription factors involved in developmental processes such as cell proliferation and differentiation is the bHLH gene superfamily. This class is present in a wide range of eukaryotes, including fungi, plants and metazoans (Jones, 2004). bHLHs are regulators of nervous system development; they control aspects such as neural fate commitment, cellular subtype specification, migratory behaviour and axonal guidance (Bertrand et al., 2002; Guillemot, 2007). Moreover, recent studies have detected bHLHs expressed in stem cells and neuronal precursors required for CNS regeneration (Cowles et al., 2013). These neuronal precursor proteins possess a highly conserved bHLH DNA-binding domain which, as the name suggests, is composed of a basic region for DNA binding and two α -helices, interrupted by a variable loop region. They normally activate their target genes as dimeric complexes. Some bHLHs also include additional domains involved in protein–protein interactions such as the ‘leucine zipper’, PAS (Per–Arnt–Sim) or ‘orange’.

Molecular phylogenetic analysis has revealed 45 orthologous families of bHLHs in metazoans (Simionato et al., 2007), all included in higher-order groups named A, B, C, D, E and F, which are related through evolution and have similar structural characteristics (Atchley and Fitch, 1997; Jones, 2004). Groups A and B bind to core DNA sequences named ‘Eboxes’ (CANNTG). Group C includes a PAS domain in addition to the HLH domain, binding to the ACGTG or GCGTG core sequences. Members of Group D are unable to bind DNA, but act as negative regulators of group A proteins. Group E contains the proteins related to the *Drosophila* HER (Hairy and enhancer of split bHLH) proteins and also includes two different domains: ‘orange’ and WRPW. Proteins with a COE domain but lacking a basic domain form the last group, F.

GPCRs form the largest and most diverse super family of transmembrane receptors, and they are very abundant in most eukaryotic species. Nevertheless, their total number differs widely between species and subtypes. Within neural tissue they are involved in sensory functions such as photosensitivity or olfactory sensation and also in the regulation of homeostasis. GPCRs are known to share a highly conserved seven transmembrane domain (7TM) that can be used as a motif in genomic analysis to search for and identify this superfamily of genes. Despite the conservation of this domain, GPCRs bind to a very diverse array of ligands and intracellular signalling molecules. This specificity is due to their highly variable N-termini that allow us to classify them into five main families following the so-called GRAFS system: glutamate, rhodopsin, adhesion, frizzled and secretin (Fredriksson et al., 2003; Schiöth and Fredriksson, 2005). The rhodopsin family (also known as class A) is the largest. Rhodopsins have short N-termini but interact with a great variety of ligands (such as amines, purines, lipids, peptides or glycoproteins) through the extracellular portion of the transmembrane domain. This family can be further subdivided in four groups (α , β , γ and δ) and 13 subfamilies. The adhesion family is characterized by long N-termini with different functional domains, such as cadherin, laminin, and calcium-binding epidermal growth factor domains. The secretin 7TM domain is very similar to that of adhesion, from which it evolved. This family, belonging to class B, together with Adhesion, can be recognized by an extracellular hormone-binding domain. The glutamate family (also

known as class C) comprises neuronal modulators with a typical long N-terminal ligand-binding domain known as the ‘Venus flytrap module’. The members of the last family, frizzled (class F), act as receptors of the Wnt pathway, and are well known for their crucial roles in early embryonic development, tissue polarity and cell signalling (Strotmann et al., 2011; Krishnan et al., 2013). Understanding the evolutionary history of GPCRs and bHLHs in Xenacoelomorpha opens a window on knowledge of how upstream regulatory and downstream effector genes have evolved within this phylum.

RESULTS AND DISCUSSION

Structure of the *Symsagittifera roscoffensis* nervous system revealed through species-specific antibodies

In order to better understand the structure and development of the nervous system of the acoel *Symsagittifera roscoffensis* von Graff 1891, we generated specific antibodies against ‘pan-neuronal’ epitopes using information derived from the Genome Project. We focused our attention on two well-known proteins: the synaptic protein synaptotagmin and the RNA-binding factor Elav. Here, we provide a summary description of these expression patterns, as an example of how we can generate markers that will aid in the interpretation of gene patterns, such as those of regulatory and terminal differentiation regulators.

Adult synaptotagmin-positive structures

The adult nervous system, as revealed by the antibody against synaptotagmin, consists of six longitudinal neurite bundles (a dorsal pair, a medio-ventral pair and a ventro-lateral pair) running along most of the length of the organism. They are clearly interconnected through very thin commissures, forming a small, mostly irregular, net not completely visible at the posterior end (Fig. 2A, arrows) contrary to the observations made with the anti-5-HT antibody (Semmler et al., 2010). The only prominent commissures are located in the brain area. Immediately posterior to the brain, five commissures can be detected that directly connect all the bundles (Fig. 2A, double arrows). As we move further posterior from the rostral domain, the neurite bundles become less intensely stained. This pattern is reminiscent of that obtained with both the anti-5-HT and the anti-tyrosinated α -tubulin antibodies described in previous studies (Semmler et al., 2010). However, in those patterns at least two more commissures clearly connect the dorsal neurite bundles and the medio-ventral ones, whereas in our experiments, the commissures are less prominent. Common to all known patterns are the three commissures immediately posterior to the brain. Perhaps because of the specificity of the anti-synaptotagmin antibody, the anterior staining appears to be more detailed than has been previously reported (Fig. 2A).

Adult Elav-positive structures

The *elav* genes, which are also used as neuronal markers, encode highly conserved RNA-binding proteins that are, most probably, involved in neural development and differentiation processes across Bilateria. As has been showed for the cnidarian *Nematostella vectensis*, post-transcriptional regulation by these proteins is an ancient feature shared by Cnidaria and Bilateria (Nakanishi et al., 2012).

In *S. roscoffensis*, Elav immunoreactivity is observed in the most anterior part of adult animals (the brain); with two neurite bundles clearly visible compared with the other two pairs that are almost undetectable (Fig. 2B). Through the posterior part of the body axis the pattern becomes progressively less intense and the neurite

bundles become thinner and thinner in appearance. We note that the thickness of the dorsal neurite bundles decreases, reflecting the presence of less well-packed fibres. Anterior to the statocyst, on the ventral side, the neurite bundles converge into an area (Fig. 2B). From this structure, neural processes extend frontally, laterally and dorsally.

It is interesting to note here that when we compare the immunochemical patterns obtained by Semmler et al. (Semmler et al., 2010) in the adult *S. roscoffensis* with those obtained here with the specific anti-Elav antibody, some clear differences arise. The 5HT and the RFamide immunoreactivity patterns described in that study show an extensive set of neural structures covering the whole body of the animal, whereas the Elav immunoreactivity pattern seems to be more limited to the anterior region of the adult. The dorsal neurite bundles are notably labelled with Elav, in clear contrast to the four ventral bundles, which are almost imperceptible, unlike the data published on the use of RFamide and serotonin antibodies (Semmler et al., 2010). Double and triple staining will accurately define the relative domains identified by all these antibodies.

Exploring the genome of Xenacoelomorpha – characterizing gene families involved in neurogenesis bHLHs

We identified the putative bHLH genes in the sequenced genomes of *S. roscoffensis* and the xenoturbellid *Xenoturbella bocki* Westblad 1949. The genomes comprise 18 genes in *S. roscoffensis* and 33 in *X. bockii*. The phylogenetic analysis performed to classify them used known protein sequences from several metazoans: *Homo sapiens*, *Drosophila melanogaster*, *N. vectensis*, *Acropora digitifera* and *Hydra magnipapillata*, courtesy of Dr Gyoja (Gyoja et al., 2012) (Table 1). To the alignments, we added sequences from *Saccoglossus kowalevskii* and *Capitella teleta*, all downloaded from the PFAM 27.0 data bank (hosted at the EMBL-EBI; see supplementary material Fig. S1). The genes were categorized into six high-order groups: A, B, C, D, E and F (Atchley and Fitch, 1997). In both genomes, *S. roscoffensis* and *X. bocki*, we found no more than one bHLH member per family; we also found that some families had no members. Only one gene from family A had unclear affinity in *S. roscoffensis*. Interestingly, we found no specific cases of genes not belonging to the families already characterized in other bilaterians. The species chosen, including protostomes, deuterostomes and cnidarians, provided us with good representation – although still limited – of the Metazoa. Our trees (see supplementary material Figs S1, S2, bHLH trees) present a similar topology to those obtained in previous studies (Simionato et al., 2007; Gyoja et al., 2012).

The 18 bHLHs characterized in *S. roscoffensis* were classified as members of the following groups: 14 in group A (members of the families: ASCa, ASCb, Beta3, E12/E47, MyoD, Net, NeuroD, Beta3, PTFa, PTFb, twist, and a putative Oligo); three in group B (Max, MTF and SRBP); one in group C (ARNT); and one in group E (HES/HEY). No representatives of group D or F were identified (Table 1). In the case of group A, we found one putative Oligo/Beta3 relative that we could not allocate to one of these two related families because of its low branch support. A similar situation occurred for HES/HEY: we were not able to determine the affinities of the acoel sequence. In fact, the sequence aligned with the similarly unclear *H. sapiens* HES/HEY. Some genes – putative relatives of NeuroD, ASCa, ASCb and ARNT – were sorted, with weak statistical support. Thus, we applied BLAST to them, one by one, and aligned them with the total length of the most closely

related sequences, specifically checking by hand for conserved regions. This procedure provided support for our family assignments (supplementary material Fig. S1, SR_tree).

The 33 bHLHs found in *X. bocki*, were also classified into the different groups: 16 in group A (relatives of: ASCa, ASCb, Atonal, Beta3, E12/E47, hand, mist, MyoRa, MyoRb, neurogenin, NeuroD, NSCL, PTFa, PTFb and twist); 7 in group B (Max, MTF, Mlx, Myc, SRBP, TF4 and USF); one in group C (ARNT); and 4 in group E (HES and HEY). Two sequences, which we named Xb_Orphan-HLH1 and Xb_Orphan-HLH2, did not match with certainty any of the 48 bHLHs included in our analysis. A component clustered in the high-order group B, with weak branch support in the phylogenetic tree, seems to be a putative member of the Net family. However, careful analysis of the sequence did not allow us to assign it to the Net family. Blast analysis suggests that it could be also a diverged orthologue of NeuroD; we provisionally named it Xb_Orphan-HLH3. A similar case occurs with Xb_Orphan-HLH4, with unclear affinities in the phylogenetic tree. Again, the sequences aligned with low bootstrap support (ARNT, NeuroD and Atonal) were checked manually before being assigned to the respective families. A similar thing happened with the acoel sequence related to the HES/HEY families (supplementary material Fig. S2, Xb_tree).

Some species, such as *H. sapiens* and the cnidarian *N. vectensis*, have duplicates in some families (Simionato et al., 2007). Duplicates do not seem to be the norm in *S. roscoffensis*. We found only two orthologues in the PTFa and PTFb families. Even this was not the case for *X. bocki*, which had only one member belonging to each of these families. Therefore, we suggest that the duplications took place within Acoelomorpha and perhaps are specific to Acoela; complete genomes of Nemertodermatida are needed to clarify the issue.

Xenoturbella bocki ‘orphan’ bHLHs

As mentioned above, in *X. bocki* we identified four ‘orphan’ sequences: Xb_Orphan-HLH1, Xb_Orphan-HLH2, Xb_Orphan-HLH3 and Xb_Orphan-HLH4, which do not seem to belong to any of the known bilaterian families. Two hypotheses would explain this: the sequences are homologous to an ancestral bilaterian gene that has been lost in all the bilaterians analysed; or they have drifted enormously from known relatives within the Xenacoelomorpha lineage. The introduction of an itemized representation of species may at some point improve the classification of these four genes. We have not determined the bHLH subfamilies yet: this will require more in-depth analyses in the future.

Remarkably, the high-order groups with representation in the *X. bocki* genome also have members in *S. roscoffensis*; however, the same families are not present in both of them (Table 1). The genome of *S. roscoffensis* has fewer families than that of *X. bocki*. Specifically, the families that are present in xenoturbellids but missing in acoels are: from Group A: Atonal, hand, mist, MyoRa and neurogenin; from Group B: Mlx, Myc, TF4 and USF; and from Group C: HES and HEY. This pattern clearly shows that some losses were produced before the diversification of Acoela. In the absence of comprehensive data from Nemertodermatida, it is difficult to be more accurate as to when these losses occurred. Interestingly, there is one family, and only one, MyoD (plus the problematic case of the putative ‘Oligo’) that has a representative in *S. roscoffensis*, but not in *Xenoturbella*.

When we look globally at all the bHLH genes in both taxa, it is important to note that the bootstrap support is always higher in the *X. bocki* trees than in those including *S. roscoffensis* sequences, although using the same metazoan reference sequences for the alignments. We think this reflects the fact that Acoela appears to

Table 1. Number of bHLH family members

Species	Group A																	Group B											Group C										
	Total	ASCa	ASCB	Amber	Atonal	Bet3	Dellah	E12/E47	Hand	Mesp	Mist	MyoD	MyoRb	MyoRa	Net	NeuroD	Neurogenin	NSCL	Oligo	Paraxis	Peridot	PTFa	PTFb	SCL	Twist														
<i>Xenoturbella bocki</i>	33	2	1	0	1	1	0	1	1	0	1	0	1	1	0(1?)	1	1	1	0	0	0	1	1	0	1														
<i>Synsagittifera roscoffensis</i>	18	1(1?)	1	0	1	0	1	1	0	0	0	1	0	0	1	1	0	1	0(1?)	0	0	2	2	0	1														
<i>Homo sapiens</i> *	118	2	3	0	2	2	0	4	2	3	1	4	2	2	1	4	3	2	3	2	0	1	1	3	2														
<i>Nematostella vectensis</i> *	82	3	1	1	0	0(2?)	0(1?)	1	2(3?)	0(2?)	0(2?)	0	1	0	1	0(1?)	0(1?)	1	0	1	1	0	1	3	1														
<i>Drosophila melanogaster</i> *	59	4	0	0	3	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	2	2	1	1														
<i>Acropora digitifera</i> *	70	3	1	1	0	0	1	1	2	1	0	0	1	0	1	0	0	0	0	1	1	1	1	3	1														
<i>Hydra magnipapillata</i> *	33	3	1	N	0	0(1?)	0(1?)	1	0	0	0(1?)	0	0	0	1	0(1?)	0(1?)	0	0(1?)	1	N	0	1	0	0														

Species	Group B											Group C													
	AP 4	Fig-alpha	MAD	Max	MTTF	Mix	Mnt	Myc	SRBP	Tt4	USF	Ahr	ARNT	Bmal	Clock	Hif	Sim	SRC	Th	Emc	Pearl	HES	HEY	COE	Orphans
<i>Xenoturbella bocki</i>	0	0	0	1	1	1	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	3(4?)	1(2?)	0	2(3?)
<i>Synsagittifera roscoffensis</i>	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0(1?)	0(1?)	0	0
<i>Homo sapiens</i> *	1	1	5	1	5	2	1	4	2	1	3	2	2	2	2	3(4?)	2(3?)	3	1(2?)	4	0	7(10?)	3(6?)	4	4
<i>Nematostella vectensis</i> *	2	0	1	1	1	1	1	4(6?)	2	1	2	2	1(2?)	1(2?)	1	0(2?)	0(2?)	0	0(2?)	1(2?)	1	11(12?)	1(2?)	1	8-16
<i>Drosophila melanogaster</i> *	1	0	0	1	1	1	1	1	1	1	1	2	1	1	3	1	1	1	1	1	0	11	1(2?)	1	0
<i>Acropora digitifera</i> *	1	0	1	1	1	1	1	5	2	1	2	2	1	1	3(4?)	0	0	0	1	1	1	7(9?)	1(3?)	1	14-17
<i>Hydra magnipapillata</i> *	1	0	1	1	1	0	0	4	0	0	1	1	1	0	0	0(1?)	0(1?)	0	0(1?)	0(1?)	ND	1	0	1	6-11

The number of orthologous family members and the total number of bHLHs are indicated. The data for *X. bocki* and *S. roscoffensis* are derived from our analysis. The information from the other representative species (*), *H. sapiens*, *N. vectensis*, *D. melanogaster* and *H. magnipapillata*, was obtained from Simonato et al. (Simonato et al., 2007). The data from *A. digitifera* (plus the latest identifications in *N. vectensis*) was obtained from Gyoja et al. (Gyoja et al., 2012).

have long branches in most phylogenetic (phylogenomic) analyses; with a higher rate of nucleotide substitutions in most genes analysed. Therefore, if the bHLH sequences follow the same general pattern, one should expect that the genome of *S. roscoffensis* encodes, in general, very divergent orthologous sequences. In addition, *S. roscoffensis* belongs to the clade Crummusculata, which includes the more derived acoels, the so-called 'high acoels'. In contrast, *Xenoturbella* appear to have short branches in all molecular phylogenetic analyses.

Of the 48 bHLH families known, 45 are present in Bilateria. Of these, 44 are the families shared by the protostomes and deuterostomes (and may therefore have been present in the urbilaterian ancestor). A total of 29–33 of these families are apparently present in the ancestor of cnidarians and bilaterians (Ledent and Vervoort, 2001; Ledent et al., 2002; Simionato et al., 2007). Surprisingly, the number of families represented in our dataset for *S. roscoffensis* is not in this numeric range. In fact, in *S. roscoffensis* we detected only 16 families. In the genome of *X. bocki* there were more; a total of 29. Nevertheless, the number of families shared between *X. bocki* and the cnidarians analysed (Table 1) (Simionato et al., 2007; Gyoja et al., 2012), ranges between 18 and 23. To us, it has become clear that in the Xenacoelomorpha lineage there have been several family losses, particularly in Acoela (see Simionato et al., 2007).

However, we should be cautious about drawing very general conclusions because we have gathered complete data from the genome of only one acoel species and one xenoturbellid. We need to explore a full range of xenacoelomorphs, incorporating members of Nemertodermatida and also some more basal acoels. A more detailed characterization of bHLH complements in different members of this phylum should provide us with a better understanding of the evolutionary dynamics of this important family of transcriptional regulators. Moreover, with access to the complements and a better description of the neuroanatomy of these animals, we should start to get a better picture of how the evolution of the nervous system, and its centralization, is linked to specific changes in genome composition, in particular for regulatory genes.

GPCRs

We mined the genomes of *S. roscoffensis* and *X. bocki* for the presence of genes containing the typical GPCR 7TM domain of the GRAFS families (Fredriksson et al., 2003; Schiöth and Fredriksson, 2005), plus the members of the Dicty_CAR (see the Materials and methods for more details). GRAFS GPCRs are typically expanded in mammals, especially the rhodopsin family, although most of them have representatives in the last eukaryotic common ancestor (de Mendoza et al., 2014). For *S. roscoffensis*, we further phylogenetically analysed each family to classify all the genes found. The GPCR sequences used as references were provided by Alex de Mendoza (de Mendoza et al., 2014) and were chosen to offer an extensive evolutionary perspective. The species used in our analysis were: *A. queenslandica*, *N. vectensis*, *S. kowalevskii*, *H. sapiens*, *C. teleta* and *D. melanogaster*. The clustering was checked for all the acoel sequences by running BLAST with all the candidates against the NCBI protein database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Analysis of GPCR family numbers

In the genome of *S. roscoffensis* we found for each family the following complements: 9 glutamate, 6 frizzled, 5 adhesion/secretin and 225 rhodopsin receptors; however, we did not find any Dicty_CAR relative, making a total of 245 different GPCRs

(Table 2). In the case of *X. bocki*, the complements were: 17 glutamate receptors, 7 Frizzled, 21 Adhesion/Secretin, 258 Rhodopsin and 1 Dicty_CAR; a total of 304 GPCRs (Table 2). Note that all the *X. bocki* and the *S. roscoffensis* rhodopsin values are given as averages of both the predictions based on the genomic sequence information and the AUGUSTUS gene predictions.

The Dicty_CAR domain corresponds to a cyclic AMP receptor. Frizzled, adhesion and rhodopsin classes are thought to have evolved from such a more primitive receptor type; the former two before the split of unikonts from the common ancestor of eukaryotes, and rhodopsin in the common ancestor of the opisthokonts. After the diversification of those in metazoan, the cAMP receptor might have become somehow redundant and consequently lost or rarely present in most metazoan species (Krishnan et al., 2012).

Table 2 shows that the two species studied have fewer representatives in all the families than in our reference species, except for *D. melanogaster*. It is known that GPCRs are subjected to species-specific diversification, which is especially remarkable in *H. sapiens*, *C. teleta* and *N. vectensis*. Despite the higher complexity of genes in *X. bocki*, compared with *A. queenslandica* and *N. vectensis*, we noticed that the former has, in general, fewer members from each GPCR family (although not in all cases). We also observe this trend within the Xenacoelomorpha monophyletic group, which is where we find *S. roscoffensis*, despite it being a divergent member of Acoela; it has a nervous system morphologically more complex than *X. bocki*, but with fewer GPCRs in all the families.

Glutamate family

In the glutamate family (see supplementary material Fig. S3) we detected three glutamate metabotropic receptors (GMR); which represents a simplification compared with the eight that found in humans, although this may be a direct consequence of the duplication of genomes specific to vertebrates. Intriguingly, *S. roscoffensis* has six γ -aminobutyric acid (GABA) receptors, a greater number than the three types found in humans. It is known that GABA receptors play a role in synaptic transmission and neuronal excitability in the CNS of vertebrates (Niswender and Conn, 2010). It is remarkable that a similar specific expansion with eight GABA receptors is observed in *S. kowalevski*, which does not have a centralized nervous system and lacks neurons that release GABA (Aronowicz and Lowe, 2006; Krishnan et al., 2013). However, although *S. kowalevski* does not have an obviously centralized nervous system and *S. roscoffensis* has a clear brain, both systems are structurally simpler than vertebrates.

Frizzled family

We also found six Frizzled-like receptors: three clustered together with Frizzled 1/2/7, one with Frizzled 5/8 and two more with Frizzled 9/10 (supplementary material Fig. S4). We did not detect any Smoothed-like receptor. Interestingly, the frizzled family is one of the most evolutionarily conserved because of its fundamental role in cell signalling and tissue polarity, where they function as Wnt receptors (Lagerström and Schiöth, 2008). With 11 members in humans, 6 in the acoels studied and 7 in the xenoturbellids (Table 2), we did not observe a species-specific expansion in this subfamily (Strotmann et al., 2011).

Adhesion and secretin families

We detected five members of the adhesion/secretin family in the acoel genome, which represents a very low number when compared with the reference species, even with the xenoturbellids, in whose

Table 2. Number of GPCRs of each family in the indicated species

Species	Rhodopsins	Adhesion/Secretin	Glutamate	Frizzled	Dicty_CAR	Total
<i>Amphimedon queenslandica</i> *	147	64	43	4	0	258
<i>Nematostella vectensis</i> *	1049	58	37	6	2	1152
<i>Xenoturbella bocki</i>	258	21	17	7	1	304
<i>Symsagittifera roscoffensis</i>	225	5	9	6	0	245
<i>Saccoglossus kowalevskii</i> *	246	22	23	3	0	294
<i>Homo sapiens</i> *	297**	48	23	11	0	379
<i>Capitella teleta</i> *	980	39	51	5	1	1076
<i>Drosophila melanogaster</i> *	71	18	10	5	0	104

* From de Mendoza et al. (de Mendoza et al., 2014).

** The number of human rhodopsins varies according to the level of stringency utilized in the search. In other publications the number rises to 683 (Krishnan et al., 2012).

Note that the values corresponding to *X. bocki* and *S. roscoffensis* rhodopsin numbers are the obtained as averages of predictions derived from the raw genomic sequence and the AUGUSTUS gene predictions.

genome we detected 21 members. Of those five (supplementary material Fig. S5), only one clearly clustered with the human homologues GPR123, GPR124 and GPR125. After a BLAST search, we assigned this gene, Sr_GPR125 (supplementary material Fig. S5, fourth arrow), as a putative orthologue to human GPR125 (Bjarnadóttir et al., 2007).

Secretin receptors evolved from adhesion receptors and therefore they share a similar 7TM core sequence. For this reason, we detected and analysed the members of both families together. They are present in most bilaterians, but not in older members of the Metazoa (Strotmann et al., 2011). It would be interesting to ascertain which of the genes detected belong to the adhesion family and whether there is any secretin homologue. In fact, the gene we call Sr_orphan1_aug3sy.g27699.t1 (supplementary material Fig. S5, first arrow) would be a good candidate secretin relative because it clusters with some members of the *C. teleta* and *H. sapiens* secretin family members, although in a basal position. However, further analyses are necessary to confirm this finding.

Rhodopsin family

The rhodopsin family is the largest and most diversified, and thus the most difficult to analyse. To complicate it further, some of the reference species genomes are not yet fully annotated, whereas our acoele sequences are extremely divergent. In order to resolve these issues, we aligned all members of this subfamily with human reference sequences, which is the most reliable and studied reference genome. What we observed in this study is that most of the sequences we identified in *S. roscoffensis* do not have a direct human orthologue; they tend to form independent groups instead. In the future, the same detailed analysis of *S. roscoffensis* should be performed on the different families of GPCRs found in the genome of *X. bocki*.

Conclusions

The phylum Xenacoelomorpha is formed of animals with a relatively simple morphology, including that of their nervous systems (Raikova et al., 2000a; Raikova et al., 2000b; Raikova et al., 2004a; Raikova et al., 2004b; Achatz et al., 2013). Here, and for the first time using the genome sequence of members of this group, we have undertaken a systematic study of two superfamilies, the bHLHs and GPCRs, involved in the specification and the functionality of nerve tissue. The main aim of this study was to correlate the molecular complexity of these families with the structural complexity of the different nervous systems within the phylum. We understand that the term ‘complexity’ is a loaded one, and there are many putative definitions for such a word, especially

in the context of biological systems. However, here we use a very narrow concept of complexity: one that reflects the degree of neuronal concentration in one pole of the animal, which implies a high degree of interconnectivity and a high capacity for processing external inputs. This is a definition that would assume that ‘cephalized’ (centralized) nervous systems are structurally more complex than ‘simpler’ neuronal nets (as seen, for instance in members of the Cnidaria). Several morphological studies (mentioned above) have shown that whereas *Xenoturbella* seems to comprise an intraepidermal net of nerve cells, some members of Acoela, particularly those belonging to the clade Crucimusculata, have a condensed bilobed brain with a dense neuropile in the most anterior end of the animal (close to several sensory structures). Our molecular data show that in the case of the two families analysed, the complement of genes (and its diversity) is higher in *Xenoturbella* than in *Symsagittifera*. This would mean that there is an inverse correlation between the structural complexity of the nervous system and the number of genes involved in both patterning and downstream sensory functions. This putative paradox requires and deserves further analysis. We are well aware that the view that we have of both genomes is still incomplete. Moreover, structural complexity, perhaps just superficial, does not have to reflect the functional complexity or the complexity of circuits in the different nervous systems. The study of gene activities and analysis of their evolving regulatory interactions should give us a better, richer understanding of the evolution of these nervous systems; and perhaps it will provide insight into the putative ‘cephalizing’ trends that we observe in this and other clades (Moroz, 2012). Knowledge of the genomes is now, in any case, paving the way to decipher the evolution of the brain primordia.

MATERIALS AND METHODS

Immunohistochemistry

Immunostaining was performed following the protocols outlined in Achatz and Martinez (Achatz and Martinez, 2012). *S. roscoffensis* was incubated in primary anti-synaptotagmin (dilution 1:200) or anti-Elav (dilution 1:800) antibodies (both previously pre-absorbed) and reacted with the secondary antibody [Alexa Fluor goat anti-rabbit 532 (Molecular Probes, Eugene, OR)].

Image acquisition, processing and preparation

Confocal laser-scanning microscopy was performed with a Leica TCS SPE or a Leica SPlI microscope (Leica Microsystems, Weztlar, Germany). Images were processed using ImageJ (Abramoff et al., 2004) and finished with Adobe Photoshop CS. Diagrams were produced using Adobe Illustrator 7.0.

Genome sequences, annotation and selection

Genome assemblies including scaffolds and contigs (SCFs), transcriptome sequences assembled from RNA-seq reads (ESTs), as well as gene annotation over genome sequences (AUGs), for both *S. roscoffensis* and *X. bocki*, were downloaded from the server of the University of Greifswald (sequences will be publicly available along with the publication of the genome). SOAPdenovo (v2) (Luo et al., 2012) and SOAPdenovo-Trans (v1) (Xie et al., 2014) were used in the protocol for the genome and transcriptome assemblies, respectively. A repeat library was generated using RepeatScout (v1.0.5) (Price et al., 2005). RepeatMasker (v3.3.30) (Smit et al., 2014) was used to produce a repeat-masked genome version with the above mentioned custom library, and to produce a table that lists repeat coordinates. CEGMA (v2.4) (Parra et al., 2007) was used to produce a set of ultra-conserved core protein genes from the unmasked genome. A second initial gene set was generated by WebAUGUSTUS (Hoff and Stanke, 2013), which ran PASA (Haas et al., 2003) on assembled RNA-Seq data and the unmasked genome from *Xenoturbella*. Both gene sets were merged (non-redundantly) and used for optimizing AUGUSTUS parameters (Stanke, et al., 2008). RNA-Seq reads were mapped against the repeat masked genome, alignments were converted to hints for AUGUSTUS. PASA and CEGMA results as well as the RepeatMasker output table were also converted to hints. AUGUSTUS was run on the unmasked genome using the optimized parameter set and all available hints. Genomic and transcriptomic nucleotide sequences were translated into all possible open reading frames (ORFs) larger than 20 amino acids using custom Perl scripts; whereas AUGUSTUS predictions already included putative proteins translated from the predicted transcripts. The described annotation protocol produced the sets of putative transcript and protein sequences that were used in this study. A search for specific protein domains was performed on all the amino acid sequences from those three sets – ORFs from genome and transcriptome sequences and proteins from AUGUSTUS predictions. HMMER (v3.0) (Eddy, 2011) was used to scan the sequences using the following Pfam (v27.0) (Finn et al., 2014) hidden Markov models (HMM): PF00001 (7tm_1), PF00002 (7tm_2, Adhesion/Secretin), PF00003 (7tm_3, Glutamate), PF01534 (Frizzled) and PF05462 (Dicty_CAR) for the GPCR family; and PF00010 (HLH DNA-binding domain) for the HLH family. hmmsearch was run with the following parameters: ‘max’ (to disable all heuristics, which entails less speed but more power), ‘E 1e-5’ (maximum total E-value to report sequences) and ‘domE 1e-5’ (maximum domain E-value to report a hit). Sequences with a significant hit for any of the previously described HMM models were then filtered out from the set of genomic (SCFs), transcriptomic (ESTs) and predicted (AUGs) proteins for further analyses.

Phylogenetic analysis

The best predictions for each domain search and dataset were manually curated and aligned with ClustalW2 (Larkin et al., 2007; McWilliam et al., 2013). We have manually filtered the more informative columns of the alignment in order to produce the presented phylogenetic trees. Then, final predictions for each family, or subfamily in the case of GPCRs, were aligned against a database of protein sequences from different species, including *H. sapiens*, *D. melanogaster*, *S. kowalevski*, *N. vectensis* and *C. teleta*. In the case of GPCRs, sequences from *A. queenslandica* were considered; for the bHLHs family, sequences from *A. digitifera* and *H. magnipapillata* were also considered. GPCR reference sequences were classified in families as described by de Mendoza et al. (de Mendoza et al., 2014) whereas the Gyoja and Satoh (Gyoja and Satoh, 2013) reference was used for bHLHs. bHLH sequences for *S. kowalevski* and *C. teleta* were directly downloaded from the Pfam database preserving the original access codes.

Full alignments were performed using MAFFT (v6.864b) (Kato and Toh, 2008) with the ‘max’ iterate 1000 and ‘local pair’ (equivalent to L-INS-i) options. Conserved regions from those alignments were manually selected using Geneious software (v5.3.6, Biomatters, <http://www.geneious.com>). Maximum-likelihood analysis was carried out using RAXML (v8.0.12) (Stamatakis, 2014), with the amino acid evolutionary model LG using a discrete gamma distribution of among-site variation rates and a proportion of invariable sites (‘m PROTGAMMAILG’ option). The best tree out of 100 bootstrap replicates (N 100) was chosen. A neighbour-joining approach was also considered and the trees were inferred by Neighbour-Joining Tree

Builder (njtree v1.9.1) (Li, 2006). Finally, the computed trees were edited and processed into images with FigTree (v1.4.1) (Rambaud, 2007), with further refinements and details added using Adobe Illustrator 7.0.

All the alignments generated with the manually filtered conserved columns can be found in Fasta format as supplementary material Figs S6 to S10. All the sequences included in our study keep the original reference number assigned in the preliminary genome/transcriptome analyses. The whole genome/transcriptome original sequences will be made publicly available along with the forthcoming publication of the Xenacoelomorpha Genome Project consortium paper.

Acknowledgements

This work was presented at the ‘Evolution of the First Nervous Systems II’ meeting, which was supported by the National Science Foundation (NSF). The Meeting was held at the Whitney Marine Laboratory, St. Augustine, Florida. We would like to thank specifically Alex de Mendoza (IBE CSIC-UPF, Barcelona) and Fuki Gyoja (Okinawa Institute of Science and Technology) for access to their GPCR and bHLH databases. Moreover, A. de Mendoza was especially helpful through the whole process on analysing and classifying the GPCRs. We appreciate the support given by Manel Bosch and the access to the CLSM Microscopy facility of the Universitat de Barcelona. We would like to thank the referees for the suggestions and comments that were instrumental in the improvement of the submitted manuscript. Finally, we acknowledge the continuous, insightful, discussions with all the members of the Xenacoelomorpha genome consortium.

Competing interests

The authors declare no competing or financial interests.

Author contributions

All authors contributed to the design of the experiments and to the writing of the manuscript. A.P. generated raw-sequencing data and provided the assemblies. K.H. provided all gene models. J.F.A. did the initial screening of functional domains over the genomic and transcriptomic sequences. E.P.-A. and B.G. performed all the genomic and phylogenetic analysis. E.P.-A. and B.G. performed all the immunochemical analysis appearing in this paper. M.C. contributed to the design and testing of the anti-peptide antibodies. All authors contributed to the writing of the manuscript. P.M. directed the whole project.

Funding

We would like to acknowledge the generous support obtained from the EU Programme ‘ASSEMBLE’, which covered some of the costs associated to our acoel collection trips. A grant from the Spanish Ministry of Economy (BFU2012-32806) funds the work in the laboratory of P. Martinez. Funds from the Generalitat de Catalunya (2009-SGR-001018) were also used in this project. E.P.-A. has a PhD (APIF-UB) fellowship from the Universitat de Barcelona.

Supplementary material

Supplementary material available online at <http://jeb.biologists.org/lookup/suppl/doi:10.1242/jeb.110379/-DC1>

References

- Abramoff, M. D., Magalhaes, P. J. and Ram, S. J. (2004). Image Processing with ImageJ. *BioPhotonics International* **11**, 36–42. Available at: <http://imagej.nih.gov/ij/>
- Achatz, J. G. and Martinez, P. (2012). The nervous system of *Isodiametra pulchra* (Acoela) with a discussion on the neuroanatomy of the Xenacoelomorpha and its evolutionary implications. *Front. Zool.* **9**, 27.
- Achatz, J. G., Chiodin, M., Salvenmoser, W., Tyler, S. and Martinez, P. (2013). The Acoela: on their kind and kinships, especially with nemertodermatids and xenoturbellids (Bilateria incertae sedis). *Org. Divers. Evol.* **13**, 267–286.
- Aronowicz, J. and Lowe, C. J. (2006). Hox gene expression in the hemichordate *Saccoglossus kowalevskii* and the evolution of deuterostome nervous systems. *Integr. Comp. Biol.* **46**, 890–901.
- Atchley, W. R. and Fitch, W. M. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proc. Natl. Acad. Sci. USA* **94**, 5172–5176.
- Baguña, J. and Riutort, M. (2004). The dawn of bilaterian animals: the case of acoelomorph flatworms. *BioEssays* **26**, 1046–1057.
- Bertrand, N., Castro, D. S. and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nat. Rev. Neurosci.* **3**, 517–530.
- Bery, A. and Martinez, P. (2011). Acetylcholinesterase activity in the developing and regenerating nervous system of the acoel *Symsagittifera roscoffensis*. *Acta Zoologica* **92**, 383–392.
- Bery, A., Cardona, A., Martinez, P. and Hartenstein, V. (2010). Structure of the central nervous system of a juvenile acoel, *Symsagittifera roscoffensis*. *Dev. Genes Evol.* **220**, 61–76.
- Bjarnadóttir, T. K., Geirardóttir, K., Ingemansson, M., Mirza, M. A., Fredriksson, R. and Schiöth, H. B. (2007). Identification of novel splice variants of Adhesion G protein-coupled receptors. *Gene* **387**, 38–48.

- Carranza, S., Baguñà, J. and Riutort, M. (1997). Are the Platyhelminthes a monophyletic primitive group? An assessment using 18S rDNA sequences. *Mol. Biol. Evol.* **14**, 485-497.
- Cebrià, F. (2008). Organization of the nervous system in the model planarian *Schmidtea mediterranea*: an immunocytochemical study. *Neurosci. Res.* **61**, 375-384.
- Correa, D. D. (1960). Two new marine Turbellaria from Florida. *Bull. Mar. Sci.* **10**, 208-216.
- Cowles, M. W., Brown, D. D., Nisperos, S. V., Stanley, B. N., Pearson, B. J. and Zayas, R. M. (2013). Genome-wide analysis of the bHLH gene family in planarians identifies factors required for adult neurogenesis and neuronal regeneration. *Development* **140**, 4691-4702.
- Crezée, M. (1975). Monograph of the Solenofilomorphae (Turbellaria: Acoela). *Internationale Revue der gesamten Hydrobiologie und Hydrographie* **60**, 769-845.
- Crezée, M. (1978). *Paratomella rubra* Rieger and Ott, an amphiatlantic acoel turbellarian. *Cah. Biol. Mar.* **19**, 1-9.
- de Mendoza, A., Sebó-Pedros, A. and Ruiz-Trillo, I. (2014). The evolution of the GPCR signaling system in eukaryotes: modularity, conservation, and the transition to metazoan multicellularity. *Genome Biol. Evol.* **6**, 606-619.
- Eddy, S. R. (2011). Accelerated profile HMM Searches. *PLOS Comput. Biol.* **7**, e1002195.
- Ehlers, U. (1985). *Das Phylogenetische System der Plathelminthes*. Stuttgart: Gustav Fischer.
- Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J. et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222-D230. Available at: <http://pfam.xfam.org/>
- Fredriksson, R., Lagerström, M. C., Lundin, L. G. and Schiöth, H. B. (2003). The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256-1272.
- Grimmelikhuijzen, C. J. P., Williamson, M. and Hansen, G. N. (2002). Neuropeptides in cnidarians. *Can. J. Zool.* **80**, 1690-1702.
- Guillemot, F. (2007). Spatial and temporal specification of neural fates by transcription factor codes. *Development* **134**, 3771-3780.
- Gyoja, F. and Satoh, N. (2013). Evolutionary aspects of variability in bHLH orthologous families: insights from the pearl oyster, *Pinctada fucata*. *Zool. Sci.* **30**, 868-876.
- Gyoja, F., Kawashima, T. and Satoh, N. (2012). A genomewide survey of bHLH transcription factors in the coral *Acropora digitifera* identifies three novel orthologous families, pearl, amber, and peridot. *Dev. Genes Evol.* **222**, 63-76.
- Haas, B. J., Delcher, A. L., Mount, S. M., Wortman, J. R., Smith, R. K., Jr, Hannick, L. I., Maiti, R., Ronning, C. M., Rusch, D. B., Town, C. D. et al. (2003). Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654-5666.
- Haszprunar, G. (1996). Plathelminthes and Plathelminthomorpha- paraphyletic taxa. *J. Zool. Syst. Evol. Res.* **34**, 41-48.
- Hejnal, A. and Martindale, M. Q. (2008). Acoel development indicates the independent evolution of the bilaterian mouth and anus. *Nature* **456**, 382-386.
- Hejnal, A. and Martindale, M. Q. (2009). Coordinated spatial and temporal expression of Hox genes during embryogenesis in the acoel *Convolutriloba longifissura*. *BMC Biol.* **7**, 65.
- Hejnal, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G. W., Edgecombe, G. D., Martinez, P., Bagnuà, J., Bailly, X., Jondelius, U. et al. (2009). Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc. Biol. Sci.* **276**, 4261-4270.
- Henry, J. Q., Martindale, M. Q. and Boyer, B. C. (2000). The unique developmental program of the acoel flatworm, *Neochilidia fusca*. *Dev. Biol.* **220**, 285-295.
- Hoff, K. J. and Stanke, M. (2013). WebAUGUSTUS – a web service for training AUGUSTUS and predicting genes in eukaryotes. *Nucleic Acids Res.* **41**, W123-W128.
- Hyman, L. H. (1951). *The Invertebrates, Platyhelminthes and Rhynchocoela*, Vol. II. New York, NY: McGraw-Hill Book Co.
- Israelsson, O. (2007). Ultrastructural aspects of the 'statocyst' of *Xenoturbella* (Deuterostomia) cast doubt on its function as a georeceptor. *Tissue Cell* **39**, 171-177.
- Jondelius, U., Wallberg, A., Hooge, M. and Raikova, O. I. (2011). How the worm got its pharynx: phylogeny, classification and Bayesian assessment of character evolution in Acoela. *Syst. Biol.* **60**, 845-871.
- Jones, S. (2004). An overview of the basic helix-loop-helix proteins. *Genome Biol.* **5**, 226.
- Katoh, K. and Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* **9**, 286-298.
- Krishnan, A., Almén, M. S., Fredriksson, R. and Schiöth, H. B. (2012). The origin of GPCRs: identification of mammalian like Rhodopsin, Adhesion, Glutamate and Frizzled GPCRs in fungi. *PLoS ONE* **7**, e29817.
- Krishnan, A., Almén, M. S., Fredriksson, R. and Schiöth, H. B. (2013). Remarkable similarities between the hemichordate (*Saccoglossus kowalevskii*) and vertebrate GPCR repertoire. *Gene* **526**, 122-133.
- Lagerström, M. C. and Schiöth, H. B. (2008). Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discov.* **7**, 339-357.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R. et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.
- Ledent, V. and Vervoort, M. (2001). The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome Res.* **11**, 754-770.
- Ledent V., Paquet O., Vervoort M. (2002). Phylogenetic analysis of the human basic helix-Loop-Helix proteins. *Genome Biol.* **3**, 0030.1-0030.18.
- Li, H. (2006). *Constructing the TreeFam Database*. PhD thesis, The Institute of Theoretical Physics, Chinese Academy of Science, Beijing, China. Available at: <http://treesoft.sourceforge.net/njtree-old.shtml>
- Lundin, K. (1997). Comparative ultrastructure of the epidermal ciliary rootlets and associated structures in species of the Nemertodermatida and Acoela (Plathelminthes). *Zoomorphology* **117**, 81-92.
- Lundin, K. (2000). Phylogeny of the Nemertodermatida (Acoelomorpha, Platyhelminthes). A cladistic analysis. *Zool. Scr.* **29**, 65-74.
- Lundin, K. and Hendelberg, J. (1996). Degenerating epidermal bodies ("pulsatile bodies") in *Mearia stichopi* (Platyhelminthes, Nemertodermatida). *Zoomorphology* **116**, 1-5.
- Lundin, K. and Sterrer, W. (2001). The Nemertodermatida. In *Interrelationships of the Platyhelminthes (Systematics Association Special Volume Series)*, Vol. 60 (ed. D. T. J. Littlewood and R. A. Bray), pp. 24-27. London: Taylor and Francis.
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y. et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* **1**, 18.
- Magie, C. R., Pang, K. and Martindale, M. Q. (2005). Genomic inventory and expression of Sox and Fox genes in the cnidarian *Nematostella vectensis*. *Dev. Genes Evol.* **215**, 618-630.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. and Lopez, R. (2013). Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* **41** W597-W600. Available at: <http://www.ebi.ac.uk/Tools/msa/clustalw2/>
- Moreno, E., Nadal, M., Bagnuà, J. and Martínez, P. (2009). Tracking the origins of the bilaterian Hox patterning system: insights from the acoel flatworm *Symsagittifera roscoffensis*. *Evol. Dev.* **11**, 574-581.
- Moroz, L. L. (2012). Phylogenomics meets neuroscience: how many times might complex brains have evolved? *Acta Biol. Hung.* **63** Suppl. 2, 3-19.
- Nakanishi, N., Renfer, E., Technau, U. and Rentzsch, F. (2012). Nervous systems of the sea anemone *Nematostella vectensis* are generated by ectoderm and endoderm and shaped by distinct mechanisms. *Development* **139**, 347-357.
- Niswender, C. M. and Conn, P. J. (2010). Metabotropic glutamate receptors: physiology, pharmacology, and disease. *Annu. Rev. Pharmacol. Toxicol.* **50**, 295-322.
- Parra, G., Bradnan, K. and Korf, I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**, 1061-1067.
- Philippe, H., Brinkmann, H., Martinez, P., Riutort, M. and Bagnuà, J. (2007). Acoel flatworms are not platyhelminthes: evidence from phylogenomics. *PLoS ONE* **2**, e717.
- Philippe, H., Brinkmann, H., Copley, R. R., Moroz, L. L., Nakano, H., Poustka, A. J., Wallberg, A., Peterson, K. J. and Telford, M. J. (2011). Acoelomorph flatworms are deuterostomes related to *Xenoturbella*. *Nature* **470**, 255-258.
- Price, A. L., Jones, N. C. and Pevzner, P. A. (2005). De novo identification of repeat families in large genomes. *Bioinformatics* **21** Suppl. 1, i351-i358.
- Raikova, O. I. (2004c). *Neuroanatomy of Basal Bilaterians (Xenoturbellida, Nemertodermatida, Acoela) and its Phylogenetic Implication*. PhD thesis, Åbo Akademi, Turku, Finland.
- Raikova, O. I. (2008). Neurophylogeny of early bilaterians: Acoela, Nemertodermatida, Xenoturbella. *J. Morphol.* **269**, 1459-1460.
- Raikova, O. I., Reuter, M., Kotikova, E. A. and Gustafsson, M. K. S. (1998). A commissural brain! The pattern of 5-HT immunoreactivity in Acoela (Platyhelminthes). *Zoomorphology* **118**, 69-77.
- Raikova, O. I., Reuter, M., Jondelius, U. and Gustafsson, M. K. S. (2000a). The brain of the Nemertodermatida (Platyhelminthes) as revealed by anti-5HT and anti-FMRamide immunostainings. *Tissue Cell* **32**, 358-365.
- Raikova, O. I., Reuter, M., Jondelius, U. and Gustafsson, M. K. S. (2000b). An immunocytochemical and ultrastructural study of the nervous and muscular systems of *Xenoturbella westbladi* (Bilateria inc. sed.). *Zoomorphology* **120**, 107-118.
- Raikova, O. I., Reuter, M. and Justine, J. L. (2001). Contributions to the phylogeny and systematic of the Acoelomorpha. In *Interrelationships of the Platyhelminthes (Systematics Association Special Volume Series)* (ed. D. T. J. Littlewood and R. A. Bray), pp. 13-23. London: Taylor and Francis.
- Raikova, O. I., Reuter, M., Gustafsson, M. K. S., Maule, A. G., Halton, D. W. and Jondelius, U. (2004a). Evolution of the nervous system in Paraphanostoma (Acoela). *Zool. Scr.* **33**, 71-88.
- Raikova, O. I., Reuter, M., Gustafsson, M. K. S., Maule, A. G., Halton, D. W. and Jondelius, U. (2004b). Basiepidermal nervous system in Nemertodermatida westbladi (Nemertodermatida): GYIRFamide immunoreactivity. *Zoology* **107**, 75-86.
- Rambaud, A. (2007). *FigTree version 1.4.1*. Available at: <http://tree.bio.ed.ac.uk/software/figtree/>
- Reuter, M., Raikova, O. I., Jondelius, U., Gustafsson, M. K., Maule, A. G. and Halton, D. W. S. (2001a). Organisation of the nervous system in the Acoela: an immunocytochemical study. *Tissue Cell* **33**, 119-128.
- Reuter, M., Raikova, O. I. and Gustafsson, M. K. S. (2001b). Patterns in the nervous and muscle systems in lower flatworms. *Belg. J. Zool.* **131** Suppl. 1, 47-53.
- Ruiz-Trillo, I., Riutort, M., Littlewood, D. T., Herniou, E. A. and Bagnuà, J. (1999). Acoel flatworms: earliest extant bilaterian Metazoans, not members of Platyhelminthes. *Science* **283**, 1919-1923.
- Schiöth, H. B. and Fredriksson, R. (2005). The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen. Comp. Endocrinol.* **142**, 94-101.

- Semmler, H., Chiodin, M., Bailly, X., Martinez, P. and Wanninger, A. (2010). Steps towards a centralized nervous system in basal bilaterians: insights from neurogenesis of the acoel *Symsagittifera roscoffensis*. *Dev. Growth Differ.* **52**, 701-713.
- Sikes, J. M. and Bely, A. E. (2010). Making heads from tails: development of a reversed anterior-posterior axis during budding in an acoel. *Dev. Biol.* **338**, 86-97.
- Simionato, E., Ledent, V., Richards, G., Thomas-Chollier, M., Kerner, P., Coornaert, D., Degnan, B. M. and Vervoort, M. (2007). Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics. *BMC Evol. Biol.* **7**, 33.
- Smit, A. F. A., Hubley, R. and Green, P. (1996-2014). *RepeatMasker Open-3.0*. Available at: <http://www.repeatmasker.org/>
- Smith, J. P. S. and Tyler, S. (1985). Fine structure and evolutionary implications of the frontal organ in Turbellaria Acoela. I. *Diopisthoporus gymnopharyngeus* n.sp. *Zool. Scr.* **14**, 91-102.
- Smith, J. P. S., Tyler, S. and Rieger, R. M. (1986). Is the Turbellaria polyphyletic? *Hydrobiologia* **132**, 71-78.
- Srivastava, M., Mazza-Curll, K. L., van Wolfswinkel, J. C. and Reddien, P. W. (2014). Whole-body acoel regeneration is controlled by Wnt and Bmp-Admp signaling. *Curr. Biol.* **24**, 1107-1113.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313.
- Stanke, M., Diekhans, M., Baertsch, R. and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**, 637-644.
- Strotmann, R., Schröck, K., Bösel, I., Stäubert, C., Russ, A. and Schöneberg, T. (2011). Evolution of GPCR: change and continuity. *Mol. Cell. Endocrinol.* **331**, 170-178.
- Tyler, S. (2001). The early worm: origins and relationships of the lower flatworms. In *Interrelationships of the Platyhelminthes (Systematics Association Special Volume Series)* (ed. D. T. J. Littlewood and R. A. Bray), pp. 3-12. London: Taylor & Francis.
- Tyler, S. and Rieger, R. M. (1975). Uniflagellate spermatozoa in Nemertoderma (Turbellaria) and their phylogenetic significance. *Science* **188**, 730-732.
- Tyler, S. and Rieger, R. M. (1977). Ultrastructural evidence for the systematic position of the Nemertodermatida (Turbellaria). *Acta Zool. Fennica* **154**, 193-207.
- von Graff, L. (1891). *Die Organisation der Turbellaria Acoela*. Leipzig: W. Engelmann.
- Westblad, E. (1949). *Xenoturbella bocki* n.g. n.sp. a peculiar, primitive turbellarian type. *Ark. Zool.* **1**, 3-29.
- Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S., Huang, W., He, G., Gu, S., Li, S. et al. (2014). SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* **30**, 1660-1666.

Fig. S1. Molecular phylogenetic tree of bHLHs transcription factors, generated by maximum-likelihood using RAxML with 100 bootstrap replications. This tree is based in the alignment of bHLHs orthologous sequences of *S. roscoffensis* with those of the species *H. sapiens*, *N. vectensis*, *D. melanogaster*, *A. digitifera* and *H. magnipapillata*. The data was provided by F. Gyoja, (Gyoja et al., 2012, Simionato et al., 2007). We are keeping the original orthologous family names used in previous studies. We use the names given by Gyoja et al. (Gyoja et al., 2012) and Simionato et al. (Simionato et al., 2007), which are based on accession number or FlyBase IDs, adding an abbreviation for the species name followed by the, generic, orthologous family names. The sequences domains of two other species, *S. kowalevski* and *C. teleta* were added from the PFAM database, preserving their original identification codes. The monophyletic groups are shown and gene families are grouped by colours. Numbers at the branches are the statistical support obtained by bootstrap; only bootstrap values above 50% are shown, except for a few cases. The bHLHs family members identified in *S. roscoffensis* are indicated by black arrows.

[Download Figure S1](#)

Fig. S2. Phylogenetic analysis of *X.bocki* bHLHs transcription factors. The represented tree is generated by maximum-likelihood using RAxML. This tree is the outcome from a multiple alignment of bHLHs orthologous sequences of *X.bocki* and those of the species *H. sapiens*, *N. vectensis*, *D. melanogaster*, *A. digitifera* and *H. magnipapillata* (Gyoja et al., 2012, Simionato et al., 2007). Basically, the bHLHs sequences from these species conserve the names given by Gyoja et al. (Gyoja et al., 2012) and Simionato et al. (Simionato et al., 2007), which are based on accession number or FlyBase IDs, adding an abbreviation for the species name followed by the generic orthologous family names. Two new species sequences domains were included in our analysis (*S. kowalevski* and *C. teleta*), all downloaded from the PFAM database and keeping their original identification codes. The gene families are clustered and highlighted within colour boxes, with the inside grouping of all monophyletic groups. Numbers at the branches are the times that a node was supported in 100 bootstrap pseudoreplications; only values higher than 50% are shown, except for few cases. Black arrows designate the bHLHs family members of *X.bocki* found.

[Download Figure S2](#)

Fig. S3. Glutamate_tree. Phylogenetic relationships between glutamate GPCR family members of the species *Amphimedon queenslandica*, *Nematostella vectensis*, *Saccoglossus kowalevskii*, *H. sapiens*, *Capitella teleta*, *D. melanogaster* and the acoel *Symsagittifera roscoffensis*. GenBank or ENSEMBL proteins IDs are shown in the sequence name of the 6 first species (sequences from de Mendoza et al., 2014). Trees are generated with RAxML bestree with 100 bootstrap pseudoreplications. Bootstrap values lower than 50% are not shown. Glutamate subfamilies are highlighted using different colours. Acoel sequences are marked with a black arrow.

[Download Figure S3](#)

Fig. S4. Frizzled_tree. Phylogenetic relationships between frizzled GPCR of the species *Amphimedon queenslandica*, *Nematostella vectensis*, *Saccoglossus kowalevskii*, *H. sapiens*, *Capitella teleta*, *D. melanogaster* and the acoel *Symsagittifera roscoffensis*. GenBank or ENSEMBL proteins IDs shown in the sequence name of the 6 first species (sequences from de Mendoza et al., 2014). Trees are generated with RAxML bestree with 100 bootstrap replicates. Bootstrap values lower than 50% are not shown. Frizzled subfamilies are highlighted using different colours. Acoel sequences are marked with a black arrow.

[Download Figure S4](#)

Fig. S5. A/S_tree. Phylogenetic relationships between adhesion and secretin GPCR of the species *Amphimedon queenslandica*, *Nematostella vectensis*, *Saccoglossus kowalevskii*, *H. sapiens*, *Capitella teleta*, *D. melanogaster* and the acoel *Symsagittifera roscoffensis*. GenBank or ENSEMBL proteins IDs are shown in the sequence name of the 6 first species (sequences from de Mendoza et al., 2014). Trees are generated with $N=100$ RAxML bestree with 100 bootstrap replicates. Bootstraps <50 are generally not shown. A/S subfamilies are highlighted using different colours. Acoel sequences are marked with a black arrow.

[Download Figure S5](#)

Fig. S6. Complete multiple alignment for the analysis of the bHLHs orthologous protein sequences of the acoel *Symsagittifera roscoffensis*, performed in MAFFT with all the bHLHs sequences used in this study. The bHLHs sequences of the species *H. sapiens*, *N. vectensis*, *D. melanogaster*, *A. digitifera* and *H. magnipapillata* were provided by F. Gyoja (Gyoja et al., 2012, Simionato et al., 2007). The sequences domains of two other species, *S. kowalevski* and *C. teleta* were added from the PFAM database, preserving their original identification codes. The file in fasta format results from the selection of conserved regions using the Geneious software (<http://www.geneious.com/>).

[Download Figure S6](#)

Fig. S7. Multiple alignment used in the analysis of the bHLHs orthologous protein sequences in the xenoturbellid *X.bocki*, performed in MAFFT with the total group of bHLHs sequences used in this study. The bHLHs sequences of the species *H. sapiens*, *N. vectensis*, *D. melanogaster*, *A. digitifera* and *H. magnipapillata* were provided by F. Gyoja (Gyoja et al., 2012, Simionato et al., 2007). The sequences domains of two other species, *S. kowalevski* and *C. teleta* were added from the PFAM database, preserving their original identification codes. The file is in FASTA format and contains the conserved regions manually selected from the original alignment using the Geneious software.

[Download Figure S7](#)

Fig. S8. Multiple alignment used in the analysis of the Glutamate orthologous proteins in the acoel *S. roscoffensis*, performed in MAFFT with the reference sequences of *A. queenslandica*, *H. Sapiens*, *D. melanogaster*, *S. kowalevski*, *N. vectensis*, and *C. teleta* provided by A. de Mendoza (de Mendoza et al., 2014). The file is in FASTA format and contains the conserved regions manually selected from the original alignment using Geneious software.

[Download Figure S8](#)

Fig. S9. Multiple alignment used in the analysis of the Frizzled orthologous proteins in the acoel *S. roscoffensis*, performed in MAFFT with the reference sequences of *A. queenslandica*, *H. sapiens*, *D. melanogaster*, *S. kowalevski*, *N. vectensis*, and *C. teleta* provided by A. de Mendoza (de Mendoza et al., 2014). The file is in FASTA format and contains the conserved regions manually selected from the original alignment using Geneious software.

[Download Figure S9](#)

Fig. S10. Multiple alignment used in the analysis of the Adhesion and Secretin orthologous proteins in the acoel *S. roscoffensis*, performed in MAFFT with the reference sequences of *A. queenslandica*, *H. sapiens*, *D. melanogaster*, *S. kowalevski*, *N. vectensis*, and *C. teleta* provided by A. de Mendoza (de Mendoza et al., 2014). The file is in FASTA format and contains the conserved regions manually selected from the original alignment using Geneious software.

[Download Figure S10](#)