Charles University in Prague

Faculty of Science

Department of Philosophy and History of Science

PhD Thesis

Evolution of photoreceptors: insights from amphioxus

Pavel Vopálenský

Supervisor: **Zbyněk Kozmik, Ph.D.**
Department of Transcriptional Regulation
Institute of Molecular Genetics
Academy of Sciences of the Czech Republic

Prague 2009

I hereby declare that this thesis has been composed by Pavel Vopálenský, the undersigned, with intention to acquire the degree of Ph.D. at the Charles University in Prague. This thesis represents an original piece of work and has not been presented in any previous application for an academic degree.

Pavel Vopálenský, December 2009

Poděkování

Chtěl bych poděkovat svému školiteli Zbyňku Kozmíkovi za vynikající vedení, hodiny strávené nekonečnými diskuzemi a za jeho podporu a otevřenost vůči novým myšlenkám.

Dále bych chtěl poděkovat všem souputníkům z Oddělení transkripční regulace za vytvoření přátelského a stimulujícího prostředí, veškeré rady a za časté a vynikající moučníky.

Děkuji Prof. Nicholasovi a Lindě Holland (UCSD) za pohostinnost a trpělivost během stáže v jejich laboratoři, Prof. Thurstonu Lacalli (University of Victoria) za neocenitelnou pomoc v interpretaci expresních dat a dále Marii Tosches, Keren Guy a Raju Tomerovi z laboratoře Prof. Detleva Arendta (EMBL) za pomoc s částí projektu zabývající se organismem *Platynereis.*

Všem svým přátelům a bližním děkuji za krásné zážitky a osvěžující chvíle strávené v jejich kruhu.

V neposlední řadě děkuji svým rodičům a celé rodině za veškeré zázemí a bezmeznou podporu během studia.

Acknowledgements

I am very thankful to my supervisor Dr. Zbynek Kozmik for excellent leadership, his time spent in our endless discussions, for his support and encouragement to pursue new ideas.

Next, I thank all labmates from the Department of Transcriptional Regulation for friendly and stimulating athmosphere, all their advise and excellent cookies and cakes.

I would like to acknowledge Prof. Nicholas and Linda Holland (UCSD) for their hospitality and patience during my stage in their lab, Prof. Thurston Lacalli (University of Victoria) for his invaluable help with interpretation of expression data and Maria Tosches, Keren Guy and Raju Tomer from the lab of Prof. Detlev Arendt (EMBL) for their help with *Platynereis* part of this work.

I thank my friends for all the nice experiences and refreshing moments we could spend together.

Last but not least, special thanks to my parents and family for their constant support and encouragement during my studies.

Table of Contents

1 Introduction

The topic of eye evolution was touched as early as at the very beginning of the evolutionary theory itself. In his '*On the Origin of Species*' (Darwin, 1859), chapter VI, devoted to '*Difficulties on Theory*', Charles Darwin expresses his concern for the explanation of the occurence of complex eyes by his theory:

'*To suppose that the eye, with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selection, seems, I freely confess, absurd in the highest possible degree. Yet reason tells me, that if numerous gradations from a perfect and complex eye to one very imperfect and simple, each grade being useful to its possessor, can be shown to exist; if further, the eye does vary ever so slightly, and the variations be inherited, which is certainly the case; and if any variation or modification in the organ be ever useful to an animal under changing conditions of life, then the difficulty of believing that a perfect and complex eye could be formed by natural selection, though insuperable by our imagination, can hardly be considered real. How a nerve comes to be sensitive to light hardly concerns us more than how life itself first originated...*'

Even today, 150 years after the appearence of this fundamental book, when many of the original questions have been solved by the advances in microscopy, physiology, molecular biology and genetics, some aspects of eye evolution still remain unsolved and occasionally raise a breath of controversy. Perhaps the most discussed topic of the last decade has been the homology of eyes across animal phyla with everlasting question about 'monophyletic' or 'polyphyletic' origin of animal eyes (see *eg.* (Meyer-Rochow, 2000, Gehring, 2000) for an example of on-going discussion). A detailed view on the question of 'mono-' or 'polyphyletic' origin of eye reveals that the question is actually not properly stated and the answer strogly depends on the level of comparison, *ie.* how the eye has been defined. Obviously, at the level of organs, the compound eyes of insect, camera type eyes of molluscs or vertebrates have a clear polyphyletic origin, however, the answer is not so clear when we focus on lower levels, such as the cells or genes. Even the simplest **'**proto-eye**'** composed of a single photoreceptor cell and a shading pigment (Figure 1) present either internally or in adjacent pigment cell, (Land and Nilsson, 2002), requires proper functioning of dozens of genes involved in the phototransduction cascade, signal

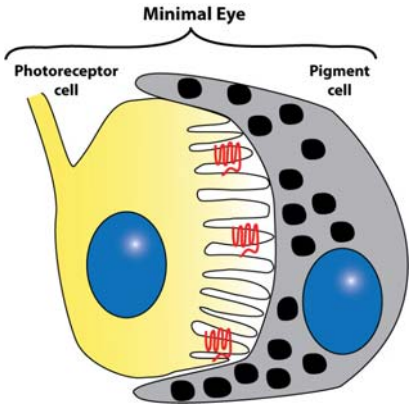


Figure 1: A schematic diagram of a 'minimal eye'
The 'minimal eye' employs a single photoreceptor cell containing photosensitive molecules and a pigment cell which restricts the incoming light to a certain direction. This simple design is exemplified by dorsal ocelli of amphioxus or paired larval eyes of *Platynereis*.

---

transmission, and pigmentation. Moreover, to properly orchestrate the development and maintenance of such an intricate structure, a tight interplay of several transcription factors is needed as well. Apparently, as the above mentioned genes participating in eye formation are not related and have different evolutionary histories, the evolution of an eye then becomes a question of evolutionary history of these separate components and their continuous or repeated integration to an organ called 'eye'.

1.1 Photoreceptors

Despite a remarkable variation in size and complexity, the common indispensable basis of all animal eyes is the photoreceptor cell containing photosensitive molecules, which are connected to a downstream phototransduction cascade. The photoreceptor cells are classified according to the morphology of membrane protrusions bearing visual pigments as 'rhabdomeric', which form microvilli, and 'ciliary', where the membrane surface is increased by folding the membrane of the cilium (Eakin, 1979, Arendt, 2003). Primary observations revealed the rhabdomeric photoreceptors being predominantly present in the eyes of invertebrates, whereas the vertebrate eyes employ the ciliary type, however, several exceptions from this rule do exist (Land and Nilsson, 2002, Arendt and Wittbrodt, 2001). Both photoreceptor cell types have always co-existed in bilaterians, as suggested by both types found in amphioxus (Lacalli, 2004) and confirmed by recent morphological and molecular studies (Arendt et al., 2004, Velarde et al., 2005). Still, it is

not clear why the two types of photoreceptors were employed in the eyes of invertebrates vs vertebrates in a mutual-exclusive way.

1.1.1 Opsins – variation on ancestral theme

The first step of photoreception is mediated by light-sensitive transmembrane proteins containing retinal chromophore, generally termed rhodopsins. They have been found in most groups of organisms including archeal prokaryotes (Blanck and Oesterhelt, 1987), unicellular eukaryotes (Nagel et al., 2002), fungi (Bieszke et al., 1999) and metazoa. The function of rhodopsins in these organisms vary from photon-driven ionic pumps in prokaryotes (Blanck and Oesterhelt, 1987), sensory molecules in fungi, to a light-gated ion channel in the eyespot of green algae (Spudich et al., 2000, Nagel et al., 2002). Metazoan opsins are seven-transmembrane proteins belonging to the superfamily of G-protein coupled receptors (GPCRs) and often are coupled to a Gα protein to mediate a phototransduction cascade. Despite these distinct functions of rhodopsins, several common features - such as a transmembrane structure, conserved retinal group covalently bound to a lysine residue and similarities in exon-intron structure suggested the possibility of common evolutionary origin of all opsins. However, a recent bioinformatic study (Larusso ND, 2008) provided strong evidence that at least prokaryotic and metazoan opsins are not homologous, therefore revealing their common features as a remarkable example of convergent evolution.

With increasing number of sequenced genomes available, opsin gene repertoire has been described in several animal species (Plachetzki et al., 2007, Raible et al., 2006, Holland et al., 2008). This information then enables the understanding of metazoan opsin evolution and origin (Figure 2), which is inseparably connected with the origin of eyes itself. The metazoan opsins seem to originate early in metazoan evolution from a single GPCR gene by acquisition of light-sensitivity. Since no opsin genes have been found in available genomic data for the choanoflagellate *Monosiga* and the poriferan *Amphimedon* (Suga et al., 2008, Plachetzki et al., 2007), it is probable that this event pre-dated the cnidarian-bilaterian ancestor. Before the split of cnidarians and bilaterians, the newly formed ur-opsin gene underwent duplication producing two ancestral classes of opsins: ciliary opsin class (c-opsin) conserved in cnidarians and bilaterians, and a second opsin gene eventually giving rise to the rest of bilaterian opsins (see below) and persisting in cnidarians as ‘cnidopsin’ class (Figure 2) (Plachetzki et al., 2007).

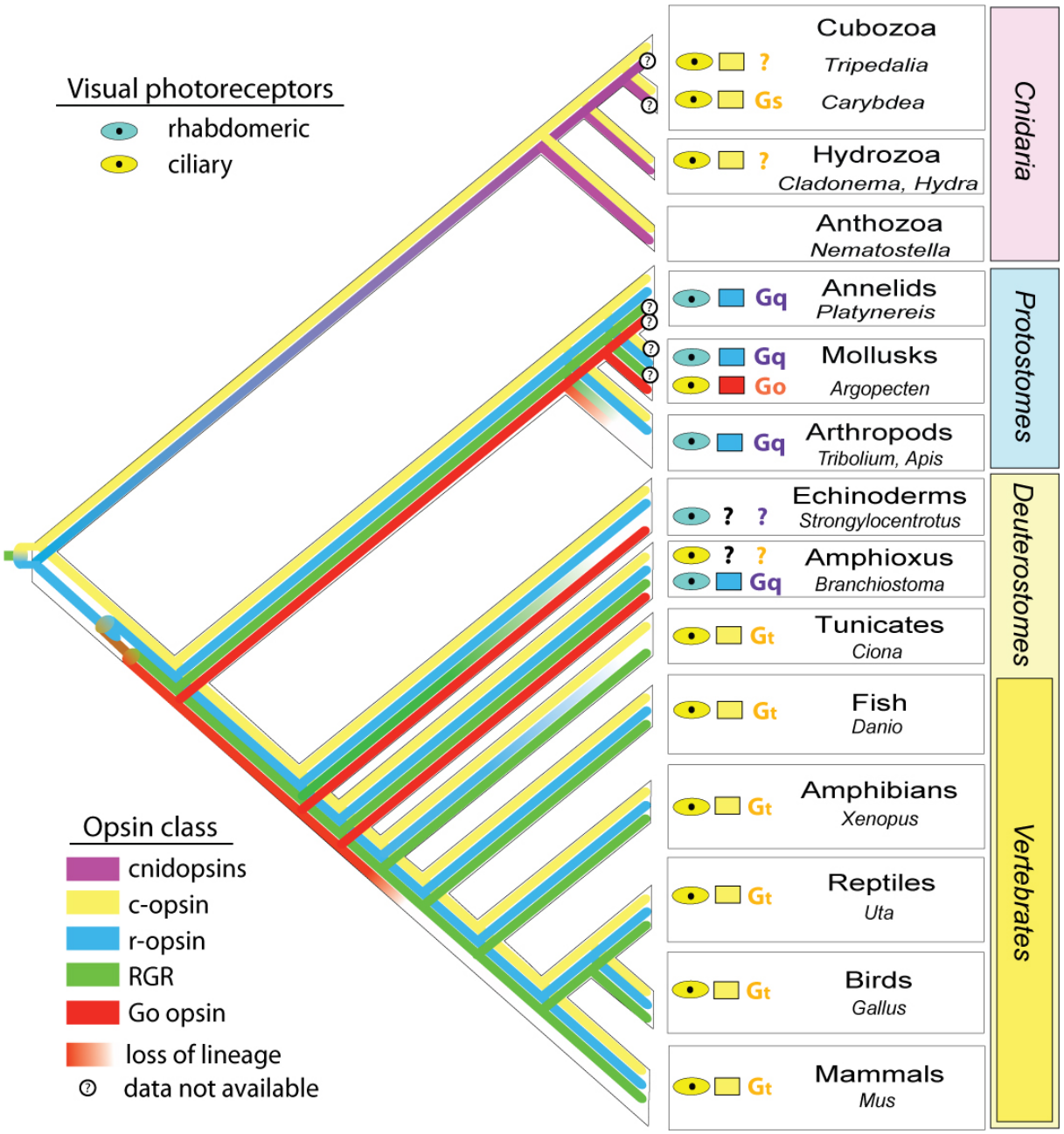


Figure 2: A schematic diagram of opsin distribution among eyes in different animal phyla
Particular opsins subfamilies are distinguished by different coloring. The lines leading to different phyla depict the hypothetical evolutionary fate of given opsin lineage based on available genomic and other data. The question-mark denotes such animals, where the presence of a given opsin lineage has not been confirmed yet. The color of the eye-like pictogram corresponds to the type of photoreceptor cell employed in the eye. The class of opsin employed in the eye is represented by the color of the rectangle next to the eye-like pictogram. If known, the Gα subunit interacting with the opsin is shown. Note that a small subset of vertebrate retinal ganglion cells expresses melanopsin coupled to a Gq signaling cascade (Panda et al., 2005). Although these cells fulfill the definition of a minimal eye, they are not the major photoreceptors of the eye and are not considered in this figure.

In cnidarians, the c-opsins are expressed in the ciliary photoreceptors of adult eyes of hydrozoan *Cladonema*, cubozoans *Tripedalia* and *Carybdea* (Suga et al., 2008, Kozmik et al., 2008, Koyanagi et al., 2008).

The bilaterian opsin repertoire comprises the c-opsin class and the second ancestral class which has diversified into several subclasses termed as rhabdomeric opsins (r-opsins), Go-coupled opsins, neuropsins and RGR (Terakita, 2005). From these, only three groups (namely c-, r- and Go-coupled opsins) seem to have been recruited for visual purposes, whereas the function of other subclasses is probably supportive, as for example photoisomerases (RGR-opsin) involved in retinal visual cycle (Radu et al., 2008). With one exception of Go-coupled opsin mediating the phototransduction in the ciliary part of the retina in scallop *Patinopecten yessoensis* (Kojima et al., 1997), the rhabdomeric and ciliary photoreceptors of bilaterians consistently employ r- and c-opsins, respectively.

1.1.2 Phototransduction – promiscuity in Gα coupling in ciliary photoreceptors

The different morphology of rhabdomeric and ciliary photoreceptors is further reflected on the level of phototransduction cascades operating in these cells, although some common elements, such as opsins or arrestins, participate in both. The rhabdomeric phototransduction cascade, which is mediated by $G\alpha_q$ and phospholipase C (Suzuki et al., 1995), seems to be evolutionarily conserved from protostomes to vertebrate retinal ganglion cells (Contin et al., 2006, Panda et al., 2005, Graham et al., 2008, Koyanagi et al., 2005). In contrast, the ciliary photoreceptors may employ both Go-opsin (Kojima et al., 1997) and c-opsins, whose Gα protein partners may be rather variable: In vertebrate rods and cones, the c-opsin couples via transducin $G\alpha_t$ to downstream hyperpolarizing phototransduction cascade. Since the $G\alpha_t$ subfamily originated together with other rod and cone specific phototransduction genes during vertebrate-specific whole-genome duplications (Nordstrom et al., 2004, Milligan and Kostenis, 2006), the $G\alpha_t$ subunits are not present in invertebrates. A question arises - what are then the Gα-proteins involved in ciliary phototransduction cascades in invertebrates and cnidarians? Apparently the members of $G\alpha_{i/o}$ protein subfamily, from which $G\alpha_t$ proteins evolved, could be plausible candidates for this function. Consistently with this assumption, a $G\alpha_{i1}$ protein subunit is expressed in the ciliary photoreceptor cells of *Ciona intestinalis* (Yoshida et al., 2002). In a

reptile *Uta stansburiana*, a vertebrate c-opsin (parietopsin) expressed in the parietal eye retina, signals via a $G\alpha_o$-protein (Su et al., 2006), and in the same cell, another vertebrate visual/nonvisual opsin – pinopsin - couples with gustducin – a third vertebrate paralogue of transducin not used in rods and cones. In contrast to its reptile counterpart, chicken pinopsin has been shown to interact with $G\alpha_{11}$ subunit (Kasahara et al., 2002), which is closely related to $G\alpha_q$ protein. An exciting surprise came from a recent study (Koyanagi et al., 2008), which revealed that phototransduction in cubomedusan *Carybdea rastonii* is mediated by $G\alpha_s$ cascade.

Taken together, these findings indicate that the coupling specificity of ciliary opsins could be rather promiscuous in comparison to the rhabdomeric opsins retaining more strictly the $G\alpha_q$ specificity. Since the origin of G-proteins pre-dated the origin of opsins, the opsin-Gα protein interaction evolved by co-option and subsequent co-evolution (Plachetzki and Oakley, 2007). Multiple co-option events during c-opsin evolution might explain their 'promiscuity' and reconcile an apparent discrepancy between the data pointing to $G_{i/o}$ phototransduction cascade in *Tripedalia* (Kozmik et al., 2008) and Gs-mediated transduction in *Carybdea* (Koyanagi et al., 2008). Although *Carybdea* and *Tripedalia* are closely related, the opsin sequences identified in the studies are rather diversified and suggest that two different phototransduction cascades in structurally the same eyes might be possible. Moreover, all the opsins found to be expressed in the eye of hydrozoan *Cladonema radiatum* and assigned to the ciliary class (Suga et al., 2008) show even more sequence diversification. Together with a low bootstrap support in phylogenetic trees and discrepancies in total number of *Hydra* and *Nematostella* opsins identified in two independent studies (Plachetzki et al., 2007, Suga et al., 2008) further analyses are required to fully resolve the relationships among cnidarian opsins, address their Gα coupling specificity and role in phototransduction.

1.2 Redeployment of a selected set of transcription factors for animal eye development

Besides the phototransduction genes, a subset of which has been described above, other components such as pigmentation genes and lens-specific genes participate in the formation of animal eyes (see appendix 8.6.3 (Vopalensky and Kozmik, 2009), for a short review). To properly orchestrate the assembly of all these components into a functional

organ, a tight regulation by a dedicated set of transcription factors is required. Given the enormous diversity of animal eyes it came as a surprise that certain transcription factors are redeployed for visual system development far more often than others. In addition to governing eye morphogenesis, some of these transcription factors are directly involved in the regulation of differentiation genes encoding structural eye components.

1.2.1 The Pax6 paradigm

One of the most remarkable examples of genes with an evolutionary conserved role in organogenesis is a paired-homeobox gene *Pax6*. Mutations in *Pax6* gene disrupt eye development in both mammals (Hill et al., 1991) and insects (Quiring et al., 1994). The ability to induce ectopic eyes through *Pax6* misexpression has furthermore been demonstrated in *Drosophila* and vertebrates (Halder et al., 1995, Chow et al., 1999). The key role of the *Pax6* gene for eye morphogenesis in such diverse species led to the proposal of *Pax6* being a 'master control gene' in animal visual system development (Gehring and Ikeo, 1999). The hypothesis proposes the origin of a regulatory cascade starting with a differentiation gene regulated by a transcription factor and subsequent intercalation of different transcription factors between these two genes. The term 'master control gene' then implies that the Pax6 transcription factor is located at the top of a gene cascade and initiates eye development in almost any tissue where it is ectopically expressed. However, neither seems to be the case. There are few notable examples, however, known so far among bilaterians of eyes developing in the absence of *Pax6*. *Pax6* is apparently not expressed in developing *Limulus* eyes (Blackburn et al., 2008), developing *Platynereis* adult eyes (Arendt et al., 2002), Hesse eye cups of amphioxus (Glardon et al., 1998) and its function is not required for the eye regeneration in planarians (Pineda et al., 2002) as well as the Bolwig organ in *Drosophila* (Suzuki and Satoh, 2000). In addition, the absence of Pax6 in ciliary photoreceptors of vertebrates has puzzled scientist for a long time, however, a recent study has shown the expression of Pax4 in these photoreceptors (Rath et al., 2009). Pax4 is a paralogue of Pax6 and is therefore orthologous to invertebrate *Pax6* genes which should be termed Pax4/6, to be formally correct.

The 'bipartite model' (Kozmik, 2005), offers an alternative explanation of frequent deployment of *Pax* genes in eye development in phylogenetically diverse species. The model proposes that the two independent DNA binding domains (paired domain and the homeodomain) within a single Pax transcription factor have been co-opted for two

essential features of the prototypical eye, production of a dark shielding pigment and expression of an opsin gene. This scenario thus elegantly explains what makes *Pax* genes ideal candidates for the key regulators in the prototypical eye, however, to explain its frequent occurence also in the developmental processes, we have to consider the new findings in the evolution of gene regulatory subcircuits.

1.2.2 Differentiation regulatory subcircuits

With the growing knowledge of gene regulatory networks architecture and evolution (Davidson, 2006) and the emerging concept of sister cell types (Arendt, 2003, Arendt, 2008), it turns out that the developmental processes may be subject of evolutionary plasticity whereas '*...the differentiating cell types seem to be islands in a sea of developmental change*' (Arendt, 2008).

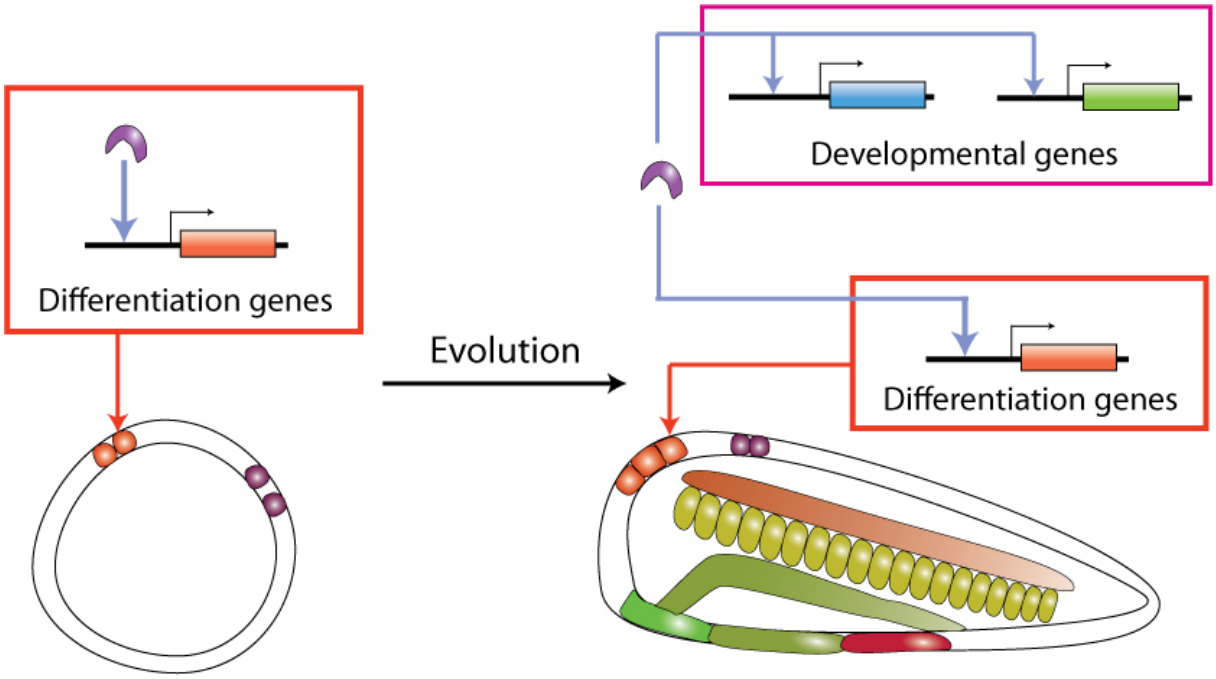


Figure 3: A simplified scheme of the evolution of gene regulatory circuits
An animal with a simple body plan possessing multifunctional differentiated cells, in which a given transcription factor directly regulates a given differentiation gene. Through the course of evolution, body plans became more elaborated due to more complex developmental processes. The regulatory relationship between the differentiation gene and the transcription factor is retained, despite the recruitment of the transcritption factor into 'higher' levels of the regulatory networks (developmental roles).

The ‘master control gene’ model (Gehring and Ikeo, 1999) and the ‘bipartite’ model (Kozmik, 2005) agree in the first step of eye evolution – an (probably stochastic) event of an opsin gene becoming regulated by a *Pax* gene. Next, being expressed in the photoreceptor cell, the regulator is ‘at hand‘ and may be recruited for a developmental regulatory task without losing its ancestral function in direct regulation of a given differentiation gene (Figure 3) by a combination of several events: (i) the transcription factor may become a novel regulator of various genes through a mutation in regulatory regions of these genes, (ii) by acquiring new cis-regulatory elements in its own promoter, the transcription factor is recruited into ‘higher‘ levels of the network, and, being a regulator of other genes (due to the events described in (i)), the network topology may be substantially changed, (iii) a mutation in the coding sequence may result in altering the protein-protein interactions, leading to the formation of new transcriptional complexes and modules, (iv) the duplication of the transcription factor may be followed by subfunctionalization of their roles. Sometimes, one of the paralogue may be more involved in the developmental function, while the other retains the role in differentiation processes. This scenario may be well illustrated by the differential expression of vertebrate *Pax6, Rx* and *Otx1/2* (see below) genes during vertebrate eye development, and their paralogues *Pax4, RaxL* and *Crx* being expreessed later and playing role in the differentiation process.
All the processes described above can lead to incorporation of a given transcription factor into developmental regulatory networks without affecting its function in regulating a key differentiation gene. In addition, the scenario does not require the transcription factor being at the top of the regulatory cascade; in contrast, it suggests that the transciption factor may occur at different positions within the network topology.

Three other transcription factors (Otx, Rx and Prox) which are repeatedly found to participate in animal eye development and photoreceptor differentiation (suggesting their potential ancestral role) are descibed in the following chapters.

1.2.3 Orthodenticle-related homeobox *(Otx)*

The first member of the *Otx* gene family – *orthodenticle (Otd*) – has been isolated from *Drosophila* and shown to be necessary for development of photoreceptors in the compound eye, Bolwig organ and the ocelli (Finkelstein et al., 1990, Vandendries et al., 1996, Royet and Finkelstein, 1995). *Otd* also participates in terminal photoreceptor differentiation. It has been shown to directly regulate opsins (Tahayato et al., 2003) and

influence the expression of genes involved in rhabdomeric phototransduction cascade (Ranade et al., 2008). In other invertebrates, the expression of *Otx* genes in photoreceptors has been reported in planarians (Umesono et al., 1999), putative eye-field precursor of the annelid *Hydroides elegans* (Arenas-Mena and Wong, 2007), ciliary and rhabdomeric photoreceptor cells of *Platynereis dumerilii*, (Detlev Arendt – personal communication), sensory pigment cells of ascidians (Wada et al., 1996) and the frontal eye region of amphioxus (Williams and Holland, 1996). Reciprocal rescue experiments with *Drosophila* and mammalian *Otx* orthologues demonstrated that at least part of ancestral genetic and biochemical interactions is still conserved between vertebrates and invertebrates (Acampora et al., 1998, Nagao et al., 1998).

Multiple vertebrate orthologoues of *otd* termed *Otx*1, *Otx*2 and *Crx*/*Otx*5 (Germot et al., 2001, Plouhinec et al., 2003) probably arose during whole genome duplication, since a single *Otx* gene is present in the genome of *C. intestinalis* (Wada et al., 2003) and amphioxus (Williams and Holland, 1998). Besides the role of *Otx* genes in early vertebrate development of anterior neural structures and the brain (Acampora et al., 2005, Simeone et al., 2002), these genes are necessary for proper development of the pineal gland and the eye (Martinez-Morales et al., 2001, Nishida et al., 2003, Plouhinec et al., 2005). Later in development, *Otx* genes play a crucial role in the terminal differentiation of photoreceptors and their maintenance during postnatal development (Nishida et al., 2003, Koike et al., 2007). The expression of *Otx* genes has been also detected in immature retinal ganglion cells (Rath et al., 2007, Martinez-Morales et al., 2001, Bovolenta et al., 1997) - putative descendants of the rhabdomeric photoreceptor line in vertebrates (Arendt, 2003).

The role of vertebrate *Otx* genes in the regulation of eye specific genes has been extensively studied and lead to the discovery of many direct target genes. *Crx,* strongly expressed in differentiated photoreceptor cells, directly regulates the phototransduction genes – rhodopsin, β–PDE, arrestin and guanylate cyclase via binding the PCE element in the promoters (Chen et al., 1997, Qian et al., 2005, Furukawa et al., 1997) (for review, see (Hennig et al., 2008). The expression of ciliary-phototransduction cascade genes in the vertebrate pineal gland is mediated by the action of *Otx* genes as well (Appelbaum and Gothilf, 2006, Takechi et al., 2008). Besides the direct regulation of photoreceptor specific genes, *Otx* genes are involved in the regulation of pigmentation. In ascidians, the *Tyrp* gene is a direct target of *Otx* (Wada et al., 2002). The vertebrate homologue *Otx*2 has been shown to bind to the promoters of *Mitf*, *tyrosinase* and *Tyrp1* (Martinez-Morales et al., 2003) as well as *Tyrp2* in the retinal pigmented epithelium (Takeda et al., 2003).

1.2.4 Retinal homeobox *(Rx)*

Rx transcription factors are characterized by a Q-50 homeodomain, OAR domain in the C-terminal region, and the octapeptide motif missing in vertebrate Qrx/RaxL paralogues (see below) (Mathers et al., 1997, Wu et al., 2009).

During vertebrate development, *Rx* genes are expressed in the anterior forebrain, retinal primordia and pineal gland (reviewed in (Bailey et al., 2004)). The over-expression of *Rx* in *Xenopus* leads to ectopic formation of retinal tissue (Mathers et al., 1997). In zebrafish *chokh* mutants, a non-sense mutation in *Rx3* paralogue leads to the loss of eyes (Loosli et al., 2003) and *Rx* knock-out mice lack the eye and the anterior brain structures (Mathers et al., 1997).

The second vertebrate paralogue of Rx is termed Qrx in mammals (Wang et al., 2004) or RaxL (Rx-L) in birds (Chen and Cepko, 2002) and amphibians (Pan et al., 2006, Wu et al., 2009). Intriguingly, the Qrx/RaxL gene is missing in rodent genomes, despite of a good overall synteny in the locus (Wang et al., 2004). Due to the lack of the octapeptide motif, which has been shown to interact with *groucho* co-repressors (Eberhard et al., 2000), Qrx/RaxL proteins have positive transactivation properties (Wang et al., 2004, Wu et al., 2009). *Qrx/RaxL* genes are generally expressed at later developmental stages than 'canonical' vertebrate *Rx* paralogues and seem to control the photoreceptor cell differentiation and maintainance (Wu et al., 2009).

Similar to vertebrates, the ascidian homologue of *Rx* is expressed in the anterior brain and the knock-down resulted in the loss of photoreceptor cells (D'Aniello et al., 2006). In contrast, *Rx* is not expressed in planarian eyes (Salo et al., 2002) and genetic studies in *Drosophila* have shown a clear dispensability of *Rx* for compound eye development (Davis et al., 2003). The explanation of this result came from the emerging concept of sister cell types (Arendt, 2003) and the fundamental discovery of *Rx* expression in the ciliary photoreceptors of *Platynereis* brain (Arendt et al., 2004). These findings led to the identification of *Rx* as ciliary photoreceptor-specific marker.

1.2.5 Prospero-related homeobox *(Prox)*

Prospero-related homeobox (*Prox*) is an atypical homeodomain protein, which, in addition to the homeodomain, contains an evolutionary highly conserved 'prospero domain' (Ryter et al., 2002) which affects the DNA-binding specificity of the homeodomain (Yousef and Matthews, 2005). Homeo- and prospero domain form a single

structural unit, whose conformational changes play a role in the regulation of nuclear export/import of Prox protein (Bi et al., 2003). Although Prox transcription factor has been studied for more than a decade, its DNA-binding and transactivation properties remain elusive. The weak Prox/DNA interaction is reflected by different DNA-binding motifs identified in independent studies (Cook et al., 2003, Lengler et al., 2001, Cui et al., 2004, Hassan et al., 1997, Yousef and Matthews, 2005). It appears that the DNA-binding specificity as well as the transactivation properties are probably refined by the interaction with other transcription factors, and the resulting protein complex can also alter the total transactivation output (Hassan et al., 1997, Yousef and Matthews, 2005).

In *Drosophila*, *prospero* plays a crucial role in CNS and eye development (Doe et al., 1991, Matsuzaki et al., 1992). In addition, it has been shown to bind to the promoters of rhodopsin Rh5 and Rh6 genes in R7 type photoreceptors to mediate transcriptional repression of these genes (Cook et al., 2003).

Two paralogues of *Prox* have been identified in vertebrates (Oliver et al., 1993, Nishijima and Ohtoshi, 2006). The expression patterns of *Prox1* in the mouse, chicken and zebrafish are rather similar and include the pancreatic tissue, the lens, and inner nuclear layer of the retina (Oliver et al., 1993, Tomarev et al., 1996, Pistocchi et al., 2008b). Besides the retinal horizontal and amacrine cells, *Prox1* is expressed in a small subset of retinal ganglion cells in mouse (Dyer et al., 2003), nevertheless, the expression of *Prox1* in ganglion cells has not been observed in the chicken (Tomarev et al., 1996). Similar to mouse, during newt eye regeneration, *Prox1* is expressed in amacrine, horizontal, bipolar and gagnlion cells, but not in photoreceptor cell layer (Markitantova et al., 2003). The expression of *Prox2* in zebrafish has been observed in the lens and cranial ganglia, however, *Prox2* was never expressed in the neural retina (Pistocchi et al., 2008a). In mouse, *Prox2* transcript has been shown to be present in adult testes and in the developing nervous system of E9.5 embryos. The function of *Prox2* in mouse remains unknown, since no apparent phenotype has been observed in *Prox2* mutant mice (Nishijima and Ohtoshi, 2006).

Based on molecular fingerprint comparison, the horizontal, amacrine and retinal ganglion cells have been suggested to be vertebrate descendents of rhabdomeric photoreceptor cells (Arendt, 2003). *Prox* is one of the proposed evolutionarily conserved markers of these cells, however, this assumption is based on a limited set of animals and needs further support from non-vertebrate deuterostomes and lophotorchozoans. Interestingly, no *Prox* homologue has been found in any cnidarian nor sponge genome

sequenced so far (Ryan et al., 2006, Larroux et al., 2008), suggesting the origin of this TF in early bilaterians.

1.3 Model organisms for studies of eye evolution

To reconstruct evolutionary history using molecular data, a broad phylogenetic sampling is desirable. If a feature is found in animals positioned at different branches of the phylogeny, then, according to the parsimony principle, this feature is likely to be ancestral. In the field of eye evo-devo biology, majority of molecular data comes from vertebrates and *Drosophila* as a representative an invertebrate species. However, with respect to the eye, vertebrates form a rather uniform group within the deuterostomes and in addition, the whole-genome duplication events which took place at the very begining of vertebrate evolution (Ohno, 1970), produced multiple paralogues of many genes (see (Nordstrom et al., 2004) for an example of phototransduction genes), making the generation and interpretation of the data rather demanding. *Drosophila*, despite its invaluable importance in many fundamental discoveries, is still a rather derived species within the ecdysozoa, with an atypical development including the head structures (Klingler, 2004). Fortunately, new model organisms are being established in ecdysozoa (eg. *Tribolium* (Wang et al., 2007)) and, more importantly, also within the lophotorchozoa (Tessmar-Raible and Arendt, 2003), a group which has been until recently rather neglected in the molecular studies. Non-vertebrate deuterostomes - sea urchins, cephalochordates and tunicates - represent another group of animals, the study of which may substantially contribute to our insight into eye evolution. The identification, cloning and expression of eye-related genes in the cephalochordate amphioxus form the main body of the experimental part of this thesis.

1.3.1 Amphioxus as a model organism

Since the first description by Pallas in 1774, who originally classified it as a snail (*Limax lanceolatus*) (Pallas, 1774), the cephalochordate amphioxus has become one of the key animal models used in evolutionary studies. Due to its vertebrate-like appearance, cephalochordates have been for long time considered as the closest invertebrate relatives to vertebrates. Recently, this phylogenetic position has been proven false (Delsuc et al., 2006) – leaving cephalochordates as sister group of '*Olfactores*' – a newly named clade formed

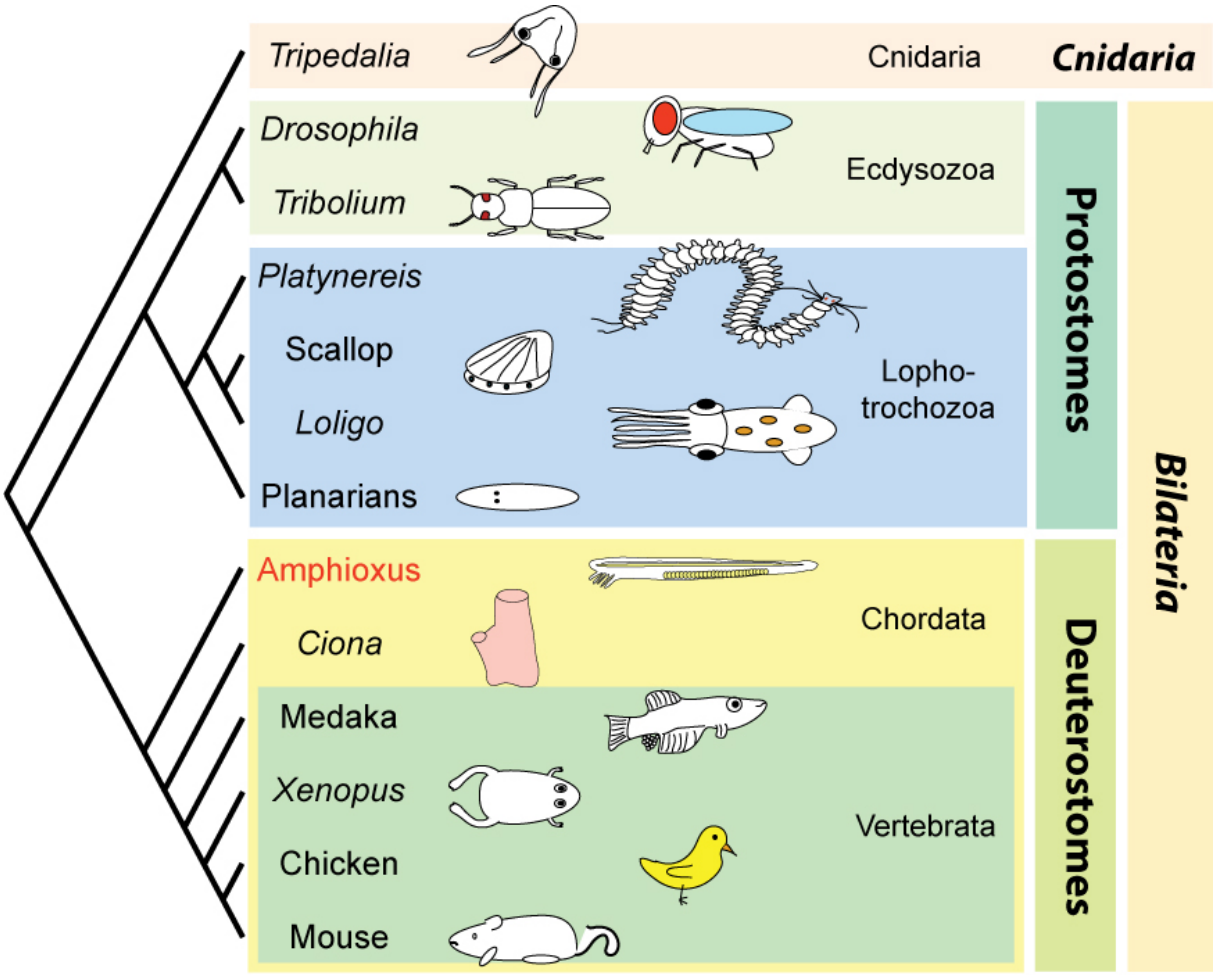


Figure 4: **Phylogenetic position of amphioxus (*Cephalochordata*) within eumetazoa and its relations to other model organisms**

---

by tunicates and vertebrates (Figure 4). Regardless of this change, amphioxus remains a crucial model organism for understanding chordate evolution and vertebrate origin, since, compared to common tunicate models *eg. Ciona sp*., retains vertebrate-like morphology to the adulthood and possesses a slowly evolving underived genome (Putnam et al., 2008).

Despite its long tradition in evolutionary and developmental studies (the oldest articles concerning amphioxus photoreceptors date back to 1897), amphioxus as a model organism still suffers from several disadvantages. First of all, although the efforts to establish amphioxus laboratory cultures have been pursued with partial success (Garcia-Fernandez et al., 2008, Zhang et al., 2007), their widespread use is still underway and the main source of biological material comes from seasonal in-field collections. Second, the unavailibility of fresh 'on-demand' eggs in laboratory conditions is the main limitation for the use of advanced methodology such as transgenesis and other reverse genetic tools.

Although protocols for generating transgenic amphioxus and morpholino-based gene knockdown were developed (Holland and Yu, 2004, Yu et al., 2004) their aplication is limited for only a few nights during the summer spawning season. Given these obstacles, the main body of data coming from amphioxus model is based on expression patterns during embryonic and early larval development.

1.3.2 Amphioxus photoreceptive structures

Amphioxus possesses four putative photoreceptive organs of both ciliary and rhabdomeric type (Figure 5), extensively reviewed in (Lacalli, 2004). Hesse (or dorsal) ocelli scattered along the neural tube are formed by rhabdomeric-type photoreceptor cell and a pigment cell, while the pigment spot of the frontal eye is associated with ciliary photoreceptors. Unpigmented photoreceptors include rhabdomeric Joseph cells (Eakin and Westfall, 1962) and cells of the lamellar body, which were proposed to be ciliary (Ruiz and Anadon, 1991).

1.3.2.1 Shielding pigments

The first documented observations of amphioxus photoreceptive structures commenced in 19th centrury by description of pigments associated with photoreceptor cells. The conspicuous black dots at the anterior tip and along the neural tube drew the attention of several German anatomists who noticed that the dark pigments of the frontal eye and of dorsal ocelli differ in their chemical composition (Krause, 1888, Hesse, 1898, Franz, 1927). Moreover, the pigments slightly differ in color (Ruppert, 1997, Franz, 1927) and the pigment of the frontal eye fluoresces, which is not observed in the case of the Hesse-cup pigment (Terio, 1964, Fritzsch, 1996). A detailed chemical analysis excluded the possibility of the Hesse eye-cup pigment being melanin, ommochrome or porphyrine (Tenbaum, 1955). However, the first larval organ of Hesse formed by one pigment cell and two adjacent photoreceptor cells (Conklin, 1932, Lacalli et al., 1994) has been shown to express *Mitf*, *tyrosinase, Tyrp-a* and *Tyrp-b* (Yu et al., 2008) - a set of genes typical for melanin synthetic pathway.

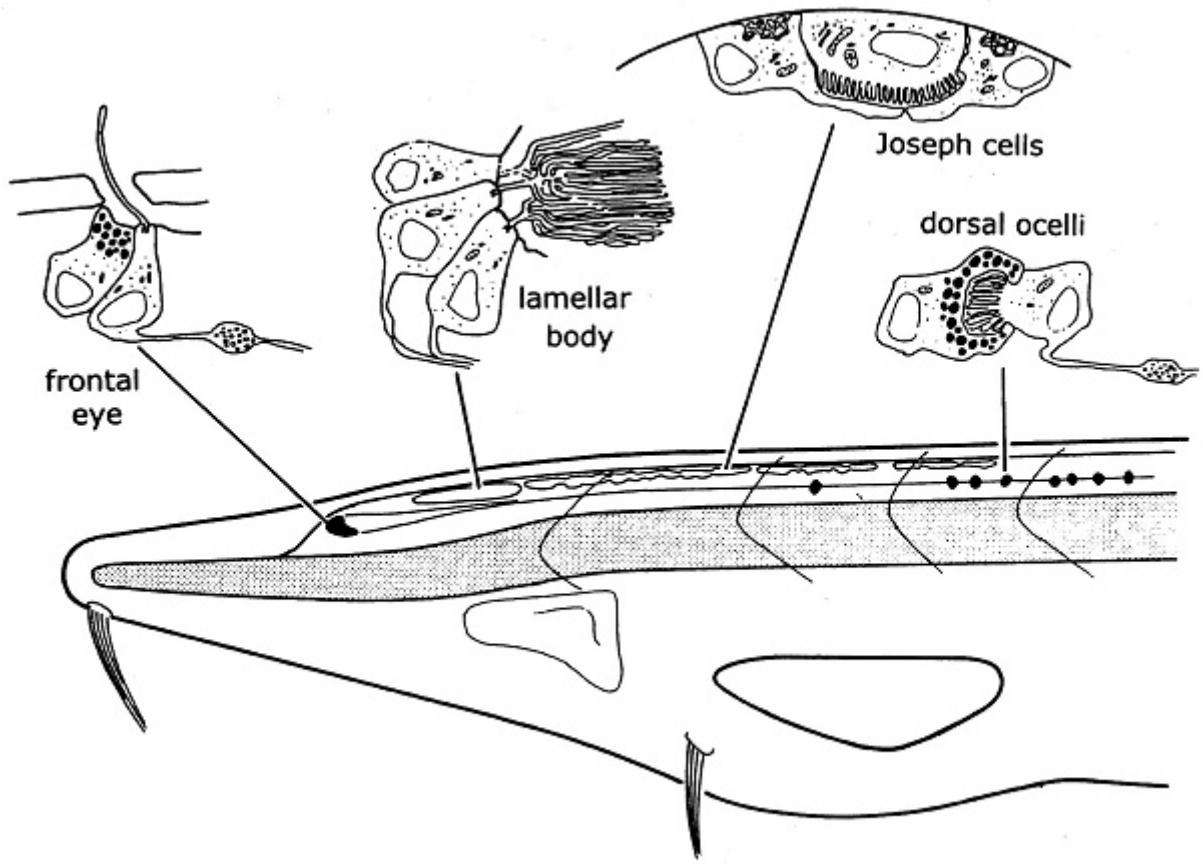


Figure 5: A schematic view of amphioxus photoreceptive structures in a 12.5d larva (adapted from (Lacalli, 2004))
Frontal eye and the lamellar body epmloy ciliary photoreceptors, whereas Joseph cells and dorsal ocelli (Hesse organs) employ rhabdomeric photoreceptors.

1.3.2.2 Photoreceptors

The frontal eye consists of several pigment cells and four rows of neurons (Figure 6) (Lacalli, 1996, Lacalli, 2004). The first two rows contain 6 and 10 receptor cells bearing cilia projecting out from the neuroporus and due to their close association with the pigment Row1 and Row2 cells were suggested to be photoreceptors. The frontal eye differentiates early in the larva and has been shown to be a functional photoreceptive structure in four days-old larvae. In a directional light, the larvae hover in the water by ciliary beating and orient so that the cilia of the frontal eye are maximally shaded by the pigment cup (Stokes, 1997). Although the frontal eye structure persist to the adulthood, whether it has any function is not known, and behavioral experiments indicate that it is not sensitive to light (Crozier, 1916).

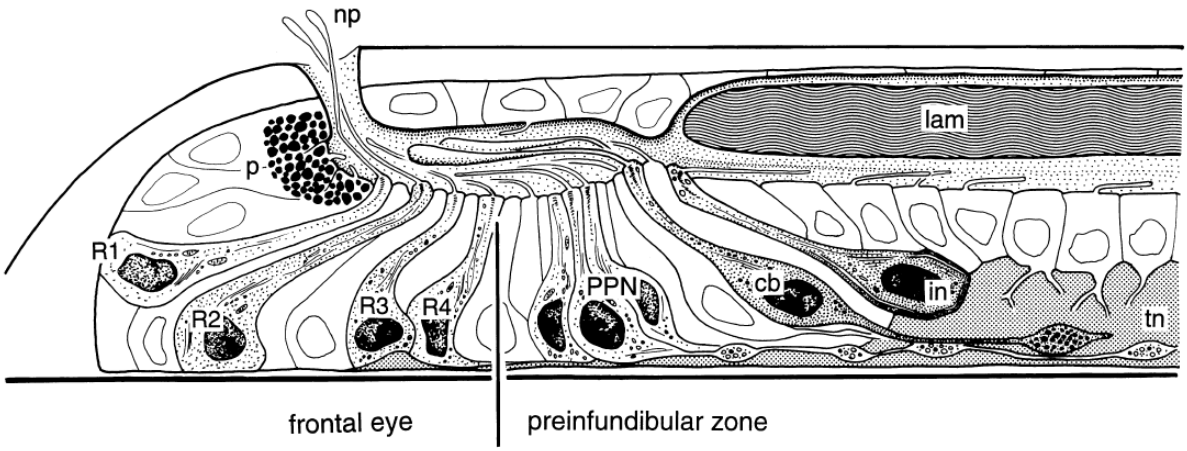


Figure 6: Organization of the cerebral vesicle of a 12.5 day amphioxus larva (adapted from (Lacalli and Kelly, 2000))
The frontal eye is composed from several pigment cells (p) and four transverse rows of receptor cells (R1-2) and neurons (R3-4). *lam* - lamellar body; *PPN* - preinfundibular projection neurons; *in* - infundibular cells; *np* – neuroporus; *cb* - ciliary bulb cells, *tn* - neuropile.

The expression of opsins or any other phototransduction cascade component in the frontal eye has not yet been adressed. Transcription factors expressed in 3-5d larva, with the photoreceptors already differentiated, include *Pax4/6* (Glardon et al., 1998) in the photoreceptor zone, and *Six3/6, Dach and Pax2/5/8* in the pigment cell zone (Kozmik et al., 2007). The expression of *Otx* in the frontal eye region has been mentioned in (Williams and Holland, 1996) and has been re-examined in this study in more detail with respect to the photoreceptor zone.

The lamellar body is a non-pigmented mass of lamellae evaginating from cilia of symmetrical groups of cells positioned dorso-laterally. These putative ciliary photoreceptors resemble the cells of lamprey pineal and the lammellar organ is therefore generally accepted as the homologue of vertebrate pineal (Ruiz and Anadon, 1991, Lacalli, 2004). The lammellar organ is well developed in larvae, however, it becomes fragmented after metamorphosis and its role in adults is not known. The cells of the lamellar organ have been shown to express *Pax4/6* (Glardon et al., 1998)

The Joseph cells are positioned in the most dorsal part of the anterior neural tube behind the lamellar body and span several anterior somites. These large rhabdomeric photoreceptors occur at later larval stages and persist through adulthood. Their bilogical function, however, remains unknown. The rhabdomeric nature of both Joseph and dorsal ocelli photoreceptors (see below) has been further confirmed by expression of amphioxus melanopsin (AmphiMop), a rhabdomeric opsin, and $G\alpha_q$ subunit, typically mediating the

phototransduction cascade in rhabdomeric photoreceptors (Koyanagi et al., 2005). The expression of any transcription factor has not been adressed yet.

The dorsal ocelli (organs of Hesse) developing in an 8d larva are scattered along the ventral part of the neural tube and consist of one rhabdomeric photoreceptor cell shielded by a single cup-shaped pigment cell. The only exception is the first Hesse organ which develops in early larva and is composed of one central pigment cell and two photoreceptors positioned symmetricaly on both sides (Eakin and Westfall, 1962, Wicht and Lacalli, 2005). Behavioral and physiological studies (Crozier, 1916) confirmed dorsal ocelli as functional photoreceptive structures in adult animal, however, with very slow physiological response (Guthrie, 1975). The biological role of dorsal ocelli is not known, although the distribution along the entire length of the body suggests their involvement in monitoring the proportion of amphioxus body burried in the sand.

The first organ of Hesse has been shown to express *Six4/5, Pax2/5/8, Eya* (Kozmik et al., 1999, Kozmik et al., 2007), but apparently does not express *Pax4/6* (Glardon et al., 1998).

1.3.3 Amphioxus and the origin of the vertebrate eye

The key phylogenetic position together with a detailed morphological description of different structures in amphioxus cerebral vesicle opened the discussion about homology and the origin of vertebrate eyes. Perhaps the most spread and generally accepted hypothesis proposes the lamellar body being homologous to the pineal organ or median eye. Due to a high level of morphological similarity of membrane folds observed in amphioxus lamellar body to the pineal organ of lamprey, these organs are considered homologous. The biological function of the lamellar body is not known, however, its rather big size, stacking of membrane, and lack of dark pigment potentially maximizes its sensitivity to light. This suggests a possible role in circadian perception or determining vertical position in water column according to decreasing/increasing light intensity (Lacalli, 2008).

The frontal eye counterparts were suggested to be the '*paired lateral eyes or, more precisely, the eyes plus the ventral structures that link them, the optic stalks and chiasmatic ridge*' (Lacalli, 1996). The homology of amphioxus frontal eye and lateral eyes of vertebrates is supported by the fact that if a developmental mutation 'cyclopea' occurs in vertebrates, the position of resulting single eye at the ventral anterior margin of the forebrain corresponds to the position of the frontal eye in amphioxus (Lacalli, 1994). Next,

the sequential organization of Rows 1 - 4 in amphioxus frontal eye resembles different layers of the vertebrate retina. Besides the proposed homology of photoreceptor cells of Row1 and Row2 to vertebrate photoreceptors, Row3 and Row4 neurons could roughly correspond to vertebrate horizontal/amarcrine and bipolar/ganglion cells, respectively, based on morphological similarity (Lacalli, 1996, Lacalli et al., 1994). Since detailed morphological studies of amphioxus frontal eye do not suggest its origin by reduction of a larger structure, the ancestral chordate condition was most probably a small single eye and the pair of lateral eyes represents a vertebrate innovation (Lamb et al., 2007).

An alternative hypothesis proposing both the pineal organ and lateral eyes being derived from the lamellar body, has been formulated (Satir, 2000), based on several facts: The cilia of the lamellar body photoreceptors are of 9+0 type (Eakin and Westfall, 1962), which are also present in rod outer segments, whereas the cilia found in the photoreceptors of the frontal eye have 9+2 organization, typical for olfactory or motile cilia. According to this hypothesis, the phenomenon of larval ciliary hovering and orientation towards a light source could be explained either by providing maximal shielding of frontal eye photoreceptive cilia, or alternatively, by maximizing the exposure of the lamellar body to light. Next, this hypothesis also proposes that the inverted vertebrate retina originated by lateral evagination of walls of a structure similar to the lamellar body and subsequent formation of optic stalk and eyecup with inverted position of photoreceptors.

The rhabdomeric photoreceptors – Joseph cells and dorsal ocelli are generally considered as homologues to intrinsically photosensitive retinal ganglion cells (Arendt, 2003, Provencio et al., 2000), which is well supported by the expression of r-opsin (melanopsin) and $G\alpha_q$ (Koyanagi et al., 2005).

2 Aims

Despite of amphioxus evolutionary importance, its photoreceptors have not yet been subjected to molecular studies. With the advantage of recent completion of *B. floridae* genome, the putative genes playing a role in phototransduction as well as key eye-specific regulators could be easily identified and characterized. To get primary insight into ancestral chordate condition with respect to eye evolution, the work aims at the following points:

a) using bioinformatics identify opsins and Gα subunit genes in the complete genome of *B. floridae*
b) based on available literature data, select a set of photoreceptor-specific transcription factors which may be involved in opsin gene regulation in amphioxus
c) clone and characterize the expression of transcription factors identified in b)
d) develop suitable tools and design an experimental setup for testing the regulation of opsin genes by selected transcription factors

3 Experimental procedures

3.1 Obtaining animals

Adult *B. floridae* animals were obtained during summer spawning season (June-August) in Tampa Bay, Tampa, Florida. Ethanol and methanol-fixed larvae and embryos were kindly provided by Linda Z Holland (Scripps Institution of Oceanography, UCSD, California).

Fixed embryos of *Platynereis dumerilii* stored in 100% methanol for immunohistochemistry were provided by Maria A. Tosches (Developmental Biology Unit, EMBL Heidelberg).

Adult wild-type C57/Bl6 mouse was obtained from IMG animal facility.

3.2 Molecular cloning and vectors

All standard molecular cloning methods were performed as described in (Ausubel et al., 1995).

3.2.1 RNA isolation and cDNA synthesis

Total RNA was isolated from fresh samples or from specimens stored at -20°C in RNA Later (Ambion) using TRIzol reagent (Invitrogen) according to manufacturer`s protocol. Glycogen (15µg/ml) was added to TRIzol to increase the yield of RNA.
Random-primed cDNA was synthetised by PrimeScript reverse transcriptase (TaKaRa) using random hexanucleotides for priming.

3.2.2 Molecular cloning and vectors

For synthesis of probes for RNA *in situ* hybridization, DNA fragments were inserted into pBluescriptIIKS+ (Stratagene) or pCR4-TOPO (Invitrogen) cloning vectors. DIG-labelled RNA was synthetised using appropriate (Sp6, T7 or T3) RNA-polymerase (Roche) and RNA DIG-labelling mix (Roche).

3.2.3 Oligonucleotides

For the list of nucleotides used in this work see appendix 8.2.

3.3 RNA in situ hybridization

Whole-mount RNA *in situ* hybridization to amphioxus larvae was performed according to standard protocol (Holland et al., 1996). The only modification was ommiting the levamisol in washes on day 3, step vi.

3.4 Preparation of antisera

3.4.1 Expression strains and constructs

For overexpression of protein fragments, pET system (Novagen) was used. Selected coding sequences were cloned into pET42a(+) vector to create proteins containing 6xHis-GST fused to protein fragment of interest. The resulting plasmids were transformed into *E. coli* BL21 DE3-RIPL strain and inducibility was tested in small scale experiments. The clone providing the best expression was selected for large scale production of protein.

3.4.2 Purification of protein fragments for rabbit immunization

A total volume of 500ml fresh LB medium without antibiotics was inoculated by overnight culture grown in LB medium supplemented with 12.5 μg/ml chloramphenicol and 30 μg/ml kanamycin. Bacteria were grown at 37°C, 200 RPM until OD_{600} reached 0.6, induced by 0.5 mM IPTG for 3 hours. Cells were harvested at 6000xg for 20 minutes and the pellet stored at -80 °C until further processing.

The pellet was resuspended in Lysis buffer (6M guanidine hydrochloride, 0.1M NaH_2PO_4, 0.01M Tris.Cl, pH 8.0, supplemented with fresh β-mercaptoethanol to final concentration of 20mM). The suspension was sonicated 6x20s and incubated for 3 hours at room temperature. The resulting lysate was centrifuged at 10 000x g for 10 minutes and the supernatant mixed with Ni-NTA agarose beads (QiaGen) previously equilibrated with Urea buffer (8M urea, 20 mM Tris.Cl, 50 mM NaH_2PO_4, 100 mM NaCl, pH 8.0, supplemented with fresh β-mercaptoethanol to a final concentration of 20mM). The suspension was incubated on a rotating platform overnight at room temperature. The beads with bound proteins were washed 2 times with 40 ml Urea buffer and loaded onto disposable chromatographic column (Bio-Rad). The column was washed with Urea buffer with decreasing pH (8.0 – 6.8) and His-tagged protein was eluted by Urea buffer pH 4.2 into several 1ml aliquots. After elution, pH was immediately increased to 7.5 by 1M Tris.Cl, pH 8. Protein concentration was estimated by Protein Assay Reagent (Bio-Rad).

For immunization, virgin female rabbits New Zealand White (Charles Rivers laboratories, Germany) were used. Rabbits were immunized 3-4 times in one month intervals with 300-500 µg of purified protein mixed with Freunds Adjuvans (Sigma) in each immunization step.

3.5 Immunohistochemistry

For immunohistochemistry, specimens were fixed in 4% PFA in 1xPBS for 1 hour at room temperature, washed 3x in 1x PBS and stored in 1x PBS for short-term storage at 4 °C or transferred to 100% methanol for long-term storage at -20 °C.

Specimens were transferred to 1x PBT (1x PBS, 0.1% Tween 20) through 50% and 25% methanol in 1xPBS. Specimens were washed 3x 20 minutes in 1x PBT, blocked in Blocking solution (10% BSA in 1x PBS) for 1 hour at room temperature and incubated with primary antibodies diluted in Blocking solution over-night at 4 °C. Next day, specimens were washed 3-4 times in 1x PBT for 20 minutes at room temperature and incubated with secondary antibodies in Blocking solution for 2 hours at room temperature. Secondary antibodies were washed away 3x 20 minutes in 1x PBT. Nuclear labeling was carried out by incubation with 1µg/ml DAPI in 1x PBS and washing 3x 5 minutes at room temperature. For fluorescence/confocal microscopy, the specimens were mounted in VectaShield (Vector Laboratories, Inc) using small coverslips as spacers between microscopy slide and coverslip.

The antibodies used in this work were as follows:

Antibody	Antigen	Dilution WB	Dilution IHC	Reference
AmphiMop	amphioxus melanopsin	-	1:500	(Koyanagi et al., 2005)
Melanopsin	Mouse melanospin	-	1:1000	(Provencio et al., 2002)
Ac-tub	acetylated β-tubulin	-	1:500	Sigma-Aldrich T6793
Prox1Mel	Human Prox1 HD-prospero domains	1:1000	1:500	(Duncan et al., 2002)
Prox1	Mouse Prox1 C-terminal peptide	-	1:1000	Chemicon AB5475
Pdu-C-Prox	*Platynereis* Prox - prospero domain	1:500*	1:250*	This work
Amphi-C-Prox	*B. floridae* Prox - prospero domain	1:500*	1:250*	This work
Amphi-N-Otx	*B. floridae* Otx - N-terminus up to HD	1:500*	1:250*	This work
Amphi-C-Otx	*B. floridae* Otx - from HD to the end	1:500*	1:250*	This work
Amphi-N-Rx	*B. floridae* Rx - N-terminus up to HD	1:500*	1:250*	This work
Amphi-C-Rx	*B. floridae* Rx - from HD to the end	1:500*	1:250*	This work

*) Dilutions listed for 1:1 mixture of raw rabbit sera with sterile 100% glycerol.

Confocal scans were acquired on Leica Sp5 microscope equipment and processed in Leica Application Suite 1.8.0 software.

Flat-mount staining with two primary antibodies from the same species was performed as follows: Adult C57/Bl6 mice were sacrificed by cervical dislocation, eyes removed and fixed in 4% PFA in 1x PBS for 1.5 hour. Retinae were dissected, washed 3x 20 minutes in 1x PBT, blocked 1 hour in Blocking solution and incubated overnight at 4° C with primary antibody directed against a nuclear protein. Next day, retinae were washed 5x 20 minutes in 1x PBT and incubated 2 hours at room temperature with secondary antibody in Blocking solution. Retinae were thoroughly washed 5x 30 minutes and incubated overnight at 4°C with the second primary antibody (directed against cytoplasmatic/membrane protein). All subsequent washes were performed analogous to the washes after primary antibody. Specimens were then transferred to 1x PBS, stained with DAPI and washed 3x 10 min in 1x PBS before mounting in Vectashield.

3.6 Luciferase reporter assays

For all luciferase reporter assays, Human Embryonic Kidney (HEK293T) cells (ATCC) were used. The cells grown in a 24-well plate were transfected at 50-60% confluency with 300 ng of total plasmid DNA using polyethylenimine (PEI), according to standard protocols. The DNA mixture used for transfection typically contained 100ng reporter vector based on pGL3-Basic (Promega), 100ng TF expression vector, 5ng normalization reporter vector pRL-TK and pBSII as filler plasmid to adjust the amount of DNA to 300ng per well. Two days after transfection, the cells were lyzed for 20 minutes using Passive lysis buffer (Promega) and assayed for firefly and renilla luciferase in Luminoskan TL (LabSystems) luminometer using STOP-Glo kit (Promega).

3.7 Bioinformatics

For producing multiple sequence alignments, ClustalX (Larkin et al., 2007) and Muscle (Edgar, 2004) software were used. The alignments were edited by BioEdit (Hall, 1999) and phylogenetic trees were inferred using Phylip Package (Felsenstein, 2005).

Searches in *Branchiostoma floridae* JGI Genome V1.0 and V2.0 (http://genome.jgi-psf.org/Brafl1/Brafl1.home.html) were performed by on-line user interface using *tblastn* a *blastp*. Missing exons and unappropriate model prediction were corrected manually and the resulting putative transcripts translated. For the refinement of predicted genes to V2.0, the assembly was downloaded and local *tblastn* was performed in BioEdit.

4 Results

To get a preliminary insight into amphioxus photoperceptive capabilities, we first investigated the key components common for both rhabdomeric and ciliary phototransduction cascade – the opsins and Gα protein subunits.

4.1 Opsin genes in Branchiostoma floridae genome

A total number of 20 different *bona fide* opsins was identified (Holland et al., 2008) in JGI *B. floridae* V1.0 genome release. All predicted protein sequences possess the key lysine residue at the end of seventh transmembrane helix and other features allowing the classification of these sequences as opsins (Figure 7). Notably, the genomic organization of predicted opsin genes shows several clusters of related genes, suggesting their origin by a recent duplication (Figure 8).

Later analysis of V2.0 genome release reduced the number to 18 putative opsin genes: Model pairs 87094/110002 (AmphiOp6 group) and 70446/70447 (c-opsin group) map to the same genomic position in V2.0 assembly. From these 18 opsins identified, seven are clear orthologues of previously identified *B. belcheri* opsins (Koyanagi et al., 2002, Koyanagi et al., 2005).

Figure 7 (next page): *B. floridae* opsin genes
Phylogenetic tree and alignments of the opsin family.The phylogenetic tree was inferred by the neighbor-joining (NJ) method using murine adenosine and chicken serotonine receptors as outgroup sequences. The numbers above each node represent the percentage of bootstrap probability based on 500 replicates. Putative *B. floridae* opsins clustered consistently in tree inferred by the Maximum likelihood method. Previously published *B. belcheri* opsins are black-framed and opsins identified in this work are typed in red and numbered according to the *B. foridae* v1.0 gene model ID from which the protein sequence was derived. Multiple sequence alignment of the opsins represented in the NJ tree: The black lines above the alignment represent the extent of transmembrane helices III, IV and VII. The color lines on both sides of the alignment mark opsin subfamilies. Amino acid fingerprint allowing opsin classification is indicated by black arrowhead and numbered according to bovine rhodopsin protein sequence. The lysine residue K296 is critical for covalent binding of retinal via Schiff base linkage, which is stabilized by counterion E113 or E181 (Terakita, 2004). The (E/D)RY triade highly conserved amongst GPCRs is important for G-protein interaction (Franke, 2002). The residues G121, P171, W175 were found to be evolutionary trace residues typical for opsin family but not for GPCRs in general (Madabushi et al. 2004). The position of ‘HPK’ and ‘NKQ’ motif typical among rhabdomeric and ciliary opsins, respectively (Arendt, 2004) are red-framed. The classification of opsin families is represented in the color boxes on the right side. Some V1.0 gene models represent the same gene due to mapping to the same genomic position in V2.0 (black bars) or overall similarity and suggestions from expression data (brown bars). Models with frame-shift mutation in EST/RT-PCR sequence are marked with asterisk. See appendix 8.5.1 for sequence accession numbers.

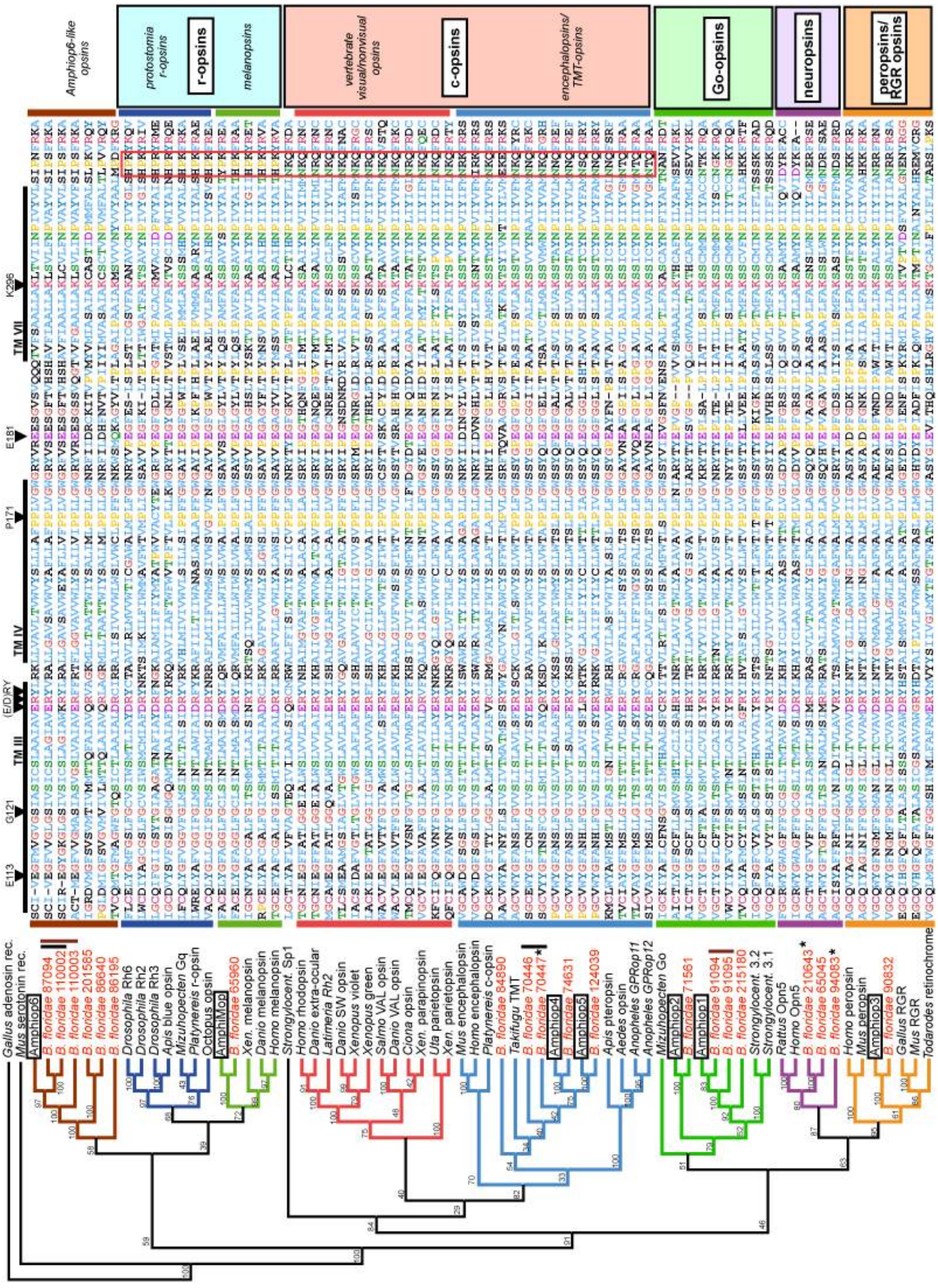


Figure 7: *B. floridae* opsin genes (see previous page for figure legend)

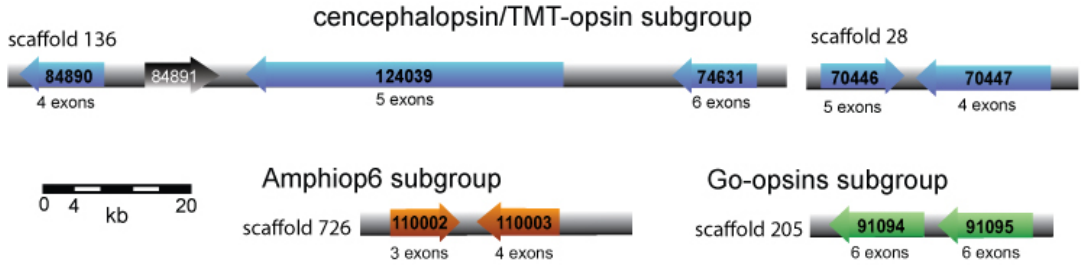


Figure 8: Genomic organization of *B. floridae* opsin genes
Clusters of opsin genes found in *B. floridae* genome suggest a recent origin by tandem gene duplication. Note that except for the three c-opsin gene models positioned at scaffold 136, the clusters at scaffolds 28, 205 and 726 most probably represent the same gene due to artifacts of inproper assembly in V1.0.

4.2 Expression of opsin genes in B. floridae

The expression of putative opsins identified *in silico* was addressed in adult and late neurula (12 and 14h) stages by RT-PCR and subsequent sequencing (Table 1), data not shown). From 18 predicted opsin genes, at least 16 are expressed in adult animal, and a subset of this number is expressed in late neurula. The failure of confirming the expression of 110003 (AmphiOp6) and 91095 (Go-opsins) may reflect a redundant prediction of these gene models caused by inaccurate contig assembly in both V1.0 and V2.0 release. If so, gene models 110003 and 91095 might represent highly polymorphic alleles of gene models 87094 and 91094, respectively, reducing the total number of distinct opsins to 16 genes (brown bars in Figure 7).

To further confirm the expression of opsins, *B. floridae* EST database available at NCBI (www.ncbi.nih.gov) and *B. floridae* cDNA database (http://amphioxus.icob.sinica.edu.tw/) were screened with predicted protein sequences. Only two models yielded a hit – EST clone CAXF6053 (gastrula stage) for model 84890 (c-opsins) and CAXC16239 (neurula stage) for model 210643 (neuropsins).

The sequencing of retrieved EST clones as well as the RT-PCR products, however, suggested that some of the expressed *bona fide* opsin genes may not produce a functional protein. The EST clone of 210643 contains a 5bp deletion leading to deleterious frame-shift and similarly the RT-PCR product sequences of models 94083 and 70447 revealed point deletions (marked with asterisks in Figure 7). Since the RT-PCR products sequenced to confirm their identity were typically 150-250bp long and did not cover the entire CDS, further cloning and sequencing of predicted CDS may reveal more indels, thus reducing the total number of functional opsin genes even further.

Model ID	Group	RT-PCR								UCSC GenomeWiki	*B. belcheri* best match	
		Embryo		Adult						Sequence confirmed	Name	Name	% AA Identity
		12h	14h	WA	A	O	M	V	T				
87094	AmphiOp6			YES						YES	MEL6	Amphiop6	67.6
110003	AmphiOp6										MELx		
201585	AmphiOp6			YES									
86640	AmphiOp6	YES	YES	YES						YES			
86195	AmphiOp6	YES		YES						YES			
65960	r-opsins	YES		YES	YES	YES	YES		YES	YES	MelMop	AmphioMop	88.8
84890	c-opsins	YES		YES	YES	YES	YES	YES	YES	YES + EST	TMTy		
70446	c-opsins			YES						YES	TMTx		
74631	c-opsins			YES						YES	Enceph4	Amphiop4	83.3
124039	c-opsins			YES						YES	TMT5	Amphiop5	73.3
71561	Go-opsins			YES	YES	YES		YES	YES	YES	PER2	Amphiop2	61.9
91094	Go-opsins		YES	YES			YES	YES		YES	PER1	Amphiop1	76.0
91095	Go-opsins												
215180	Go-opsins	YES	YES	YES						YES			
210643	Neuropsins	YES		YES						EST	NEUR1a		
65045	Neuropsins	YES	YES	YES						YES	NEUR1b		
94083	Neuropsins	YES		YES				YES	YES	YES			
90832	Peropsins	YES	YES	YES	YES	YES	YES	YES		YES	PER3	Amphiop3	84.9

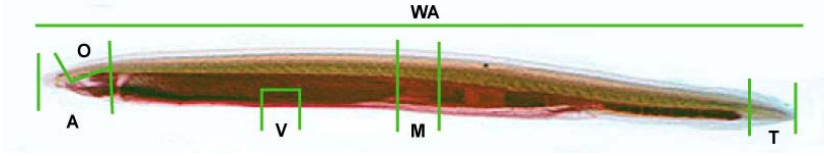


Table 1: Analysis of *B. floridae* opsins expression
A total number of 18 *B. floridae bona fide* opsins was screened for mRNA expression at different embryonal stages and in different parts of adult body (see the scheme below the table for annotation) by RT-PCR. In most cases the identity of PCR product was confirmed by sequencing. If the results were not conclusive or the experiment was not performed no indication of expression is provided. The correpsonding 12 opsin genes proposed in UCSC opsin resource (http://genomewiki.ucsc.edu/index.php/Opsin_evolution:_key_critters_(deuterostomes)) and seven *B. belcheri* opsins identified in (Koyanagi et al., 2002) with respective AA identity are listed.

4.3 G_α subunits in Branchiostoma floridae *genome*

Bioinformatic search for G-protein alpha subunits in *B. floridae* genome V1.0 resulted in identification of ten putative genes (Figure 9), however, further re-examination of V2.0 assembly revealed only seven unique Gα subunits. Every identified Gα subunit could be clearly assigned to one of the five Gα subfamilies (GNAI, GNAS, GNAQ, GNA12 and GNAV) comprising the encestral metazoan Gα complement (Oka et al., 2009). In bilaterians, the GNAI sufamily includes GNAI and GNAO genes (plus GNAT genes in vertebrates), a situation also confirmed here for amphioxus. From all the predicted protein models only one (ID 89300) could not be assigned to any of these subfamilies due to limited sequence information; and also the screen of cDNA/EST databases to identify the missing sequence information was unsuccessful.

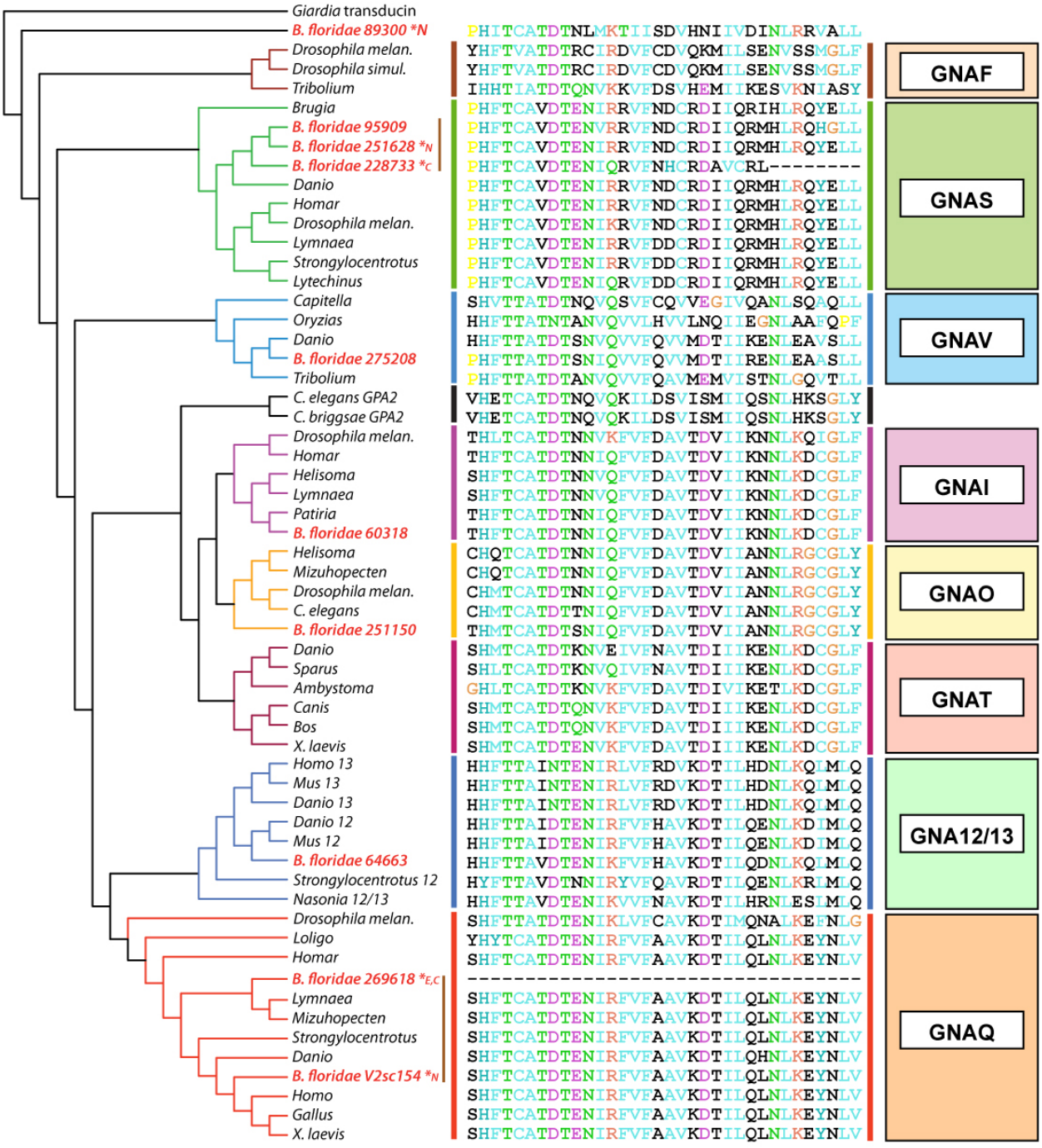


Figure 9: Amino-acid sequence alignment of Gα subunit C-termini

The phylogenetic tree was inferred by the neighbor-joining method (NJ) using *Giardia* 'transducin' as an outgroup sequence. The missing sequence at either N- or C-termini did not allow to infer the phylogenetic tree of all sequences at once, therefore the tree shown here is based on two separate reconstructions. For this reason, the bootstrap support values are not shown here, however, the tree reflects the topology of the two separate trees. Missing sequence information in amphioxus predicted Gα subunits is marked by asterisk and a letter denoting the missing part (E – exon, N – N-terminal portion, C – C-terminal portion). The V1.0 predicted models mapping to the same genomic position in V2.0 assembly are marked with brown vertical bar. The gene model 251628 (GNAS subfamily) does not map to the same position, however, the overall sequence similarity suggests its identity with model ID 95909. See appendix 8.5.2 for sequence accession numbers.

4.4 Identification and cloning of AmphiRx

Blasting *B. floridae* V1.0 genome database with *Platynereis dumerilii Rx*, *Mus musculus Rx* and *Ciona intestinalis Rx* protein sequences resulted in finding a single putative CDS of AmphiRx covered by three exons. The predicted protein sequence contains all features characteristic for invertebrate and vertebrate Rx proteins (appendix 8.3): a 60AA long Q50 paired-type homeodomain, an octapeptide positioned at the N-terminus and Rx and OAR domains within the C-terminal portion. Apparently, this gene is not homologous to the partial sequence named '*B. lanceolatum* Rx homeobox' deposited in the NCBI database (accession nr. DQ401115). The full-length CDS of *AmphiRx* was cloned from cDNA prepared from an adult animal, suggesting its expression and function at this stage.

4.5 Identification and cloning of AmphiProx

Blasting mouse, *Xenopus* and *Drosophila* Prox/Prospero protein sequence against *B. floridae* V1.0 genome release yielded putative *AmphiProx* CDS covered by three exons. Several EST clones containing *AmphiProx* CDS were identified and revealed additional exons and partial 5` and 3` UTRs. The full-length CDS was cloned from adult animal suggesting its role in adult stages, however *AmphiProx* mRNA seems to be also maternaly deposited, since one of the ESTs containing *AmphiProx* (BFEG036i02) was derived from egg mRNA.

Interestingly, the exon-intron organization of *AmphiProx* is not similar to other deuterostome *Prox* genes, and a detailed look at other *Prox* genes reveals that the genomic organization is highly diverse within the deuterostomes: Whereas vertebrate Prox genes are rather uniform, having four coding exons, the two *C. intestinalis* paralogues *Prox-a* and *Prox-b* of Prox are coded by 9 and 11 exons, respectively, and *S. purpuratus* Prox CDS is covered by four exons.

Several conserved domains described in vertebrate Prox amino acid sequences (Q-rich, P-rich) are not found in AmphiProx and other non-vertebrate sequences (see appendix 8.4 for AA sequence alignment). On the other hand, a short stretch of sequence (AKRARVEN) in the N-terminal portion of the protein seems to be conserved from vertebrates to insects. The conservation is disrupted only in *Ciona* Prox variants reflecting again a rather high level of sequence diversification in this species.

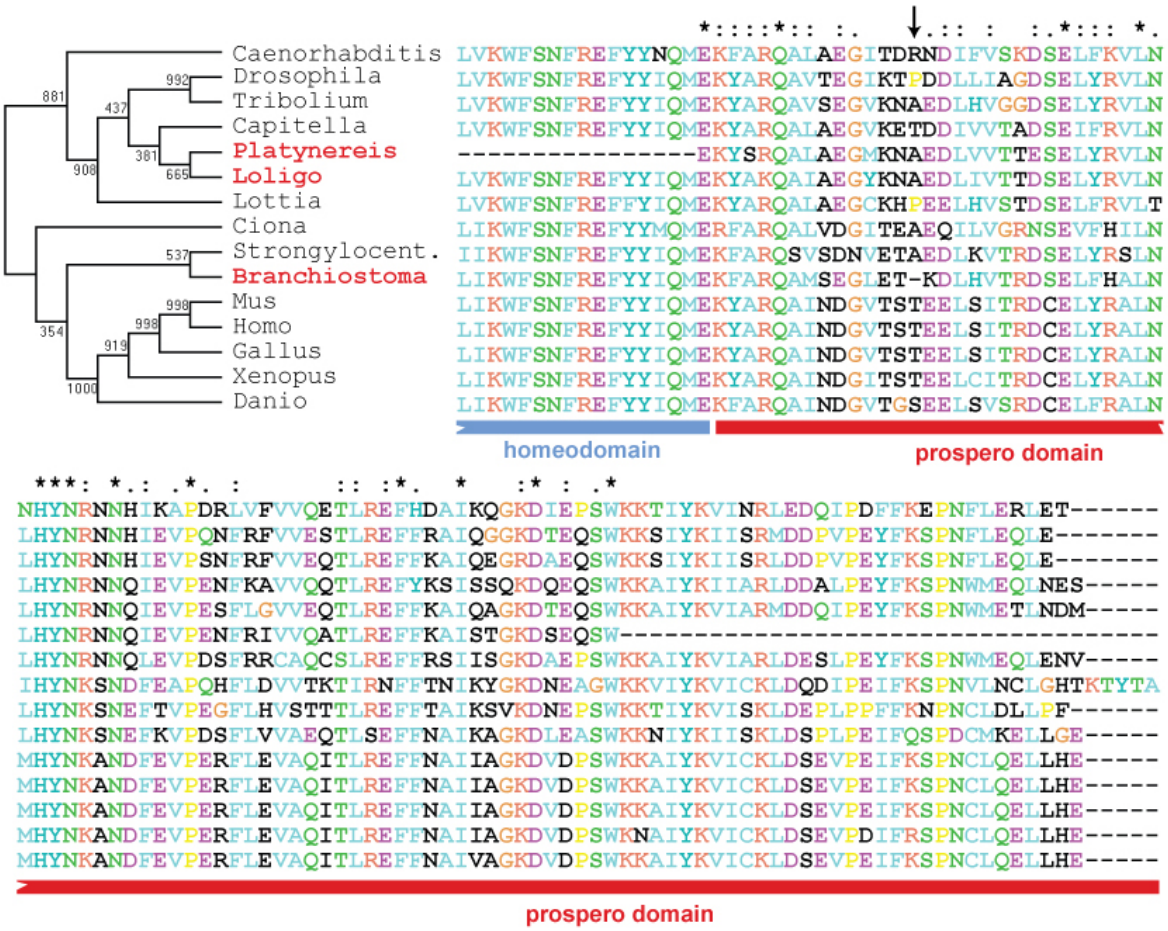


Figure 10: Phylogenetic analysis and AA-sequence alignment of C-terminal portion of Prox protein fragments prepared in this work.
An NJ-tree was inferred using ClustalX (1000 bootstrap steps, all gaps excluded). The alignment shown covers the last 16 AA of the homeodomain followed by the prospero domain. The missing sequence at the beginning of Pdu-Prox and at the end of *Loligo* Prox is caused by unsucessful extension of the CDS fragment. Note the high level of sequence conservation between *Platynereis* and *Loligo* protein sequences. Also note the single AA deletion at position 14 of AmphiProx (black arrow).

The prospero domain adjacent to the homeodomain at the C-terminus represents an ultimate hallmark of Prox proteins and is highly conserved in all species examined. Intriguingly, amphioxus prospero domain contains a single amino acid deletion at position 14 not observed in any other bilaterian species (Figure 10).

4.6 Preparation of AmphiOtx, AmphiRx and AmphiProx antisera

For the preparation of antisera for AmphiRx and AmphiOtx, two independent protein fragments were selected. To avoid cross-detection of other homeodomain proteins, the

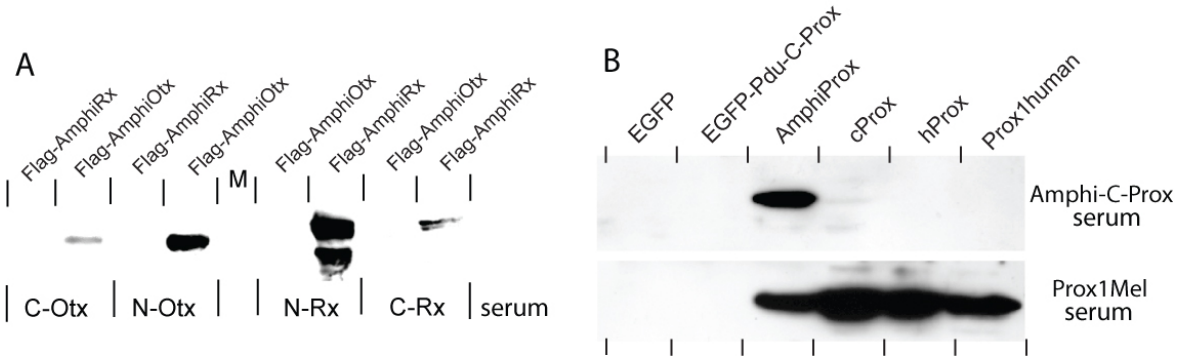


Figure 11: Western-blot analysis of rabbit polyclonal sera
HEK293T cell lines were transfected with an expression vector carrying Flag-tagged full-length coding sequences of the proteins indicated. Two days after trensfection a whole-cell extract was prepared and tested by western-blotting. The crude rabbit antisera were used at 1:500 dillution. (A) Testing the antisera against AmphiRx and AmphiOtx. Any given antiserum was tested on a double-lane blot containing the appropriate protein and another protein not used for immunization as a negative control. The lower band in N-Rx detected AmphiRx lane may be caused by protein degradation. (B) The Prox1 serum prepared against human PROX1 homeodomain-prospero domain fragments (Duncan et al., 2002) crossreacts with chicken Prox and AmphiProx, but not with the prospero domain of *Platynereis* Prox (see below). The antiserum raised against AmphiProx prospero domain only, however, does recognize only AmphiProx and not its vertebrate counterparts.

---

homeodomain was excluded, and either N-terminal or C-terminal portions of the protein were used for immunization. For Prox protein, only the C-terminal portion of the protein, *ie* the prospero domain was used. After the 3rd immunization boost, the rabbit polyclonal sera were tested for immunoreactivity with respective Flag-tagged full length AmphiRx, AmphiOtx and AmphiProx proteins (Figure 11) expressed in HEK293 cell line.

The strong conservation of the prospero domain suggested that some of the previously published Prox antibodies raised against this domain might work with AmphiProx as well. Indeed, the antibody against the protein fragment of human Prox containing the homeodomain and prospero domain cross-reacts with chicken and AmphiProx (Figure 11 B).

4.7 Re-examination of AmphiOtx expression in the frontal eye area

Although the expression pattern of *AmphiOtx* has been published more than a decade ago (Williams and Holland, 1996), its expression in the frontal eye, despite being mentioned in the publication, was not the main goal of the study. Since the expression of this gene in the cerebral vesicle has not been a subject of any further study since then, a

more detailed examination of *AmphiOtx* expression and extension to later developmental stages was needed to understand its role in amphioxus photoreceptor development. The expression of *AmphiOtx* in the cerebral vesicle was observed in three major areas (Figure 12):

1) The most anterior part of the cerebral vesicle

The expression domain of *AmphiOtx* at 30h stage includes the very most anterior cells of the cerebral vesicle, where Row1 and Row2 photoreceptor cells and pigment cells originate. In a 3.5d larva, where these cells have already differentiated, the shape of the *AmphiOtx*-positive cells (which project forward to the pigment cup) confirms their identity as Row1 photoreceptors. The anterior expression domain includes more cells than just Row1, most probably Row2, but not Row3, which lies more caudally. The exact information about the identity of these more caudal cells is hard to assess due to differences between stages examined here and a 12.5 larva, for which the detailed anatomical studies have been done (Lacalli, 1996). From the staining, it is also not clear, whether the signal covers also the pigment cells, which lie more dorsally to Row1, closer to the neuroporus. Thus, *AmphiOtx* may or may not be expressed in these pigment cells.

2) The posterior part of the cerebral vesicle

Based on the relative position to the point where the nerve cord narrows, the AmphiOtx-positive cells lie just in front of the infundibular cells. In a 12.5 days-old larva, this area contains a variety of cell types, *eg* balance organ or PPN2 cells, however, at the stages examined here the main question arises, whether these cells are already post-mitotic or not. The region in between the anterior and posterior part of the cerebral vesicle is not rich in differentiated neurons and probably contains still proliferating cells awaiting later differentiation. Generally, the differentiation in the cerebral vesicle occurs mainly at the anterior and posterior ends with the middle region containing udifferentiated cells (Holland and Holland, 2006). As the cerebral vesicle lenghtens, the cell divison stops in most of the area except for the cells positioned just anteriorly of the preinfundibular region, which undergo another burst of division. At the 30h stage, when there is still one round of division to come, these cells are still expressing rather high level of *AmphiOtx*. However, the expression ceases at later stage (3.5 days old larva), when these cells have exited the cell cycle. Taken together, the expression of *AmphiOtx* in the most-posterior domain of the cerebral vesicle may reflect the developmental role of Otx in specifying the cerebral vesicle.

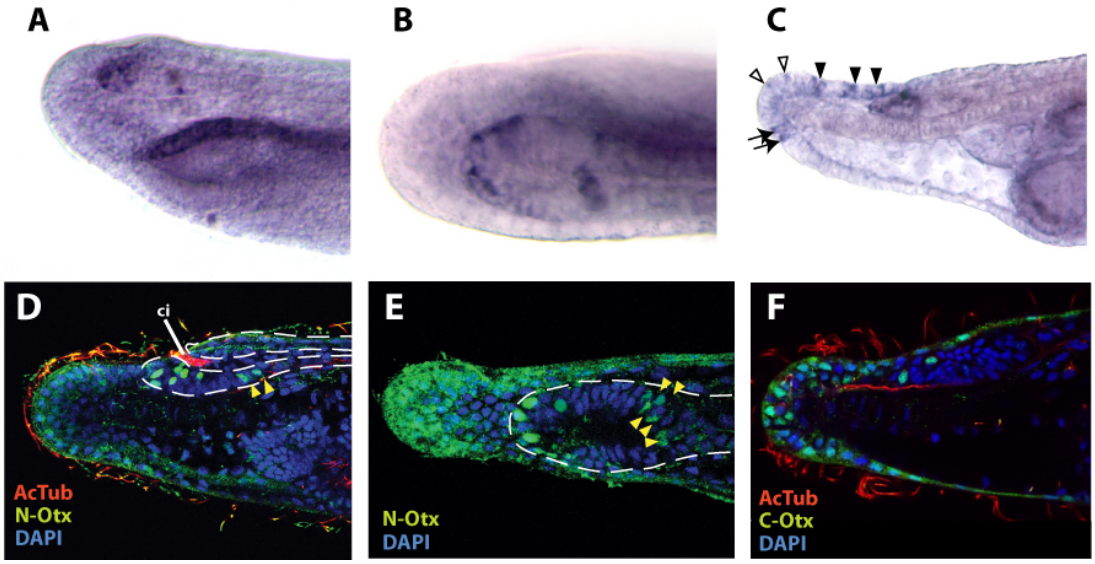


Figure 12: *AmphiOtx* expression in the cerebral vesicle detected by RNA *in situ* hybridization and immunohistochemistry
Lateral (A) and dorsal (B) view of a 30h larva showing the expression of AmphiOtx in the anterior-most and posterior-most domains of the cerebral vesicle. Later on, in a 3.5d larva (C) where the pigment spot has already formed, the expression is still apparent in the most-anterior part, however, only weakly detectable in the most-anterior part. The expression of *AmphiOtx* also occurs in a few rostral epithelial cells – dorsal and ventral groups pioneering the anterior-dorsal nerve (black arrowheads and black arrows, respectively), and a distinct population of rostral sensory cells at the tip (white arrowheads). Immunohistochemical staining with rabbit antisera prepared in this work confirmed the results of in situ hybridization. (D) A confocal-microscopy image created by overlying several Z-stack confocal planes showing the nuclear signal of AmphiOtx visualised by antibody against its N-terminal portion. Besides the signal in the most anterior rows of cells underlying the bundle of outcoming cilia (ci, stained by acetylated β-tubulin antibody), the signal also occurs in the cells lying in front of the infundibular organ (yellow arrowheads). (E) the same specimen as in (D), dorsal view, several confocal planes overlayed. The cytoplasmatic/background staining in the rostrum and weakly apparent also on the border of the specimen may be caused by nonspecific binding of the antisera. (F) IHC staining with antibody against the C-terminal portion of AmphiOtx resembles the staining with antibody raised against N-terminal portion and also confirms the expression of AmphiOtx in rostral epithelial cells.

---

3) Epithelial sensory cells

The expression of *AmphiOtx* could be also observed in several epithelial sensory cells of the rostrum (Lacalli, 2004). One group of these cells (black arrowheads in Figure 12 C) lies at the dorsal surface of the rostrum and pioneers the anterodorsal nerve. Another group of cells positioned at the tip of the rostrum (white arrowheads in Figure 12 C) are the rostral epithelial sensory cells pioneering the rostral nerve. Since the number of the cells stained is very low in comparison to the total number of this cell type at later stages, it is possible that more cells progressively develop starting with a few cells shown here. The

most ventral cells (black arrows in Figure 12 C) are distinct from the rostral cells and their axons probably pioneer the ventral branch of the anterior-dorsal nerve.

The rabbit antisera prepared against fragments of AmphiOtx protein were tested on 30h larvae, to confirm their tissue specificity. Despite a cytoplasmatic background staining occuring mainly in the epithelial cells of the specimens, the nuclear signal recapitulated the mRNA signal observed by *in situ* hybridization.

4.8 Expression of AmphiRx during amphioxus development

To get primary insight into the developmental expression of *AmphiRx*, RT-PCR analysis was performed on mRNA isolated from different developmental stages. The transcript was first detected in 9h embryonic cDNA library with continuous expression through later stages to adulthood.

The spatial pattern of *AmphiRx* was assessed by RNA *in situ* hybridization using a full-length CDS probe. No signal was detected in early stage embryos (9-20h) probably due to very low expression not detectable by RNA *in situ* hybridization. The earliest stage, where *AmphiRx* transcript was detected, was late neurula (24h). The signal was restricted to the anterior part of the animal and no signal was observed in any other cells throughout the body. The strongest signal at the developmental stages from late neurula to early larva (30h) is located in the anterior part of the ventral cerebral vesicle, spreading more caudally than the anterior domain of *AmphiOtx* expression. Although *AmphiRx* expression seems not to be present in the anterior-most cells in 24 – 30h larvae (Figure 13 A-D), these cells express *AmphiRx* later (3.5d larva, Figure 13 E). Plastic-embedded sections of a 3.5d larva (Figure 13 G-L) showed a clear *AmphiRx* signal in the cells lying just under the pigment spot (Figure 13 G – I and L), which belong to Row1 and Row2 putative photoreceptor cells, and a rather weak signal in a few rows of cells positioned more caudally (Figure 13 G, K and L). The pigment cell bodies seem to lack *AmphiRx* expression. The signal was also observed in the lamellae of the lamellar body (Figure 13 J), but not in the bodies of the lamellar cells, which lie dorsally on both sides of the lamellar body. Rostral expression of *AmphiRx* seems to include the tip-positioned receptors connected the rostral nerve but not the sensory cells pioneering the anterodorsal nerve (Figure 13 D).

From the two different antisera raised against N- and C-terminal regions of AmphiRx, the N-terminal serum showed a stronger signal in western-blot analysis and was applicable for IHC staining (Figure 13 F). Although the N-Rx serum produced a nuclear

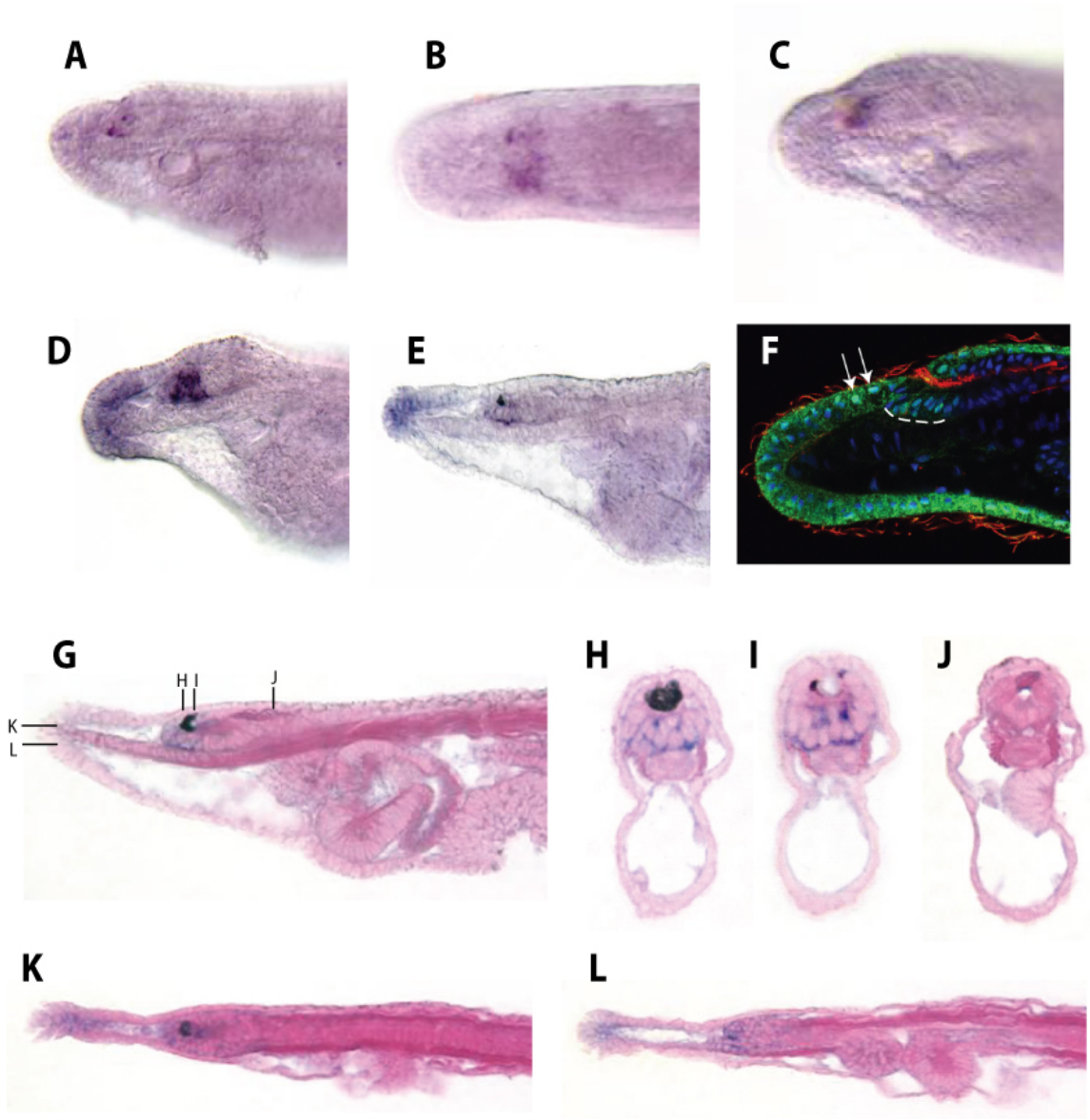


Figure 13: Expression of *AmphiRx* during amphioxus development
Lateral (A) and dorsal (B) view of a 24h larva showing the onset of *AmphiRx* expression detectable by RNA *in situ* hybridization. The signal is present in symmetrical lateral areas of the anterior part of the developing cerebral vesicle; (C) and (D) 30h larvae showing a strong signal in the anterior half of the cerebral vesicle, possibly excluding the most-anterior group of cells. Note the strong signal in the rostrum of (D); (E) *AmphiRx* expression in a 3.5d larva with the pigment cells already differentiated, the strongest signal is apparent in the cells underlying the pigment cup including Row1, 2 and probably Row3 and the intensity of the signal decreases with more posterior position. A strong signal is also present at the tip of the rostrum, and possibly at the dorsal side closer to the neuroporus; (F) IHC staining with the antiserum prepared against the N-terminal part of AmphiRx. Despite the high background signal, a nuclear signal is still observable in several cell at the most anterior tip of the cerebral vesicle. White arrows point to two nuclei of the dorsal rostral epithelial cells. (G) sagittal section of specimen (E) through the middle of the cerebral vesicle; (H) and (I) - cross-sections of an *AmphiRx* stained 3.5 days larva at the levels of the most anterior tip (H) and most posterior end (I) of the pigment cup. The signal is present in the cytoplasm and apices of the Row1 and Row2 cells; (J) cross section at the level of the lamellar body. The signal seems to be present in the lamellae, but not in the bodies of the lamellar cells. (K and L) frontal sections showing the gradient of signal decreasing from the anterior-most cells to the posterior part of the cerebral vesicle.

signal in the anterior part of the cerebral vesicle, which is in accordance with the RNA *in situ* results, the overall background in the epithelim was rather high. For this reason, it was hard to clearly distinguish the nuclear signal observed in the dorsal part of the rostral epithelium, however, since the signal increases in the position of the nuclear DAPI staining, it is most probably a real nuclear AmphiRx signal.

4.9 Co-expression of Prox with rhabdomeric opsins

Although the expression pattern of *AmphiProx* has been adressed by RNA *in situ* hybridization (data not shown) in early developmental stages, no expression was detected in the first organ of Hesse, which is the only rhabdomeric photoreceptor structure present at these stages. For adressing possible expression of AmphiProx in amphioxus photoreceptors, the immunohistochemistry in combination with confocal microscopy was used. This methodology also allowed studying fully developed Hesse organs in adult animals.

4.9.1 Expression of *AmphiProx* in rhabdomeric photoreceptors

Since the surface layer of adult animals is hardly penetratable for antibodies, for IHC staining of adult amphioxus neural tissue, the neural tube has been always dissected from fixed animals and pulled out of the neural sheath. The dissected neural tubes are suitable for mounting to perform confocal microscopy scanning.

To visualize the rhabdomeric photoreceptor cells, the previously published antibody against *B. belcheri* melanopsin (Koyanagi et al., 2005), which is the only rhabdomeric opsin in *B. floridae* genome, was used. This antibody clearly stains bodies of cells closely associated with the pigment cups as was observed in *B. belcheri*. Since AmphiMop antibody has been raised in mouse and Prox antisera used here originate from rabbit, the co-expression of Prox and melanopsin could be studied by double IHC. Both the Prox1Mel antibody (Duncan et al., 2002) and Amphi-C-Prox (this work) stained consistently the nuclei of Joseph cells as well as the dorsal ocelli (Figure 14). No difference in the signal was observed in dorsal ocelli positioned in the anterior, middle or tail part of the neural tube. Prox signal was observed in all melanopsin-positive cells, and in addition, a few Prox-positive nuclei were also scattered in the surrounding neural tissue. Using IHC, the coexpression of AmphiProx and AmphiMop has been confirmed also in the first organ of Hesse (data not shown).

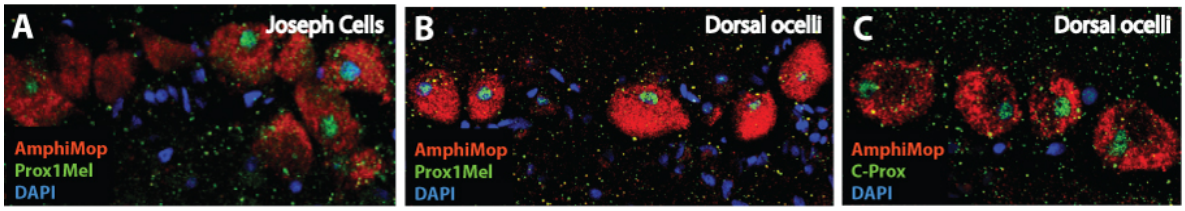


Figure 14: Expression of *AmphiProx* in amphioxus rhabdomeric photoreceptors
AmphiMop antibody staining was used to mark amphioxus rhabdomeric photoreceptors. Prox1Mel antibody detected nuclear signal of Prox protein in (A) Joseph cells as well as (B) dorsal ocelli. Staining with Amphi-C-Prox antibody led to the same result, shown here only for dorsal ocelli (C).

4.9.2 Cloning and expression pattern of *Platynereis dumerilii Prox (Pdu-Prox)*

The encouraging results of *AmphiProx/AmphiMop* coexpression strongly supported the hypothesis of Prox being a conserved marker of rhabdomeric photoreceptor cell type (proposed in (Arendt, 2003)). Still, the available data come only from ecdysozoa (*Drosophila*), vertebrata (chicken) and amphioxus (this work). To expand the dataset to more animal branches, *Platynereis dumerilii* was selected for adressing possible Prox expression in rhabdomeric photoreceptors. *P. dumerilii* belongs to annelids, thus the data would represent the first Prox expression data within the lophotrochozoans, and in addition, it has been shown that *Platynereis* retains many of the urbilaterian ancestral features (Raible and Arendt, 2004).

A short fragment of *Pdu-Prox* was cloned by degenerate PCR (see appendix 8.2 for primer sequences) from a 48/72h cDNA library mixture (kindly provided by Keren Guy, Arendt lab, EMBL, Heidelberg) and extended at the 3`end by SMART RACE procedure. The 3´ partial mRNA sequence of total length 1117 bp contains a 309 bp region coding the C-termiunus of putative Pdu-Prox. The corresponding amino-acid sequence covers the last amino acid of homeodomain and the entire prospero domain (Figure 10), which enables to clearly classify the cloned gene as a member of Prox transcription factor family. The classification is also supported by phylogenetic analysis (Figure 10).

RNA *in situ* hybridization (performed by Keren Guy, Arendt lab, EMBL Heidelberg) revealed the expression of *Pdu-Prox* in both larval and adult eyes, which consist of rhabdomeric photoreceptors (Figure 15 A). Using a new microscopy/image analysis procedure – '*in silico* expression profiling' (developed by Raju Tomer Arendt lab,

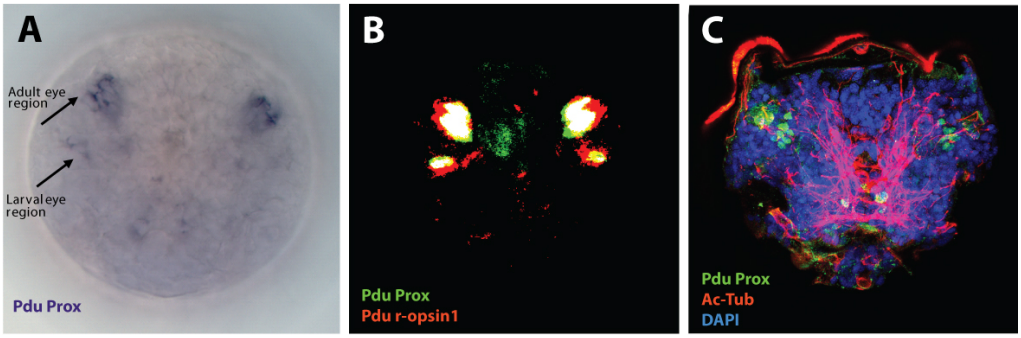


Figure 15: Expression of Pdu-Prox in larval and adult eyes of *P. dumerilii*
(A) RNA *in situ* hybridization signal was observed in both larval and adult eyes area in a 48h larva; (B) Prox expression co-localizes with Pdu-r-opsin expression (data generated by Keren Guy and Raju Tomer, EMBL Heidelberg); (C) antibody against Pdu-Prox prospero domain labels the nuclei of rhabdomeric photoreceptors, larval eyes area is not visible in this confocal plane.

EMBL Heidelberg) *Pdu-Prox* was shown to be coexpressed with *Platynereis* rhabdomeric opsin (Figure 15 B).

Rabbit polyclonal antibody against Pdu-Prox prospero domain was prepared and tested as described previously for amphioxus Prox, Rx and Otx antibodies. Together with Pdu-Prox antibody, two antibodies against Pdu-Rx were also prepared (for reasons, see discussion). All the antisera detect their corresponding protein in western-blot (Figure 16). Immunohistochemistry with rabbit serum generated against the prospero-domain of Pdu-Prox confirmed the RNA *in situ* hybridization results by the presence of nuclear signal in the position corresponding to adult and larval eyes (Figure 15 C).

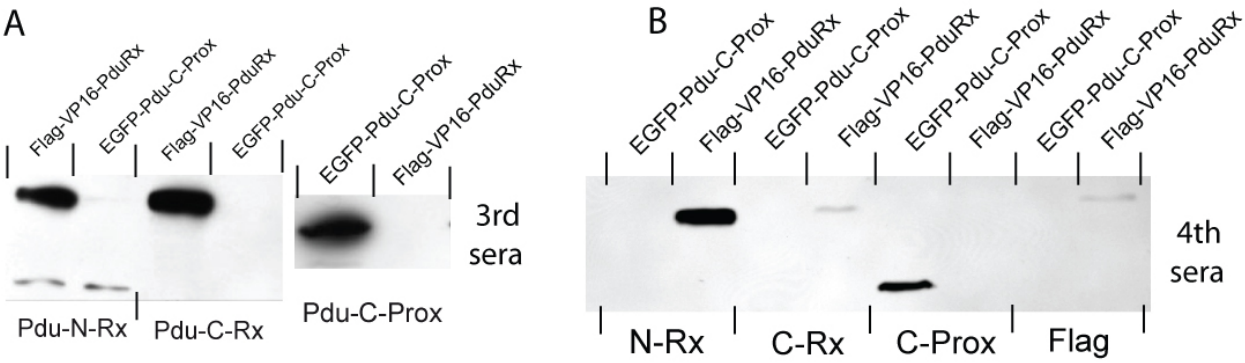


Figure 16: Western-blot analysis of rabbit polyclonal sera against Pdu-Prox and Pdu-Rx
(A) sample sera after the third immunization showed a good reactivity to their corresponding proteins; (B) after the fourth immunization, the specificity of N-Rx antibody was increased (no bands of lower MW), however, the sensitivity of C-Rx antibody apparently decreased. C-Prox antibody does not seem to be affected by the fourth immunization.

4.9.3 Expression of Prox1 and melanopsin in mouse RGCs

Prox/melanopsin coexpression was also tested in mouse retinal ganglion cells (RGCs), which have been suggested to be the vertebrate descendants of rhabdomeric photoreceptor lineage (Arendt, 2003). A small subset of RGCs has been shown to express melanopsin and control the circadian rythms (Provencio et al., 2002). In another study (Dyer et al., 2003), a small subset of RGC has been also shown to express Prox1, which raises the question, whether these cells are identical with the melanopsin-positive subset of RGCs. The double IHC staining using the melanopsin and Prox1 antibody, however, showed that these two cell populations are not identical (Figure 17).

4.10 Transactivation properties of AmphiOtx, AmphiRx and AmphiProx

The previous results and data from literature suggested that AmphiOtx, AmphiRx and AmphiProx might play a conserved role in regulation of c- and r-opsins (see discussion for details). For better understanding of this process and designing future experiments to test this hypothesis, pilot experiments have been done to assess the general transactivation properties of these transcription factors. For this purpose, several artifical reporter vectors have been prepared and tested: The K50 and Q50 homeodomain proteins, represented by AmphiOtx and AmphiRx, respectively, have been shown to interact with DNA binding sites containing TAATC or TAATT (Hanes and Brent, 1991). The reporter constructs used here beared a minimal promoter fused to three copies of these DNA-binding motifs dimerized in a head-to-head manner allowing the binding of homodimerized transription

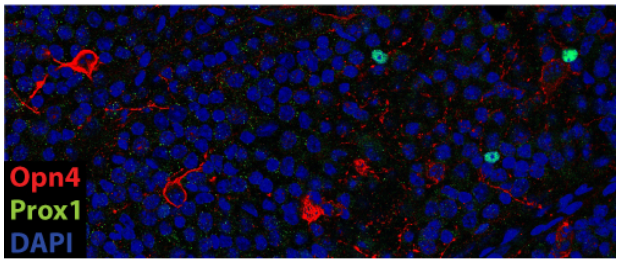


Figure 17: Prox1 and melanopsin are not coexpressed in mouse RGC
Sequential double IHC was performed with antibody against mouse melanopsin (Provencio et al., 2002) and antibody against mouse Prox1 and retinae were flatmounted for confocal microscopy. Prox1 signal does not localize to the nuclei of melanopsin (Opn4)-positive cells.

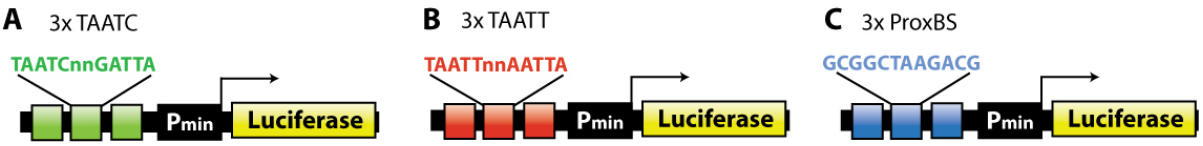


Figure 18: Artificial reporters used for characterization of AmphiOtx, AmphiRx and AmphiProx
The reporter constructs for testing AmphiOtx and AmphiRx contained multimerized DNA-binding motifs TAATC (A) and TAATT (B), which have been shown to interact with K50 and Q50 homeodomain proteins in previous reports (Hanes and Brent, 1991); (C) three copies of *prospero*-binding site from *Drosophila* Rh6 gene promoter were used to create the Prox-responsive reporter.

factors (Figure 18 A and B). Since no consensus sequence is known for Prox-type homeobox proteins, three copies of the *prospero*-binding element from *Drosophila* Rh6 gene promoter (Cook et al., 2003) were fused to minimal promoter to create a Prox-responsive reporter construct (Figure 18 C).

4.10.1 AmphiOtx

Amphioxus Otx protein has strong activatory properties, as has been shown for other members or Otx transcription factors. The activating properties are encoded mainly by the C-terminal portion (from HD to the stop codon) of the protein, as shown by Gal4-reporter assay (Figure 19 A). Interestingly, the Gal4-fusion of the full-length AmphiOtx did not

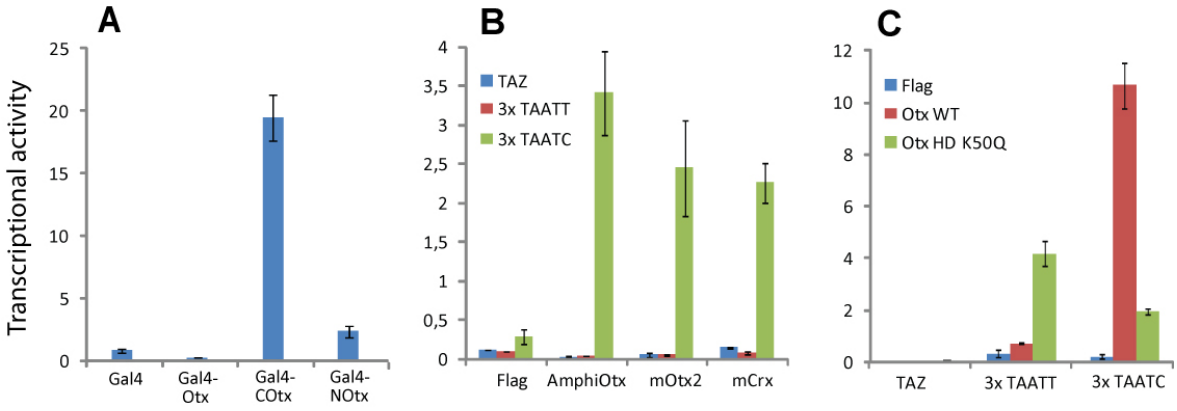


Figure 19: Transcriptional properties of AmphiOtx:
(A) The activation properties of AmphiOtx are mostly present in its C-terminal region, although the N-terminus has weak activation properties as well. The fusion of Gal4 to full-length AmphiOtx resulted in no activation. (B) AmphiOtx, as a K50 homeodimain protein, activates only artificial reporters containing 3x palindromic TAATC motif, but not a reporter containing TAATT sites. (C) by substituting K50 for Q50 in AmphiOtx homeodomain, this selectivity is reversed, however, the selectivity of Q50 homeodomain seems to be weaker than the selectivity of K50 HD. In all cases, empty TF-expression vectors (Gal4, Flag) and empty luciferase reporter vectors (TAZ) were used as negative controls.

lead to activation, probably due to interference between the two DNA-binding domains (Gal4 and HD). AmphiOtx as well as mouse Otx2 and Crx (all K50 homeodomain proteins) were able to activate the K50-responsive reporter, but not the Q50-responsive reporter (Figure 19 B). By mutating the intrinsic K50 for Q50 within the AmphiOtx homeodomain, the selectivity of DNA binding was reversed (Figure 19 C). This result is in agreement with the observations reported previously in other biological system (Hanes and Brent, 1991) and also proves the functionality and specificity of the reporter constructs generated in this work.

4.10.2 AmphiRx

The N-terminal and C-terminal protein portions adjacent to the HD mediate a weak repression in Gal4-fusion reporter tests (Figure 20 A) and weak repressive properties were also observed with full-length AmphiRx protein. Interestingly, in the case of AmphiRx, repressive effects were observed on both K50- and Q50-responsive reporters (Figure 20 B). The repressive properties, however, can be overcome by fusing AmphiRx to a strong activatory domain derived from VP16. The resulting VP16-AmphiRx protein is capable of activating both K50- and Q50-responsive reporters, although the selectivity for the TAATC DNA-binding motif was lower than that of AmphiOtx (Figure 20 C).

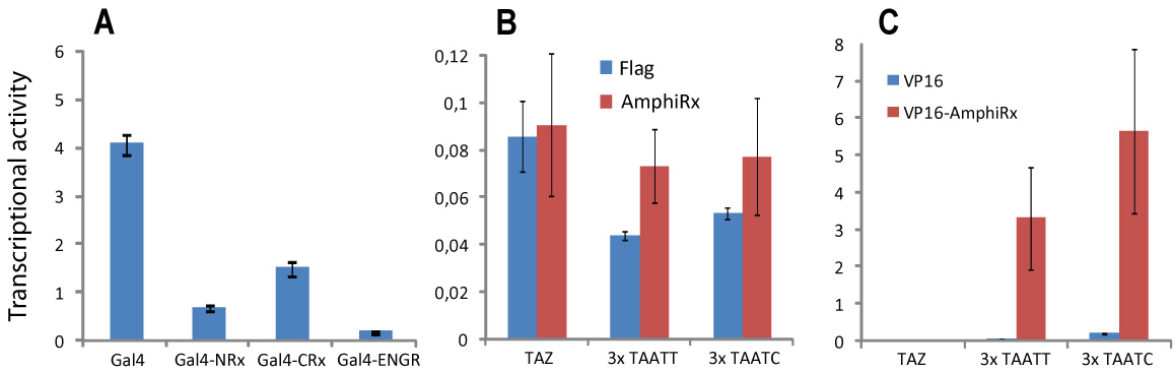


Figure 20: Transcriptional properties of AmphiRx:
(A) Fusion of Gal4-DNA-binding domain to C-Rx and N-Rx protein fragments led to repression of luciferase expression in Gal4-reporter assay, suggesting repressive effect of naked N-Rx and C-Rx protein fragments. Gal4-*engrailed* fusion (Gal4-ENGR) serves as a positive control for repressive effect; (B) Since the K50- and Q50- responsive reporters contransfected with whole AmphiRx expression vector showed the same luciferase activity at a very low level, whole AmphiRx protein does not have any apparent activating or reppresive properties. (C) Fusing AmphiRx to the strongly activating VP16 domain shows that AmphiRx recognizes both TAATT an TAATC DNA binding motifs.

4.10.3 AmphiProx

The transactivation properties of AmphiProx were primarily tested on previously reported natural targets: *Drosophila* Rh6 gene promoter, which is repressed by *Drosophila prospero* (Cook et al., 2003), and chicken βB1-crystallin promoter, which has been shown to be activated by chicken Prox (Cui et al., 2004). The same results, namely activation of βB1-crystallin promoter and repression of *Drosophila* Rh6 gene promoter, were observed with AmphiProx (Figure 21 A), suggesting that Prox transcription factors probably do not act alone on the promoters, but their transactivation properties are context dependent. This assumption was partially confirmed by an artificially constructed Prox-responsive reporter. Although it is strongly activated by VP16-AmphiProx fusion protein (Figure 21 B), suggesting that AmphiProx recognizes and binds to the GCGGCTAAGACG element, native AmphiProx does not activate nor represses the same reporter, suggesting that its action on Rh6 and βB1-crystallin is mediated by additional interaction partners occupying the promoter.

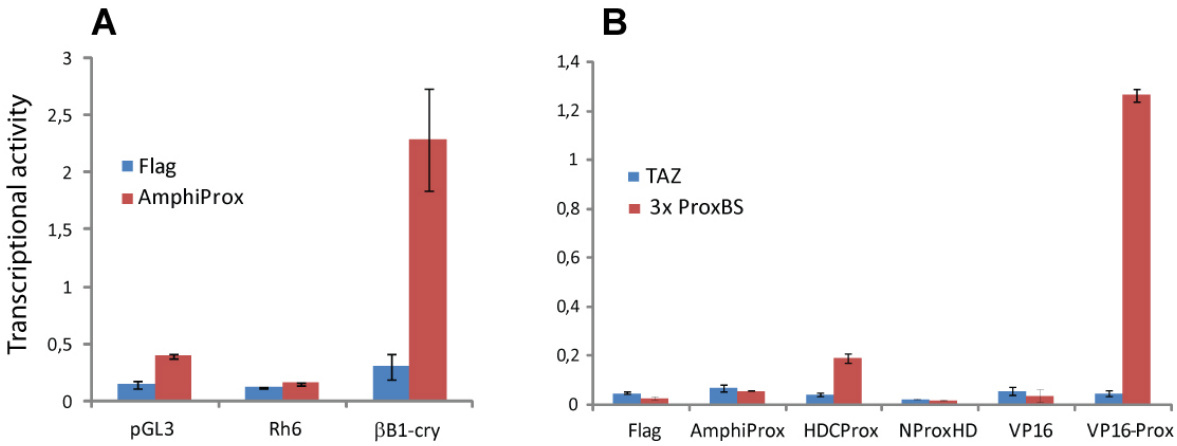


Figure 21: Transcriptional properties of AmphiProx:
(A) Whereas AmphiProx acts as a repressor at *Drosophila* Rh6 promoter, it activates the chicken βB1-crystallin promoter, suggesting its ambiguous transcriptional activities dependent on promoter context. (B) As a full-length protein, AmphiProx does not have any effect on artificial reporter vector containg 3x palindromic TAAGACG (Prox-responsive element derived from *Drosophila* Rh6 promoter). The recognition of 3x TAAGACG binding motif has been also confirmed with strongly activating VP16-AmphiProx fusion. Dissection of AmphiProx identified an intrinsic activation region within homeodomain-prosepro domain portion.

5 Discussion

5.1 *Opsins and G_α subunits in* B. floridae *genome*

The genome-wide survey of amphioxus genome revealed an unexpected expansion and diversification of opsin genes. Amphioxus opsin complement includes all major opsin subfamilies, which have been suggested to be present at the base of bilaterians: r-opsins, c-opsin, Go-opsins, and peropsin/RGR-opsins (Terakita, 2005, Raible et al., 2006, Plachetzki et al., 2007). In addition, amphioxus possesses members of an RGR-opsin-related subclass termed *neuropsins* (*Opn5*), identified previously in mammals (Tarttelin et al., 2003). Although a member of this subfamily has not been found in *Ciona* (tunicates) or *Strongylocentrotus* (echinodermates), its presence in amphioxus pushes back the date of origin of this group to the chordate ancestor.

Probably the most interesting class of opsins identified in *B. floridae* genome is the AmphiOp6 group with a single member isolated previously from *B. belcheri* (Koyanagi et al., 2002) and at least four paralogues identified in *B. floridae*. No other members of this subfamily have been found in any other animal, suggesting its origin within cephalochordate lineage from the r-opsin class before the split of different amphioxus species. Despite being most closely related to r-opsins, this opsin class substantially differs, mainly by the modification of the HPK tripeptide motif at the interface between the 7th transmbebrane helix and the cytoplasmatic loop. Since this motif has been shown to be one of the most essential sites for interaction with Gα and has been strongly conserved in the r-opsin class, the possible $G\alpha_q$ interaction might have been lost in AmphiOp6 opsins. Since the genomic survey of Gα subunits did not reveal any additional $G\alpha_q$-related subunit, which could hypothetically explain the preservation of signaling capabilities by co-evolution of opsin/Gα protein sequences, the interaction partner of AmphiOp6 opsins and possible phototransduction cascade, if any, remains an open question.

Since the Gα subunits play a role in multiple cellular signaling pathways in different body parts and cDNA/EST screen indeed confirmed the expression of most of the identified genes at various developmental stages, there was no point in adressing the temporal expression of these genes. For the same reason, the spatial expression of Gα subunits with regards to possible phototransduction cascades is not very informative until the tools allowing to adress a direct co-expression with opsins are generated. Mutliple sequence alignment of amphioxus Gα subunits revealed, rather expectedly, a very strong

conservation of C-terminal peptides, further corroborating the classification and raising the possibility to use commercially available antibodies, frequently raised against these evolutionarily conserved C-termini.

5.2 Expression of AmphiOtx and AmphiRx

The examination of expression of *AmphiOtx* and *AmphiRx* in the cerebral vesicle extended the current dataset available for key developmental genes expressed during amphioxus CNS development (integrated with previous studies to create an expression map in Figure 22).

The expression of AmphiRx in differentiated ciliary photoreceptors of the frontal eye is in a good agreement with the conservation of Rx as a marker of ciliary photoreceptor cell types (Arendt, 2003). Unfortunately, the expression of AmphiRx in the lamellar body which is the second ciliary photoreceptive structure was not conclusive, due to the presence of RNA *in situ* signal only in the area of lamellar folds, but not in the bodies of the lamellar cells. The AmphiRx antibodies generated during this work may help, after relevant developmental stage fixed for IHC becomes available, since only a clear nuclear signal is expected by AmphiRx IHC.

Besides the developmental expression pattern of *AmphiOtx* published previously (Williams and Holland, 1996) the data presented here reveal a possible role of AmphiOtx in differentiated ciliary photoreceptors of the frontal eye. On the other hand, AmphiOtx has not yet been detected in the rhabdomeric photoreceptors of the Hesse ocelli (data not shown), which is in contrast with an almost universal expression of *Otx* in both types of photoreceptors observed in multiple species across animal phyla.

5.3 Expression of AmphiProx and Pdu-Prox

The study of expression of amphioxus and *Platynereis* Prox brought a substantial extension of Prox/r-opsin co-expression data which have been so far available only for *Drosophila* and chicken. The co-expression of Prox and r-opsin in the rhabdomeric photoreceptors of these primitive, underived animals strongly support the previous suggestion that Prox is an evolutionarily conserved marker of rhabdomeric photoreceptors.

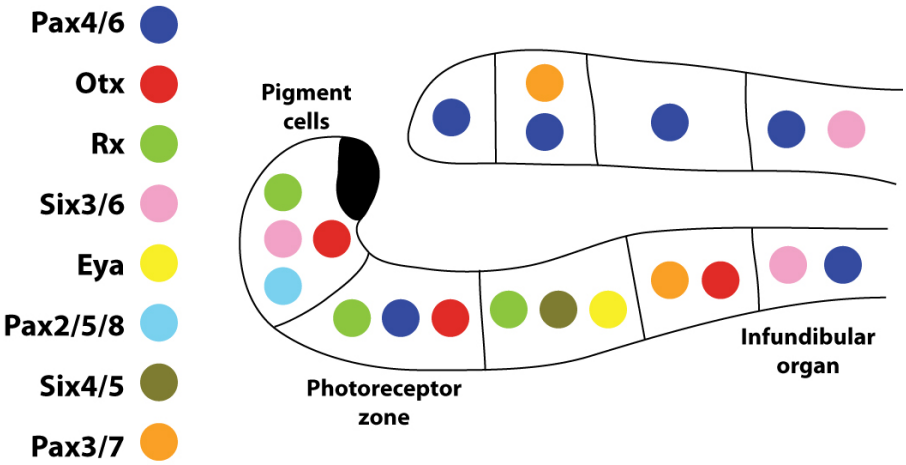


Figure 22: Summary of expression of different TFs in the cerebral vesicle of a 3d amphioxus larva (based on (Kozmik et al., 2007, Williams and Holland, 1996) and this work).

5.4 Model for evolution of transcriptional regulation of opsin genes

Available Rx, Otx and Prox expression data (reviewed in Introduction) together with the data retrieved in this work lead to the proposal of a new model for the evolution of transcriptional regulation of opsin genes (Figure 23). The model is an extension of the model proposed in (Vopalensky and Kozmik, 2009) and also integrates the paxcentric (bipartite) model of eye evolution (Kozmik, 2005) with recent notion of *Pax4* expression in vertebrate ciliary photoreceptors (Rath et al., 2009).

Frequent deployment of *Pax* and *Otx* genes in the differentiation of both types of photoreceptors, and the conserved expression of *Rx* in ciliary and *Prox* in rhabdomeric photoreceptors probably reflects their involvement in the regulation of key differentiation/structural genes functioning in these cells. Now, if we consider the original thoughts on evolution of differentiation gene batteries together with the evolution of opsin genes proposed in (Plachetzki et al., 2007), a scenario for evolution of opsin regulation emerges as depicted in (Figure 23). Since the split of r- and c-opsins predates the bilaterians as well as cnidarians, the regulation of these genes by Otx and Pax transcription factors must have been established very early in the metazoan evolution. Later on, after the c-opsin class diverged from the second lineage (giving eventually rise to all other opsin classes), Rx transcription factor was recruited for the regulation of c-opsin in ciliary photoreceptors. The rhabdomeric opsins, which diversified from the opsin lineage just before the protostome-deuterostome split, acquired in the newly formed cell type a new transcriptional regulator – Prox. Although the model does not consider other opsin

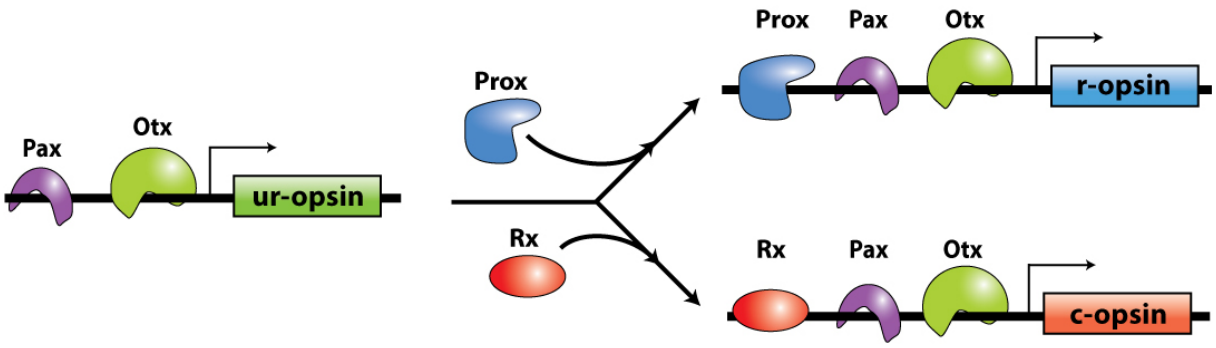


Figure 23: Proposed moded for evolution of transcriptional regulation of opsin genes
This hypothetical model is based on repeated occurence of a given transcription factor in given types of photoreceptor cells, and on an assumption that the reason for keeping the transcription factor in the cells is its involvement in the regulation of a key differentiation gene – opsin. The ur-opsin gene has been already regulated by the Otx gene and by Pax (not excluding other transcription factors involved). After the duplication and diversification of r- and c-opsins, the regulation by Otx and Pax has been preserved in both lineages. Retinal homeobox Rx has been recruited for regulation of c-opsin, whereas Prox homeobox has been co-opted for the regulation of r-opsin genes. With increasing complexity of animal body plans, all the transcription factors have consequently acquired additional roles in eye development.

---

lineages (*eg*. the Go-opsins), these could be regulated by Otx and Pax transcription factors as well, based on the same assumption used to create this model. In addition, one has to keep in mind the enormous stochasticity and plasticity of evolutionary process, with many gains and losses of function occuring in different animal phyla. Thus, the testing and verification of the model will require the collection of co-expression and regulatory data from multiple different species. Then, applying the maximum parsimony principle, the evolutionary ancestry (this model) *vs*. independent recruitment of the proposed TFs for opsin regulation could be assessed in more detail.

Obviously, there are already discrepancies between the prediction based on the model and experimental observations. Two pieces of data from this work are not in agreement with the proposed model. First, the *AmphiOtx* has not been detected in the rhabdomeric photoreceptors by *in situ* hybridization or IHC with AmphiOtx-specific antibodies. Since the dataset of *Otx* expression in both ciliary and rhabdomeric photoreceptor cell types covers multiple distinct branches of animal phylogeny, this unexpected result may point to an amphioxus-specific loss of this ancestral feature. Second, in murine ipRGC, the Prox1-positive cells are not identical with the melanopsin-positive cells. Melanopsin expression in the ipRGC might be hypothetically driven by Prox2 paralogue which is not recognized by the antibody used here. However, since Prox2 is not expressed in the neural retina of zebrafish and Prox2$^{-/-}$ mice do not have any

apparent retinal phenotype, this explanation is not very plausible. The lack of Prox1/melanopsin co-expression in mouse retina might rather reflect the change of melanopsin regulation in the mouse retina, due to requirements for more complex circadian regulation of melanopsin expression in nocturnal animals.

5.5 Future prospects

The proposed model predicts a universal employment of several transcription factos in the regulation of opsin genes in any animal phylum possessing a given photoreceptor cell type and opsin. Since several species- or phylum-specific losses of this regulatory relationship may occur, for testing the model a broad phylogenetic sampling is desireable, covering the species from all major branches of the animal kingdom: ecdysozoans, lophotrochozoans, deutorostomes and cnidarians.

Unfortunately, the ecdysozoan model organisms with advanced genetic tools, such as *Drosophila* or *C. elegans*, are not suitable for testing the model, since these organisms have lost the c-opsin (*Drosophila*) or all opsin genes (*C. elegans*). On the other hand, an emerging insect model *Tribolium castaneum* possesses a small number of opsins (two r-opsins and one c-opsin in total) and the ease of RNAi knock-down and transgenesis makes it an ideal model organism to test the proposed model *in vivo*.

Within the lophotrochozoans, the most plausible organism for testing the model is the annelid *Platynereis dumerilii* due to its extensive molecular dataset as well as current efforts in developing genomic tools (genome sequence, EST and BAC libraries etc.). Another major lophotrochozoan group of animals – the molluscs – do not include any well-established model organism, however, at least the expression data for *Otx*, *Rx, Prox* and opsins could be extended by basic molecular tools (RNA *in situ* hybridization, immunohistochemistry) in scallops, which have been shown to possess both ciliary and rhabdomeric photoreceptors (Kojima et al., 1997). Since the best-established cnidarian models *Nematostella* and *Hydra* do not possess any obvious photoreceptive structures, another cnidarian species *Tripedalia cystophora*, which has been recently shown to possess ciliary photoreceptors employing a ciliary opsin (Kozmik et al., 2008), may provide very valuable information: Although the functional regulatory studies are not possible in *Tripedalia*, only the mere expression *Otx* and *Rx* homologues (which are present in cnidarian genomes) in the c-opsin-positive photoreceptors would support or weaken the proposed model outside the bilateria. The last major branch of animal kingdom – the deuterostomes – is represented by the chordate amphioxus, for which the expression data

regarding the proposed model have been provided in this work. Since the possibilities to identify and test amphioxus cis-regulatory modules of opsin genes in homologous system are very limited, to reveal the direct involvement of a specific TF in regulation of amphioxus opsins requires a combination of multiple approaches:

Since the genome of *B. floridae* is known, it can be easily searched for known DNA-binding consensus sequences in the proximity of the genes of interest (opsins). Since a majority of false-positive results is produced by this approach, the phylogenetic footprinting analysis has been proven very helpful in identifying the real and functional TF-binding sites within the cis-regulatory modules. The attempts to find evolutionary conserved elements in non-coding regions between amphioxus and tunicates or vertebrates failed due to high level of diversification. A possible solution of this problem could rely in comparison of different amphioxus species eg. *B. floridae* and *B. lanceolatum*, whose split occured about 120 millions of years ago (Nohara et al., 2005). Unfortunately, the BAC library for *B. lanceolatum* is still underway and although the BAC library from a more distant amphioxus *B. belcheri* has been constructed (Wang et al., 2005), its availability for scientific community is very limited. In addition, whether the evolutionary distances between these amphioxus species are suitable for cis-regulatory elements comparison is still not clear, since our preliminary analyses of *B. belcheri* Pax4/6 gene BAC (available at NCBI) and *B. floridae* Pax4/6 genomic locus revealed a rather high level of sequence diversification.

5.5.1 Cell-culture based screen for TF-binding sites

To bypass these obstacles, a cell culture-based system for adressing TF-binding sites has been designed and is currently being developed in the laboratory. The system (Figure 24) is based on cell lines with inducible expression of VP16-fusion of a selected transcription factor. Such cells are transfected with luciferase reporter constructs containing random fragments of BAC clones of a gene, for which the TF-binding sites are to be identified. Since the pilot experiments with amphioxus Rx, Otx and Prox transcription factors have shown that their endogenous transactivaion capability may not be strong enough to be easily measured, the VP16-fusion is necessary for reproducible and standard read-out. Due to strong activating properties of VP16-fusion, once the TF binds specifically to its DNA-binding motif, a strong luciferase signal will be produced, regardless of the transactivating properties of the TF itself. The size of BAC DNA

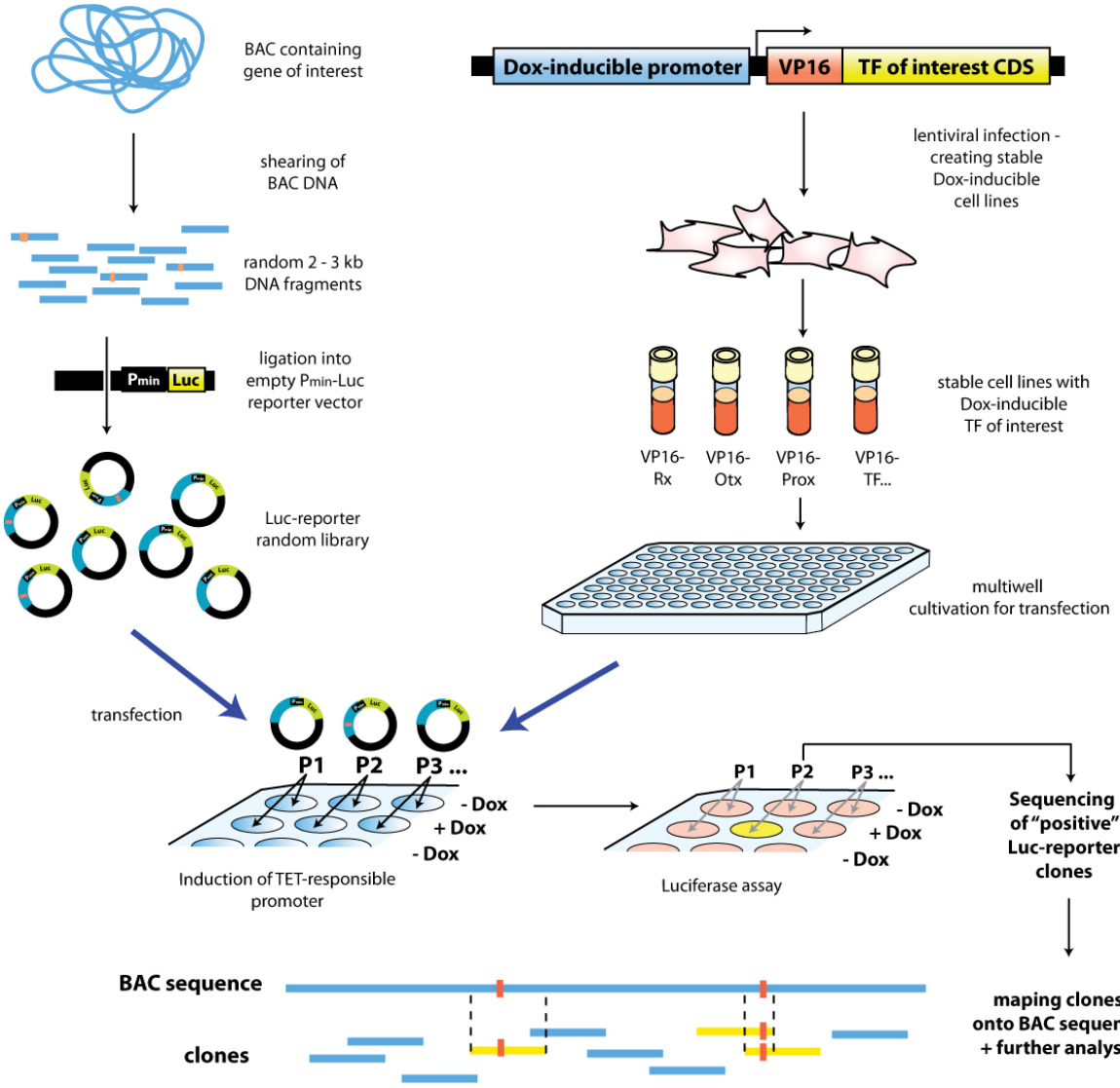


Figure 24: A flow-chart of medium-throughput screen for TF-binding sites within BAC DNA
A BAC clone containing the gene, whose regulatory elements are being addressed (*eg.* opsin), is mechanically sheared to produce 2-3kb fragments. These fragmets are cloned into a reporter vector containing a minimal promoter and luciferase gene, thus producing a library of reporters containing random BAC fragments upstream of the Luc reporter gene. These reporter constructs are transfected into stable cell lines containing a tightly regulated cassette expressing VP16-fusion of the transcription factor of interest upon induction with doxycycline (Dox) molecule. As the VP16 domain converts any transcription factor to a strong activator, a strong luciferase signal should be produced in wells which have been transfected with a reporter clone bearing a DNA fragment containing the TF-binding site. The clones producing a strong luciferase signal are sequenced, mapped to the original BAC sequence and subjected to further more detailed analysis.
Once the panel of Dox-inducible VP16-TF cell lines is generated, this approach allows a fast screen of rather large segments of genomic DNA (~ 150kb per one BAC) for TF-binding sites and substantially narrows down the size of the genomic sequence to be analyzed in more detail.

fragments and the number of random reporter constructs screened is balanced to reach a coverage factor 2-3x. Such a system allows a fast, medium-throughput screen of rather large segments of genomic DNA (~ 150kb per one BAC) for functional TF-binding sites and substantially narrows down the size of the genomic sequence to be analyzed in more detail. Moreover, it also enables to find the DNA-binding sites of those TFs for which the DNA-binding consensus is not known and could not be therefore identified *in silico*.

5.5.2 Chromatin immunoprecipitation (ChIP)

A second methodology, which could adress the involvement of a given transcription factor in regulation of a certain gene even without the availability of transgenic methods, is chromatin immunoprecipitation (ChIP). If a target gene is known and putative TF-binding sites identified in its regulatory regions, standard ChIP procedure can be performed. However, a broader insight into the portfolio of downstream genes on a whole-genome level could be acquired by sequencing immunoprecipitated DNA fragments (ChIP-Seq). For this technique, the crucial requirements are antibodies recognizing the TF of interest and sufficient amount of biological material. A sufficient amount of biological material is available for laboratory-held model organisms such as *Platynereis* and *Tribolium*. For amphioxus, since all the transcription factors and multiple opsins are expressed in adult animals, the primary ChIP experiments can also be carried out on adult animals which are available in substantial amounts during the collecting season. The necessary antibodies for testing the Rx and Prox part of the proposed model in amphioxus and *Platynereis* have been generated in this work.

6 Conclusions

In this work, a bioinformatic survey of amphioxus genome led to the identification of amphioxus opsin and Gα subunit complement. Although most of the opsin genes have been shown to be expressed in different developmental stages, further experiments, mainly the spatial expression pattern, will be needed for understanding their biological role.

The set of selected transcription factors (Otx, Rx and Prox) has been cloned from amphioxus and the expression pattern has been addressed in different developmental stages. The preparation of rabbit polyclonal antibodies against these transcription factors has proven to be a very useful methodology enabling to study gene expression at cellular resolution. Together with advances in confocal microscopy, a panel of such antibodies raised against important developmental genes will allow generation of a detailed expression map of amphioxus cerebral vesicle. The antibodies developed here also represent a powerful tool for other techniques, such as chromatin immunoprecipitation.

Identification of *Prox* expression in rhabdomeric photoreceptors of amphioxus led to molecular cloning and of an orthologous gene from a lophotrochozoan *P. dumerilii*. The co-expression of Prox with r-opsin in *Platynereis* and amphioxus brought a new piece of evidence to the proposal of Prox as an evolutionarily conserved marker of rhabdomeric photoreceptors.

The molecular data gathered in this work together with available literature data led to the proposal of a model for evolution of transcriptional regulation of opsin genes. The data and tools generated in this work and will be not only relevant for the proposed model, but also in a more general sense for illuminating the gene regulatory networks underlying the development of amphioxus photoreceptive structures and therefore may provide a substantial insight into the origins of the vertebrate eye.

7 References

ACAMPORA, D., ANNINO, A., TUORTO, F., PUELLES, E., LUCCHESI, W., PAPALIA, A. & SIMEONE, A. (2005) Otx genes in the evolution of the vertebrate brain. *Brain Res Bull,* 66**,** 410-20.

ACAMPORA, D., AVANTAGGIATO, V., TUORTO, F., BARONE, P., REICHERT, H., FINKELSTEIN, R. & SIMEONE, A. (1998) Murine Otx1 and Drosophila otd genes share conserved genetic functions required in invertebrate and vertebrate brain development. *Development,* 125**,** 1691-702.

APPELBAUM, L. & GOTHILF, Y. (2006) Mechanism of pineal-specific gene expression: the role of E-box and photoreceptor conserved elements. *Mol Cell Endocrinol,* 252**,** 27-33.

ARENAS-MENA, C. & WONG, K. S. (2007) HeOtx expression in an indirectly developing polychaete correlates with gastrulation by invagination. *Dev Genes Evol,* 217**,** 373-84.

ARENDT, D. (2003) Evolution of eyes and photoreceptor cell types. *Int J Dev Biol,* 47**,** 563-71.

ARENDT, D. (2008) The evolution of cell types in animals: emerging principles from molecular studies. *Nat Rev Genet,* 9**,** 868-82.

ARENDT, D., TESSMAR-RAIBLE, K., SNYMAN, H., DORRESTEIJN, A. W. & WITTBRODT, J. (2004) Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science,* 306**,** 869-71.

ARENDT, D., TESSMAR, K., DE CAMPOS-BAPTISTA, M. I., DORRESTEIJN, A. & WITTBRODT, J. (2002) Development of pigment-cup eyes in the polychaete Platynereis dumerilii and evolutionary conservation of larval eyes in Bilateria. *Development,* 129**,** 1143-54.

ARENDT, D. & WITTBRODT, J. (2001) Reconstructing the eyes of Urbilateria. *Philos Trans R Soc Lond B Biol Sci,* 356**,** 1545-63.

AUSUBEL, F. M., BRENT, R., KINGSTON, R. E., MOORE, D. D., SEIDMAN, J. G., SMITH, J. A. & STRUHL, K. (Eds.) (1995) *Current Protocols in Molecular Biology,* USA, John Wiley & Sons, Inc.

BAILEY, T. J., EL-HODIRI, H., ZHANG, L., SHAH, R., MATHERS, P. H. & JAMRICH, M. (2004) Regulation of vertebrate eye development by Rx genes. *Int J Dev Biol,* 48**,** 761-70.

BELLINGHAM, J., WHITMORE, D., PHILP, A. R., WELLS, D. J. & FOSTER, R. G. (2002) Zebrafish melanopsin: isolation, tissue localisation and phylogenetic position. *Brain Res Mol Brain Res,* 107**,** 128-36.

BI, X., KAJAVA, A. V., JONES, T., DEMIDENKO, Z. N. & MORTIN, M. A. (2003) The carboxy terminus of Prospero regulates its subcellular localization. *Mol Cell Biol,* 23**,** 1014-24.

BIESZKE, J. A., BRAUN, E. L., BEAN, L. E., KANG, S., NATVIG, D. O. & BORKOVICH, K. A. (1999) The nop-1 gene of Neurospora crassa encodes a seven transmembrane helix retinal-binding protein homologous to archaeal rhodopsins. *Proc Natl Acad Sci U S A,* 96**,** 8034-9.

BLACKBURN, D. C., CONLEY, K. W., PLACHETZKI, D. C., KEMPLER, K., BATTELLE, B. A. & BROWN, N. L. (2008) Isolation and expression of Pax6 and atonal homologues in the American horseshoe crab, Limulus polyphemus. *Dev Dyn,* 237**,** 2209-19.

BLANCK, A. & OESTERHELT, D. (1987) The halo-opsin gene. II. Sequence, primary structure of halorhodopsin and comparison with bacteriorhodopsin. *EMBO J,* 6**,** 265-73.

BOVOLENTA, P., MALLAMACI, A., BRIATA, P., CORTE, G. & BONCINELLI, E. (1997) Implication of OTX2 in pigment epithelium determination and neural retina differentiation. *J Neurosci,* 17**,** 4243-52.

CONKLIN, E. G. (1932) The embryology of amphioxus. *Journal of Morphology,* 54**,** 69-151.

CONTIN, M. A., VERRA, D. M. & GUIDO, M. E. (2006) An invertebrate-like phototransduction cascade mediates light detection in the chicken retinal ganglion cells. *FASEB J,* 20**,** 2648-50.

COOK, T., PICHAUD, F., SONNEVILLE, R., PAPATSENKO, D. & DESPLAN, C. (2003) Distinction between color photoreceptor cell fates is controlled by Prospero in Drosophila. *Dev Cell,* 4**,** 853-64.

CROZIER, W. J. (1916) The photoreceptors of amphioxus. *Anat Rec,* 11**,** 520.

CUI, W., TOMAREV, S. I., PIATIGORSKY, J., CHEPELINSKY, A. B. & DUNCAN, M. K. (2004) Mafs, Prox1, and Pax6 can regulate chicken betaB1-crystallin gene expression. *J Biol Chem,* 279**,** 11088-95.

D'ANIELLO, S., D'ANIELLO, E., LOCASCIO, A., MEMOLI, A., CORRADO, M., RUSSO, M. T., ANIELLO, F., FUCCI, L., BROWN, E. R. & BRANNO, M. (2006) The ascidian homolog of the vertebrate homeobox gene Rx is essential for ocellus development and function. *Differentiation,* 74**,** 222-34.

DARWIN, C. (1859) *On the Origin of Species by Means of Natural Selection,* London, Murray.

DAVIDSON, E. H. (2006) *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution,* Burlington, MA, Academic Press (Elsevier).

DAVIS, R. J., TAVSANLI, B. C., DITTRICH, C., WALLDORF, U. & MARDON, G. (2003) Drosophila retinal homeobox (drx) is not required for establishment of the visual system, but is required for brain and clypeus development. *Dev Biol,* 259**,** 272-87.

DELSUC, F., BRINKMANN, H., CHOURROUT, D. & PHILIPPE, H. (2006) Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature,* 439**,** 965-8.

DOE, C. Q., CHU-LAGRAFF, Q., WRIGHT, D. M. & SCOTT, M. P. (1991) The prospero gene specifies cell fates in the Drosophila central nervous system. *Cell,* 65**,** 451-64.

DUNCAN, M. K., CUI, W., OH, D. J. & TOMAREV, S. I. (2002) Prox1 is differentially localized during lens development. *Mech Dev,* 112**,** 195-8.

DYER, M. A., LIVESEY, F. J., CEPKO, C. L. & OLIVER, G. (2003) Prox1 function controls progenitor cell proliferation and horizontal cell genesis in the mammalian retina. *Nat Genet,* 34**,** 53-8.

EAKIN, R. E. & WESTFALL, J. A. (1962) Fine Structure of Photoreceptors in Amphioxus. *J Ultrastructure Research,* 6**,** 531-532.

EAKIN, R. M. (1979) Evolutionary Significance of Photoreceptors - Retrospect. *American Zoologist,* 19**,** 647-653.

EBERHARD, D., JIMENEZ, G., HEAVEY, B. & BUSSLINGER, M. (2000) Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *EMBO J,* 19**,** 2292-303.

EDGAR, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res,* 32**,** 1792-7.

FELSENSTEIN, J. (2005) PHYLIP (Phylogeny Inference Package). 3.6 Distributed by the author ed., Department of Genome Sciences, University of Washington, Seattle.

FINKELSTEIN, R., SMOUSE, D., CAPACI, T. M., SPRADLING, A. C. & PERRIMON, N. (1990) The orthodenticle gene encodes a novel homeo domain protein involved in the development of the Drosophila nervous system and ocellar visual structures. *Genes Dev,* 4**,** 1516-27.

FRANZ, V. (1927) Morphologie der Akranier. *Ergebnisse der Anaomie und Entwicklungsgeschichte,* 27**,** 464-692.

FRITZSCH, B. (1996) Similarities and Differences in Lancelet and Craniate Nervous Systems. *Israel Journal of Zoology,* 42**,** S-147-160.

FURUKAWA, T., MORROW, E. M. & CEPKO, C. L. (1997) Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell,* 91**,** 531-41.

GARCIA-FERNANDEZ, J., JIMENEZ-DELGADO, S., PASCUAL-ANAYA, J., MAESO, I., IRIMIA, M., MINGUILLON, C., BENITO-GUTIERREZ, E., GARDENYES, J., BERTRAND, S. & D'ANIELLO, S. (2008) From the American to the European amphioxus: towards experimental Evo-Devo at the origin of chordates. *Int J Dev Biol*.

GEHRING, W. (2000) Reply to meyer-rochow. *Trends Genet,* 16**,** 245.

GEHRING, W. J. & IKEO, K. (1999) Pax 6: mastering eye morphogenesis and eye evolution. *Trends Genet,* 15**,** 371-7.

GERMOT, A., LECOINTRE, G., PLOUHINEC, J. L., LE MENTEC, C., GIRARDOT, F. & MAZAN, S. (2001) Structural evolution of Otx genes in craniates. *Mol Biol Evol,* 18**,** 1668-78.

GLARDON, S., HOLLAND, L. Z., GEHRING, W. J. & HOLLAND, N. D. (1998) Isolation and developmental expression of the amphioxus Pax-6 gene (AmphiPax-6): insights into eye and photoreceptor evolution. *Development,* 125**,** 2701-10.

GRAHAM, D. M., WONG, K. Y., SHAPIRO, P., FREDERICK, C., PATTABIRAMAN, K. & BERSON, D. M. (2008) Melanopsin ganglion cells use a membrane-associated rhabdomeric phototransduction cascade. *J Neurophysiol,* 99**,** 2522-32.

GUTHRIE, D. M. (1975) The Physiology and Structure of the Nervous System of Amphioxus (the Lancelet), Branchiostoma lanceolatum Pallas. IN BARRINGTON, E. J. W. & JEFFERIES, R. P. S. (Eds.) *Protochordates.* London, Symp. Zool. Soc. Lond.

HALDER, G., CALLAERTS, P. & GEHRING, W. J. (1995) Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila. *Science,* 267**,** 1788-92.

HALL, T. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* , 41**,** 4.

HANES, S. D. & BRENT, R. (1991) A genetic model for interaction of the homeodomain recognition helix with DNA. *Science,* 251**,** 426-30.

HASSAN, B., LI, L., BREMER, K. A., CHANG, W., PINSONNEAULT, J. & VAESSIN, H. (1997) Prospero is a panneural transcription factor that modulates homeodomain protein activity. *Proc Natl Acad Sci U S A,* 94**,** 10991-6.

HENNIG, A. K., PENG, G. H. & CHEN, S. (2008) Regulation of photoreceptor gene expression by Crx-associated transcription factor network. *Brain Res,* 1192**,** 114-33.

HESSE, R. (1898) The Eyes of Amphioxus. *Amer Nat,* 32**,** 447-448.

HILL, R. E., FAVOR, J., HOGAN, B. L., TON, C. C., SAUNDERS, G. F., HANSON, I. M., PROSSER, J., JORDAN, T., HASTIE, N. D. & VAN HEYNINGEN, V.

(1991) Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature,* 354**,** 522-5.

HOLLAND, L. Z., ALBALAT, R., AZUMI, K., BENITO-GUTIERREZ, E., BLOW, M. J., BRONNER-FRASER, M., BRUNET, F., BUTTS, T., CANDIANI, S., DISHAW, L. J., FERRIER, D. E., GARCIA-FERNANDEZ, J., GIBSON-BROWN, J. J., GISSI, C., GODZIK, A., HALLBOOK, F., HIROSE, D., HOSOMICHI, K., IKUTA, T., INOKO, H., KASAHARA, M., KASAMATSU, J., KAWASHIMA, T., KIMURA, A., KOBAYASHI, M., KOZMIK, Z., KUBOKAWA, K., LAUDET, V., LITMAN, G. W., MCHARDY, A. C., MEULEMANS, D., NONAKA, M., OLINSKI, R. P., PANCER, Z., PENNACCHIO, L. A., PESTARINO, M., RAST, J. P., RIGOUTSOS, I., ROBINSON-RECHAVI, M., ROCH, G., SAIGA, H., SASAKURA, Y., SATAKE, M., SATOU, Y., SCHUBERT, M., SHERWOOD, N., SHIINA, T., TAKATORI, N., TELLO, J., VOPALENSKY, P., WADA, S., XU, A., YE, Y., YOSHIDA, K., YOSHIZAKI, F., YU, J. K., ZHANG, Q., ZMASEK, C. M., DE JONG, P. J., OSOEGAWA, K., PUTNAM, N. H., ROKHSAR, D. S., SATOH, N. & HOLLAND, P. W. (2008) The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res,* 18**,** 1100-11.

HOLLAND, L. Z., HOLLAND, P. W. & HOLLAND, N. D. (1996) Whole mount in situ hybridization applicable to Amphioxus and other small larvae. IN FERRARIS, J. D. & PALUMBI, S. R. (Eds.) *Molecular Zoology, Advances, Strategies, and Protocols.* Wiley-Liss Inc.

HOLLAND, L. Z. & YU, J. K. (2004) Cephalochordate (amphioxus) embryos: procurement, culture, and basic methods. *Methods Cell Biol,* 74**,** 195-215.

HOLLAND, N. D. & HOLLAND, L. Z. (2006) Stage- and tissue-specific patterns of cell division in embryonic and larval tissues of amphioxus during normal development. *Evol Dev,* 8**,** 142-9.

CHEN, C. M. & CEPKO, C. L. (2002) The chicken RaxL gene plays a role in the initiation of photoreceptor differentiation. *Development,* 129**,** 5363-75.

CHEN, S., WANG, Q. L., NIE, Z., SUN, H., LENNON, G., COPELAND, N. G., GILBERT, D. J., JENKINS, N. A. & ZACK, D. J. (1997) Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron,* 19**,** 1017-30.

CHOW, R. L., ALTMANN, C. R., LANG, R. A. & HEMMATI-BRIVANLOU, A. (1999) Pax6 induces ectopic eyes in a vertebrate. *Development,* 126**,** 4213-22.

KASAHARA, T., OKANO, T., HAGA, T. & FUKADA, Y. (2002) Opsin-G11-mediated signaling pathway for photic entrainment of the chicken pineal circadian clock. *J Neurosci,* 22**,** 7321-5.

KLINGLER, M. (2004) Tribolium. *Curr Biol,* 14**,** R639-40.

KOIKE, C., NISHIDA, A., UENO, S., SAITO, H., SANUKI, R., SATO, S., FURUKAWA, A., AIZAWA, S., MATSUO, I., SUZUKI, N., KONDO, M. & FURUKAWA, T. (2007) Functional roles of Otx2 transcription factor in postnatal mouse retinal development. *Mol Cell Biol,* 27**,** 8318-29.

KOJIMA, D., TERAKITA, A., ISHIKAWA, T., TSUKAHARA, Y., MAEDA, A. & SHICHIDA, Y. (1997) A novel Go-mediated phototransduction cascade in scallop visual cells. *J Biol Chem,* 272**,** 22979-82.

KOYANAGI, M., KUBOKAWA, K., TSUKAMOTO, H., SHICHIDA, Y. & TERAKITA, A. (2005) Cephalochordate melanopsin: evolutionary linkage between invertebrate visual cells and vertebrate photosensitive retinal ganglion cells. *Curr Biol,* 15**,** 1065-9.

KOYANAGI, M., TAKANO, K., TSUKAMOTO, H., OHTSU, K., TOKUNAGA, F. & TERAKITA, A. (2008) Jellyfish vision starts with cAMP signaling mediated by opsin-Gs cascade. *Proc Natl Acad Sci U S A*.

KOYANAGI, M., TERAKITA, A., KUBOKAWA, K. & SHICHIDA, Y. (2002) Amphioxus homologs of Go-coupled rhodopsin and peropsin having 11-cis- and all-trans-retinals as their chromophores. *FEBS Lett,* 531**,** 525-8.

KOZMIK, Z. (2005) Pax genes in eye development and evolution. *Curr Opin Genet Dev,* 15**,** 430-8.

KOZMIK, Z., HOLLAND, N. D., KALOUSOVA, A., PACES, J., SCHUBERT, M. & HOLLAND, L. Z. (1999) Characterization of an amphioxus paired box gene, AmphiPax2/5/8: developmental expression patterns in optic support cells, nephridium, thyroid-like structures and pharyngeal gill slits, but not in the midbrain-hindbrain boundary region. *Development,* 126**,** 1295-304.

KOZMIK, Z., HOLLAND, N. D., KRESLOVA, J., OLIVERI, D., SCHUBERT, M., JONASOVA, K., HOLLAND, L. Z., PESTARINO, M., BENES, V. & CANDIANI, S. (2007) Pax-Six-Eya-Dach network during amphioxus development: conservation in vitro but context specificity in vivo. *Dev Biol,* 306**,** 143-59.

KOZMIK, Z., RUZICKOVA, J., JONASOVA, K., MATSUMOTO, Y., VOPALENSKY, P., KOZMIKOVA, I., STRNAD, H., KAWAMURA, S., PIATIGORSKY, J., PACES, V. & VLCEK, C. (2008) Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proc Natl Acad Sci U S A,* 105**,** 8989-93.

KRAUSE, W. (1888) Die Retina. II. Die Retina der Fische. *Internationale Monatschrift fur Anatomie un Physilologie,* 5**,** 132-148.

LACALLI, T., HOLLAND, N. & WEST, J. (1994) Landmarks in the Anterior Central Nervous System of Amphioxus Larvae. *Philosophical Transactions: Biological Sciences,* 344**,** 165-185.

LACALLI, T. C. (1994) Apical Organs, Epithelial Domains, and the Origin of the Chordate Central-Nervous-System. *American Zoologist,* 34**,** 533-541.

LACALLI, T. C. (1996) Frontal eye circuitry, rostral sensory pathways and brain organization in amphioxus larvae: Evidence from 3D reconstructions. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* 351**,** 243-263.

LACALLI, T. C. (2004) Sensory systems in amphioxus: a window on the ancestral chordate condition. *Brain Behav Evol,* 64**,** 148-62.

LACALLI, T. C. (2008) Basic features of the ancestral chordate brain: a protochordate perspective. *Brain Res Bull,* 75**,** 319-23.

LACALLI, T. C. & KELLY, S. J. (2000) The infundibular balance organ in amphioxus larvae and related aspects of cerebral vesicle organization. *Acta Zoologica,* 81**,** 37-47.

LAMB, T. D., COLLIN, S. P. & PUGH, E. N., JR. (2007) Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. *Nat Rev Neurosci,* 8**,** 960-76.

LAND, M. & NILSSON, D.-E. (2002) *Animal Eyes,* New York, Oxford University Press.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LARROUX, C., LUKE, G. N., KOOPMAN, P., ROKHSAR, D. S., SHIMELD, S. M. & DEGNAN, B. M. (2008) Genesis and expansion of metazoan transcription factor gene classes. *Mol Biol Evol,* 25**,** 980-96.

LARUSSO ND, R., BE, SINGH AK, OAKLEY TH (2008) Type II Opsins: Evolutionary Origin by Internal Domain Duplication? *J Mol Evol,* 66**,** 417–423.

LENGLER, J., KRAUSZ, E., TOMAREV, S., PRESCOTT, A., QUINLAN, R. A. & GRAW, J. (2001) Antagonistic action of Six3 and Prox1 at the gamma-crystallin promoter. *Nucleic Acids Res,* 29**,** 515-26.

LOOSLI, F., STAUB, W., FINGER-BAIER, K. C., OBER, E. A., VERKADE, H., WITTBRODT, J. & BAIER, H. (2003) Loss of eyes in zebrafish caused by mutation of chokh/rx3. *EMBO Rep,* 4**,** 894-9.

MARKITANTOVA, Y. V., MAKARIEV, E. O., PAVLOVA, G. V., ZINOVIEVA, R. D. & MITASHOV, V. I. (2003) Location of the Prox1 gene expression during newt lens and retina regeneration. *Dokl Biol Sci,* 391**,** 361-4.

MARTINEZ-MORALES, J. R., DOLEZ, V., RODRIGO, I., ZACCARINI, R., LECONTE, L., BOVOLENTA, P. & SAULE, S. (2003) OTX2 activates the molecular network underlying retina pigment epithelium differentiation. *J Biol Chem,* 278**,** 21721-31.

MARTINEZ-MORALES, J. R., SIGNORE, M., ACAMPORA, D., SIMEONE, A. & BOVOLENTA, P. (2001) Otx genes are required for tissue specification in the developing eye. *Development,* 128**,** 2019-30.

MATHERS, P. H., GRINBERG, A., MAHON, K. A. & JAMRICH, M. (1997) The Rx homeobox gene is essential for vertebrate eye development. *Nature,* 387**,** 603-7.

MATSUZAKI, F., KOIZUMI, K., HAMA, C., YOSHIOKA, T. & NABESHIMA, Y. (1992) Cloning of the Drosophila prospero gene and its expression in ganglion mother cells. *Biochem Biophys Res Commun,* 182**,** 1326-32.

MEYER-ROCHOW, V. B. (2000) The eye: monophyletic, polyphyletic or perhaps biphyletic? *Trends Genet,* 16**,** 244-5.

MILLIGAN, G. & KOSTENIS, E. (2006) Heterotrimeric G-proteins: a short history. *Br J Pharmacol,* 147 Suppl 1**,** S46-55.

NAGAO, T., LEUZINGER, S., ACAMPORA, D., SIMEONE, A., FINKELSTEIN, R., REICHERT, H. & FURUKUBO-TOKUNAGA, K. (1998) Developmental rescue of Drosophila cephalic defects by the human Otx genes. *Proc Natl Acad Sci U S A,* 95**,** 3737-42.

NAGEL, G., OLLIG, D., FUHRMANN, M., KATERIYA, S., MUSTI, A. M., BAMBERG, E. & HEGEMANN, P. (2002) Channelrhodopsin-1: a light-gated proton channel in green algae. *Science,* 296**,** 2395-8.

NISHIDA, A., FURUKAWA, A., KOIKE, C., TANO, Y., AIZAWA, S., MATSUO, I. & FURUKAWA, T. (2003) Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. *Nat Neurosci,* 6**,** 1255-63.

NISHIJIMA, I. & OHTOSHI, A. (2006) Characterization of a novel prospero-related homeobox gene, Prox2. *Mol Genet Genomics,* 275**,** 471-8.

NOHARA, M., NISHIDA, M. & NISHIKAWA, T. (2005) New complete mitochondrial DNA sequence of the lancelet Branchiostoma lanceolatum (Cephalochordata) and the identity of this species' sequences. *Zoolog Sci,* 22**,** 671-4.

NORDSTROM, K., LARSSON, T. A. & LARHAMMAR, D. (2004) Extensive duplications of phototransduction genes in early vertebrate evolution correlate with block (chromosome) duplications. *Genomics,* 83**,** 852-72.

OHNO, S. (1970) *Evolution by gene duplication,* Berlin, Springer-Verlag.

OKA, Y., SARAIVA, L. R., KWAN, Y. Y. & KORSCHING, S. I. (2009) The fifth class of Galpha proteins. *Proc Natl Acad Sci U S A,* 106**,** 1484-9.

OLIVER, G., SOSA-PINEDA, B., GEISENDORF, S., SPANA, E. P., DOE, C. Q. & GRUSS, P. (1993) Prox 1, a prospero-related homeobox gene expressed during mouse development. *Mech Dev,* 44**,** 3-16.

PALLAS, P. (1774) Spicilegia zoologica, quibus novae imprimus et obscurae animalium species iconibus, descriptionibus atque comentariis illustrantur. *G.A.Lange Berlin,* I(10).

PAN, Y., NEKKALAPUDI, S., KELLY, L. E. & EL-HODIRI, H. M. (2006) The Rx-like homeobox gene (Rx-L) is necessary for normal photoreceptor development. *Invest Ophthalmol Vis Sci,* 47**,** 4245-53.

PANDA, S., NAYAK, S. K., CAMPO, B., WALKER, J. R., HOGENESCH, J. B. & JEGLA, T. (2005) Illumination of the melanopsin signaling pathway. *Science,* 307**,** 600-4.

PINEDA, D., ROSSI, L., BATISTONI, R., SALVETTI, A., MARSAL, M., GREMIGNI, V., FALLENI, A., GONZALEZ-LINARES, J., DERI, P. & SALO, E. (2002) The genetic network of prototypic planarian eye regeneration is Pax6 independent. *Development,* 129**,** 1423-34.

PISTOCCHI, A., BARTESAGHI, S., COTELLI, F. & DEL GIACCO, L. (2008a) Identification and expression pattern of zebrafish prox2 during embryonic development. *Dev Dyn,* 237**,** 3916-20.

PISTOCCHI, A., GAUDENZI, G., CARRA, S., BRESCIANI, E., DEL GIACCO, L. & COTELLI, F. (2008b) Crucial role of zebrafish prox1 in hypothalamic catecholaminergic neurons development. *BMC Dev Biol,* 8**,** 27.

PLACHETZKI, D. C., DEGNAN, B. M. & OAKLEY, T. H. (2007) The origins of novel protein interactions during animal opsin evolution. *PLoS ONE,* 2**,** e1054.

PLACHETZKI, D. C. & OAKLEY, T. H. (2007) Key transitions during the evolution of animal phototransduction: novelty, "tree-thinking," co-option, and co-duplication. *Integrative and Comparative Biology,* 47**,** 759-769.

PLOUHINEC, J. L., LECONTE, L., SAUKA-SPENGLER, T., BOVOLENTA, P., MAZAN, S. & SAULE, S. (2005) Comparative analysis of gnathostome Otx gene expression patterns in the developing eye: implications for the functional evolution of the multigene family. *Dev Biol,* 278**,** 560-75.

PLOUHINEC, J. L., SAUKA-SPENGLER, T., GERMOT, A., LE MENTEC, C., CABANA, T., HARRISON, G., PIEAU, C., SIRE, J. Y., VERON, G. & MAZAN, S. (2003) The mammalian Crx genes are highly divergent representatives of the Otx5 gene family, a gnathostome orthology class of orthodenticle-related homeogenes involved in the differentiation of retinal photoreceptors and circadian entrainment. *Mol Biol Evol,* 20**,** 513-21.

PROVENCIO, I., RODRIGUEZ, I. R., JIANG, G., HAYES, W. P., MOREIRA, E. F. & ROLLAG, M. D. (2000) A novel human opsin in the inner retina. *J Neurosci,* 20**,** 600-5.

PROVENCIO, I., ROLLAG, M. D. & CASTRUCCI, A. M. (2002) Photoreceptive net in the mammalian retina. This mesh of cells may explain how some blind mice can still tell day from night. *Nature,* 415**,** 493.

PUTNAM, N. H., BUTTS, T., FERRIER, D. E., FURLONG, R. F., HELLSTEN, U., KAWASHIMA, T., ROBINSON-RECHAVI, M., SHOGUCHI, E., TERRY, A., YU, J. K., BENITO-GUTIERREZ, E. L., DUBCHAK, I., GARCIA-FERNANDEZ, J., GIBSON-BROWN, J. J., GRIGORIEV, I. V., HORTON, A. C., DE JONG, P. J., JURKA, J., KAPITONOV, V. V., KOHARA, Y., KUROKI, Y., LINDQUIST, E., LUCAS, S., OSOEGAWA, K., PENNACCHIO, L. A., SALAMOV, A. A., SATOU, Y., SAUKA-SPENGLER, T., SCHMUTZ, J., SHIN,

I. T., TOYODA, A., BRONNER-FRASER, M., FUJIYAMA, A., HOLLAND, L. Z., HOLLAND, P. W., SATOH, N. & ROKHSAR, D. S. (2008) The amphioxus genome and the evolution of the chordate karyotype. *Nature,* 453**,** 1064-71.

QIAN, J., ESUMI, N., CHEN, Y., WANG, Q., CHOWERS, I. & ZACK, D. J. (2005) Identification of regulatory targets of tissue-specific transcription factors: application to retina-specific gene regulation. *Nucleic Acids Res,* 33**,** 3479-91.

QUIRING, R., WALLDORF, U., KLOTER, U. & GEHRING, W. J. (1994) Homology of the eyeless gene of Drosophila to the Small eye gene in mice and Aniridia in humans. *Science,* 265**,** 785-9.

RADU, R. A., HU, J., PENG, J., BOK, D., MATA, N. L. & TRAVIS, G. H. (2008) Retinal pigment epithelium-retinal G protein receptor-opsin mediates light-dependent translocation of all-trans-retinyl esters for synthesis of visual chromophore in retinal pigment epithelial cells. *J Biol Chem,* 283**,** 19730-8.

RAIBLE, F. & ARENDT, D. (2004) Metazoan evolution: some animals are more equal than others. *Curr Biol,* 14**,** R106-8.

RAIBLE, F., TESSMAR-RAIBLE, K., ARBOLEDA, E., KALLER, T., BORK, P., ARENDT, D. & ARNONE, M. I. (2006) Opsins and clusters of sensory G-protein-coupled receptors in the sea urchin genome. *Dev Biol,* 300**,** 461-75.

RANADE, S. S., YANG-ZHOU, D., KONG, S. W., MCDONALD, E. C., COOK, T. A. & PIGNONI, F. (2008) Analysis of the Otd-dependent transcriptome supports the evolutionary conservation of CRX/OTX/OTD functions in flies and vertebrates. *Dev Biol,* 315**,** 521-34.

RATH, M. F., BAILEY, M. J., KIM, J. S., COON, S. L., KLEIN, D. C. & MOLLER, M. (2009) Developmental and daily expression of the Pax4 and Pax6 homeobox genes in the rat retina: localization of Pax4 in photoreceptor cells. *J Neurochem,* 108**,** 285-94.

RATH, M. F., MORIN, F., SHI, Q., KLEIN, D. C. & MOLLER, M. (2007) Ontogenetic expression of the Otx2 and Crx homeobox genes in the retina of the rat. *Exp Eye Res,* 85**,** 65-73.

ROYET, J. & FINKELSTEIN, R. (1995) Pattern formation in Drosophila head development: the role of the orthodenticle homeobox gene. *Development,* 121**,** 3561-72.

RUIZ, S. & ANADON, R. (1991) The fine structure of lamellate cells in the brain of amphioxus (Branchiostoma lanceolatum, Cephalochordata). *Cell and Tissue Research,* 263**,** 597-600.

RUPPERT, E. E. (1997) Cephalochordata (Acrania). IN HARRISON, F. W. & RUPPERT, E. E. (Eds.) *Microscopic Anatomy of Invertebrates.* New York, Wiley-Liss, Inc.

RYAN, J. F., BURTON, P. M., MAZZA, M. E., KWONG, G. K., MULLIKIN, J. C. & FINNERTY, J. R. (2006) The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes: evidence from the starlet sea anemone, Nematostella vectensis. *Genome Biol,* 7**,** R64.

RYTER, J. M., DOE, C. Q. & MATTHEWS, B. W. (2002) Structure of the DNA binding region of prospero reveals a novel homeo-prospero domain. *Structure,* 10**,** 1541-9.

SALO, E., PINEDA, D., MARSAL, M., GONZALEZ, J., GREMIGNI, V. & BATISTONI, R. (2002) Genetic network of the eye in Platyhelminthes: expression and functional analysis of some players during planarian regeneration. *Gene,* 287**,** 67-74.

SATIR, P. (2000) A comment on the origin of the vertebrate eye. *Anat Rec,* 261**,** 224-7.

SIMEONE, A., PUELLES, E. & ACAMPORA, D. (2002) The Otx family. *Curr Opin Genet Dev,* 12**,** 409-15.

SPUDICH, J. L., YANG, C. S., JUNG, K. H. & SPUDICH, E. N. (2000) Retinylidene proteins: structures and functions from archaea to humans. *Annu Rev Cell Dev Biol,* 16**,** 365-92.

STOKES, M. (1997) Larval locomotion of the lancelet. *J Exp Biol,* 200**,** 1661-80.

SU, C. Y., LUO, D. G., TERAKITA, A., SHICHIDA, Y., LIAO, H. W., KAZMI, M. A., SAKMAR, T. P. & YAU, K. W. (2006) Parietal-eye phototransduction components and their potential evolutionary implications. *Science,* 311**,** 1617-21.

SUGA, H., SCHMID, V. & GEHRING, W. J. (2008) Evolution and functional diversity of jellyfish opsins. *Curr Biol,* 18**,** 51-5.

SUZUKI, M. M. & SATOH, N. (2000) Genes expressed in the amphioxus notochord revealed by EST analysis. *Dev Biol,* 224**,** 168-77.

SUZUKI, T., TERAKITA, A., NARITA, K., NAGAI, K., TSUKAHARA, Y. & KITO, Y. (1995) Squid photoreceptor phospholipase C is stimulated by membrane Gq alpha but not by soluble Gq alpha. *FEBS Lett,* 377**,** 333-7.

TAHAYATO, A., SONNEVILLE, R., PICHAUD, F., WERNET, M. F., PAPATSENKO, D., BEAUFILS, P., COOK, T. & DESPLAN, C. (2003) Otd/Crx, a dual regulator for the specification of ommatidia subtypes in the Drosophila retina. *Dev Cell,* 5**,** 391-402.

TAKEDA, K., YOKOYAMA, S., YASUMOTO, K., SAITO, H., UDONO, T., TAKAHASHI, K. & SHIBAHARA, S. (2003) OTX2 regulates expression of DOPAchrome tautomerase in human retinal pigment epithelium. *Biochem Biophys Res Commun,* 300**,** 908-14.

TAKECHI, M., SENO, S. & KAWAMURA, S. (2008) Identification of cis-acting elements repressing blue opsin expression in zebrafish UV cones and pineal cells. *J Biol Chem*.

TARTTELIN, E. E., BELLINGHAM, J., HANKINS, M. W., FOSTER, R. G. & LUCAS, R. J. (2003) Neuropsin (Opn5): a novel opsin identified in mammalian neural tissue. *FEBS Lett,* 554**,** 410-6.

TENBAUM, E. (1955) Polarisationsoptische Beiträge zur Kenntnis der Gewebe von Branchiostoma Lanceolatum (P). *Zeitschrift für Zellforschung,* 42**,** 149-192.

TERAKITA, A. (2005) The opsins. *Genome Biol,* 6**,** 213.

TERIO, B. (1964) Intorno ai fotorecettori dell´anfiosso. *Atti della Societa Peloritana di Science Fisiche, Matematche e Naturali,* 10**,** 111-125.

TESSMAR-RAIBLE, K. & ARENDT, D. (2003) Emerging systems: between vertebrates and arthropods, the Lophotrochozoa. *Curr Opin Genet Dev,* 13**,** 331-40.

TOMAREV, S. I., SUNDIN, O., BANERJEE-BASU, S., DUNCAN, M. K., YANG, J. M. & PIATIGORSKY, J. (1996) Chicken homeobox gene Prox 1 related to Drosophila prospero is expressed in the developing lens and retina. *Dev Dyn,* 206**,** 354-67.

TOMONARI, S., TAKAGI, A., AKAMATSU, S., NOJI, S. & OHUCHI, H. (2005) A non-canonical photopigment, melanopsin, is expressed in the differentiating ganglion, horizontal, and bipolar cells of the chicken retina. *Dev Dyn,* 234**,** 783-90.

UMESONO, Y., WATANABE, K. & AGATA, K. (1999) Distinct structural domains in the planarian brain defined by the expression of evolutionarily conserved homeobox genes. *Dev Genes Evol,* 209**,** 31-9.

VANDENDRIES, E. R., JOHNSON, D. & REINKE, R. (1996) orthodenticle is required for photoreceptor cell development in the Drosophila eye. *Dev Biol,* 173**,** 243-55.

VELARDE, R. A., SAUER, C. D., WALDEN, K. K., FAHRBACH, S. E. & ROBERTSON, H. M. (2005) Pteropsin: a vertebrate-like non-visual opsin expressed in the honey bee brain. *Insect Biochem Mol Biol,* 35**,** 1367-77.

VOPALENSKY, P. & KOZMIK, Z. (2009) Eye evolution: Common use and independent recruitment of genetic components. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* accepted.

WADA, S., KATSUYAMA, Y., SATO, Y., ITOH, C. & SAIGA, H. (1996) Hroth an orthodenticle-related homeobox gene of the ascidian, Halocynthia roretzi: its expression and putative roles in the axis formation during embryogenesis. *Mech Dev,* 60**,** 59-71.

WADA, S., TOKUOKA, M., SHOGUCHI, E., KOBAYASHI, K., DI GREGORIO, A., SPAGNUOLO, A., BRANNO, M., KOHARA, Y., ROKHSAR, D., LEVINE, M., SAIGA, H., SATOH, N. & SATOU, Y. (2003) A genomewide survey of developmentally relevant genes in Ciona intestinalis. II. Genes for homeobox transcription factors. *Dev Genes Evol,* 213**,** 222-34.

WADA, S., TOYODA, R., YAMAMOTO, H. & SAIGA, H. (2002) Ascidian otx gene Hroth activates transcription of the brain-specific gene HrTRP. *Dev Dyn,* 225**,** 46-53.

WANG, L., WANG, S., LI, Y., PARADESI, M. S. & BROWN, S. J. (2007) BeetleBase: the model organism database for Tribolium castaneum. *Nucleic Acids Res,* 35**,** D476-9.

WANG, Q. L., CHEN, S., ESUMI, N., SWAIN, P. K., HAINES, H. S., PENG, G., MELIA, B. M., MCINTOSH, I., HECKENLIVELY, J. R., JACOBSON, S. G., STONE, E. M., SWAROOP, A. & ZACK, D. J. (2004) QRX, a novel homeobox gene, modulates photoreceptor gene expression. *Hum Mol Genet,* 13**,** 1025-40.

WANG, W., XU, H. L., LIN, L. P., SU, B. & WANG, Y. Q. (2005) Construction of a BAC library for Chinese amphioxus Branchiostoma belcheri and identification of clones containing Amphi-Pax genes. *Genes Genet Syst,* 80**,** 233-6.

WICHT, H. & LACALLI, T. C. (2005) The nervous system of amphioxus: structure, development, and evolutionary significance. *Canadian Journal of Zoology-Revue Canadienne De Zoologie,* 83**,** 122-150.

WILLIAMS, N. A. & HOLLAND, P. W. (1998) Gene and domain duplication in the chordate Otx gene family: insights from amphioxus Otx. *Mol Biol Evol,* 15**,** 600-7.

WILLIAMS, N. A. & HOLLAND, P. W. H. (1996) Old head on young shoulders. *Nature,* 383**,** 490-490.

WU, H. Y., PERRON, M. & HOLLEMANN, T. (2009) The role of Xenopus Rx-L in photoreceptor cell determination. *Dev Biol,* 327**,** 352-65.

YOSHIDA, R., KUSAKABE, T., KAMATANI, M., DAITOH, M. & TSUDA, M. (2002) Central nervous system-specific expression of G protein alpha subunits in the ascidian Ciona intestinalis. *Zoolog Sci,* 19**,** 1079-88.

YOUSEF, M. S. & MATTHEWS, B. W. (2005) Structural basis of Prospero-DNA interaction: implications for transcription regulation in developing cells. *Structure,* 13**,** 601-7.

YU, J. K., HOLLAND, N. D. & HOLLAND, L. Z. (2004) Tissue-specific expression of FoxD reporter constructs in amphioxus embryos. *Dev Biol,* 274**,** 452-61.

YU, J. K., MEULEMANS, D., MCKEOWN, S. J. & BRONNER-FRASER, M. (2008) Insights from the amphioxus genome on the origin of vertebrate neural crest. *Genome Res,* 18**,** 1127-32.

ZHANG, Q. J., SUN, Y., ZHONG, J., LI, G., LU, X. M. & WANG, Y. Q. (2007) Continuous culture of two lancelets and production of the second filial generations in the laboratory. *J Exp Zoolog B Mol Dev Evol,* 308**,** 464-72.

8 Appendices

8.1 Abbreviations

10x g – relative centrifuge force corresponding to 10x normal gravity acceleration on Earth surface

BSA – bovine serum albumin

CDS – coding sequence

CNS – central nervous system

EST – expressed sequence tag

IHC – immunohistochemistry

ipRGC – intrinsically photosensitive retinal ganglion cells

IPTG – Isopropyl β-D-1-thiogalactopyranoside

LB medium – Luria-Bertani medium

ML – maximum likelihood method for phylogeny inference

NJ – Nejghbor-Joining method for phylogeny inference

OD_{600} – optical density measured at 600 nm

PBS – phosphate buffered saline

PFA – paraformaldehyde

RPM – rotations per minute

RGCs – retinal ganglion cells

TF – transcription factor

UTR – untranslated region

WB – western blot

8.2 Oligonucleotides

B. floridae opsins and *AmphiRx* – RT-PCR

Organism	"Genes"	Orient.	Name	Sequence (5´-3´)	Purpose
B. floridae	87094	F1	1-PV-RT	CGTCACCAGAAGTCTTAGTC	RT-PCR
B. floridae	87094-110002	R3	2-PV-RT	TCCGAACGCACCAATCAC	RT-PCR
B. floridae	110002	F1	3-PV-RT	ATCGTCACTAGGAGTCTCAG	RT-PCR
B. floridae	110003	F1	4-PV-RT	CATGTGCATCGTTTACTGTC	RT-PCR
B. floridae	110003	R	5-PV-RT	GCTTGGTGAGTATGGTGA	RT-PCR
B. floridae	175449-201585	F1	6-PV-RT	CACTACCTGGTGCTGAAC	RT-PCR
B. floridae	175449-201585	R3	7-PV-RT	TCATCACACTGACACTGAAG	RT-PCR
B. floridae	86640-86644	F3	8-PV-RT	CTTCGCACACTCATGGAC	RT-PCR
B. floridae	86640-86644	R3	9-PV-RT	CCACTGTGATCTAGGATGA	RT-PCR
B. floridae	65959-65960	F3	10-PV-RT	TCGGTCAGTTTGGAGAGT	RT-PCR
B. floridae	65959-65960	R3	11-PV-RT	TGGTCCTTCGTCGAGTTG	RT-PCR
B. floridae	86195-86253	F1	12-PV-RT	CCTGACGTTAGCAGTCTTCG	RT-PCR
B. floridae	86195-86253	R2	13-PV-RT	AGCCAATCAGAGCCACGATG	RT-PCR
B. floridae	84890-185357	F1	14-PV-RT	CATCATCTTCAGCGTCTTC	RT-PCR
B. floridae	84890-185357	R1	15-PV-RT	CTTGTCTTGGCGAACTTG	RT-PCR
B. floridae	74630-124039	F3	16-PV-RT	GCTCCTCTTACCTTACCTAC	RT-PCR
B. floridae	74630-124039	R4	17-PV-RT	GATGAAGATAGCCACGATG	RT-PCR
B. floridae	74631-84894	F1	18-PV-RT	TGTGTCCTACATCGTGAC	RT-PCR
B. floridae	74631-84894	R1	19-PV-RT	GACCATCGTAGTGACCA	RT-PCR
B. floridae	70447-206045	F1	20-PV-RT	CAACGACATCTCCTACATCA	RT-PCR
B. floridae	70447-206045	R2	21-PV-RT	AGGAAACACGTGATCAGCA	RT-PCR
B. floridae	206170-205982	F1	22-PV-RT	ACATCGTGGCCTACTTCGT	RT-PCR
B. floridae	206170-205982	R1	23-PV-RT	GGTCCAGCAGATCGTGAAC	RT-PCR
B. floridae	71561	F2	24-PV-RT	ATCCTGATGACGCTGTTCTA	RT-PCR
B. floridae	71561	R1	25-PV-RT	AGTGAAGATGTCGAGGTTG	RT-PCR
B. floridae	215180	F1	26-PV-RT	AGCATCAAGTGGACAGTAG	RT-PCR
B. floridae	215180	R1	27-PV-RT	CCACCATGAACAGTATCAC	RT-PCR
B. floridae	91094	F3	28-PV-RT	CGTCCTGAACGTCTCGAT	RT-PCR
B. floridae	91094-91095-91106	R1	29-PV-RT	CAGGATGGTGGCGATGAT	RT-PCR
B. floridae	91095-91106	F2	30-PV-RT	CTTCGGCTATGTCGTCTATG	RT-PCR
B. floridae	65045-73626	F1	31-PV-RT	CATCTCAGCGATCTACAG	RT-PCR
B. floridae	65045-73626	R1	32-PV-RT	GAGTGGTACTGCATCTAC	RT-PCR
B. floridae	210643	F1	33-PV-RT	GACGGAGAGGAAGCTCAC	RT-PCR
B. floridae	210643	R2	34-PV-RT	CCACAGACTGTTGCTCTTG	RT-PCR
B. floridae	94083	F1	35-PV-RT	CGTCATCACTTTCTGCTAC	RT-PCR
B. floridae	94083	R2	36-PV-RT	GACTTGGCGAACATGCTG	RT-PCR
B. floridae	90832	F2	37-PV-RT	AGTACCAGAACCACTTCTCA	RT-PCR
B. floridae	90832	R3	38-PV-RT	TTCCTCCAGCTCCACATCAG	RT-PCR
B. floridae	cyt. actin	F	39-PV-RT	GGACAGGTCATCACCATC	RT-PCR
B. floridae	cyt. actin	R	40-PV-RT	GAGGAGCGATGATCTTGA	RT-PCR
B. floridae	calmodulin	F	41-PV-RT	CTTCAGCCTCTTCGACAAG	RT-PCR
B. floridae	calmodulin	R	42-PV-RT	CCATCATTGTGAGGAACTC	RT-PCR
B. floridae	AmphiRx	F	300-PV	GTGGTGGACATTCTCAACG	RT-PCR
B. floridae	AmphiRx	R	301-PV	CGTCTTCGGCCGTCTCAG	RT-PCR

Cloning amphioxus transcription factors

Organism	"Genes"	Orient.	Name	Sequence (5´-3´)	Purpose
B. floridae	AmphiOtx	F	43-PV	GAGGATCCATGGCGTACATGAAGTCACCGTAC	full-length CDS
B. floridae	AmphiOtx	R	44-PV	GCAAGCTTCAGTCGGCTTCACAGAACCTG	full-length CDS
B. floridae	AmphiRx	F	ZK834A	GAGGATCCATGAACGGCCAGAGCGACAG	full-length CDS
B. floridae	AmphiRx	R	ZK834B	CGTCAAGCTTCCTGAGTTGCAAACTAACTCATG	full-length CDS
B. floridae	AmphiProx	F	60-PV	CAGGATCCATGTATCCATGTCTAACCCAGATG	full-length CDS
B. floridae	AmphiProx	R	49-PV	CGAAGCTTAGGTCTCACTCTCCAAGGAGCTC	full-length CDS

Cloning *Platynereis dumerilii* Prox

Organism	Gene	Orient.	Name	Sequence (5´-3´)	Purpose
any	Prox	F1a	129-PV	gcnaarctnatgttyttyt	degenerate Prox
any	Prox	F1b	130-PV	gcnaarttratgttyttyt	degenerate Prox
any	Prox	F2	131-PV	garttytaytayathcaratgga	degenerate Prox
any	Prox	R1a	132-PV	atngcncgraaraaytc	degenerate Prox
any	Prox	R1b	133-PV	atngcnytraaraaytc	degenerate Prox
any	Prox	R2a	134-PV	ttrtadatngmyttyttcca	degenerate Prox
any	Prox	R2b	135-PV	ttrtadatrytyttyttcca	degenerate Prox
invertebrate	Prox	F	157-PV	aayaaraayaayacngcnca	degenerate Prox
P. dumerilii	Prox 3RACE	F	160-PV	GCCAGGCCTTAGCTGAGGGCATG	3` RACE
P. dumerilii	Prox 3RACE	F	161-PV	GAGGGCATGAAGAATGCTGAGGACC	3` RACE
P. dumerilii	Prox 3RACE	F	162-PV	gaCCAGATTGAGGTCCCAGAATCCTTCC	3` RACE
P. dumerilii	Prox 5RACE	R	165-PV	TCCTCAGCATTCTTCATGCCCTCAG	5` RACE
P. dumerilii	Prox 5RACE	R	166-PV	GTAGATGGCCTTCTTCCACGACTG	5` RACE
P. dumerilii	Prox 5RACE	R	167-PV	GATGGGTCAAGTCGAAGGTCAATAGGTCAG	5` RACE

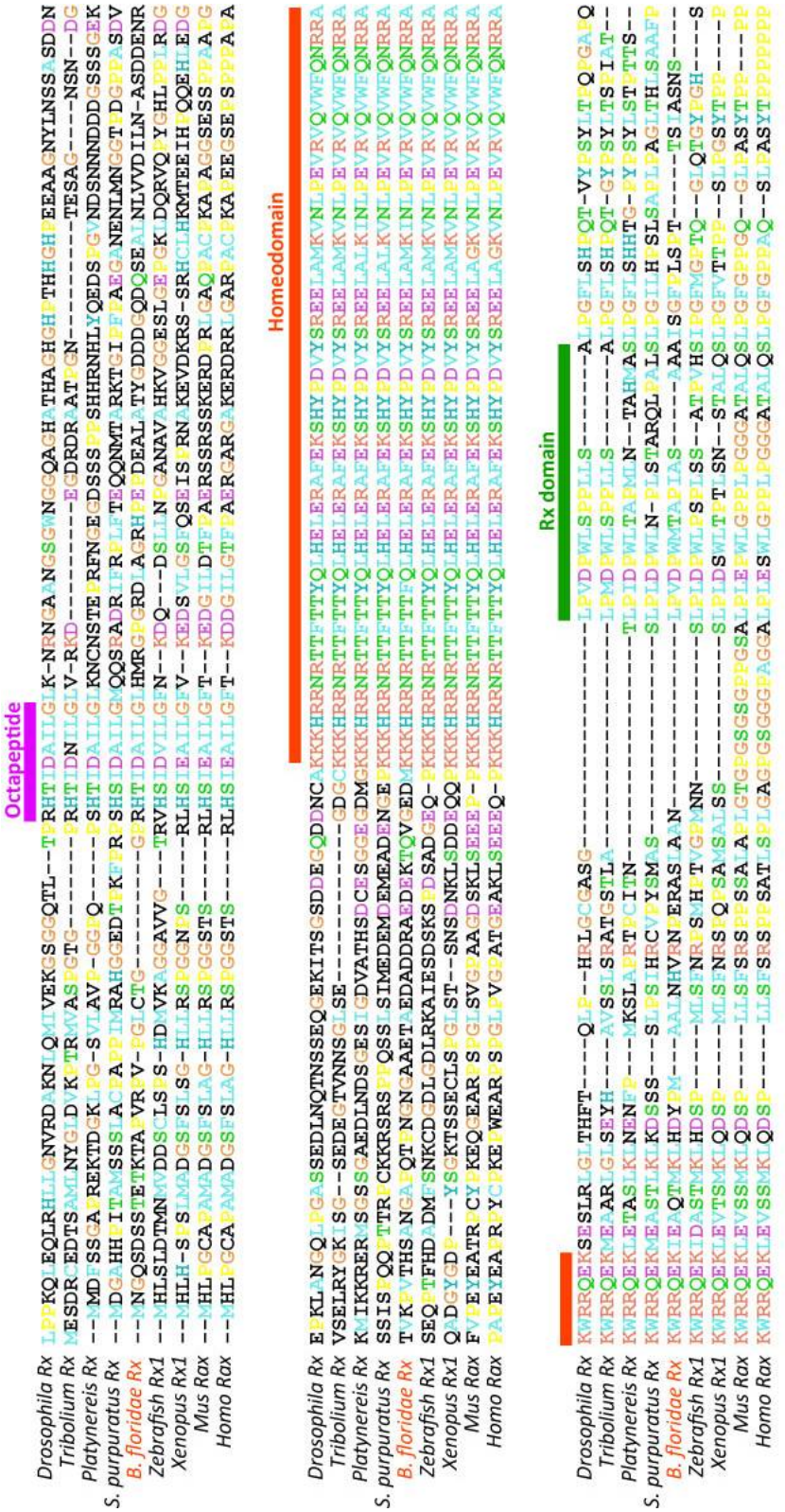
Cloning into into pET42a+ for protein expression

Organism	Gene	Orient.	Name	Sequence (5´-3´)	Purpose
B. floridae	AmphiOtx	F	43-PV	GAGGATCCATGGCGTACATGAAGTCACCGTAC	N-Otx AB
B. floridae	AmphiOtx	R	63-PV	GCAAGCTTCTAAGGAGGCGGCACTGTCCCGTC	N-Otx AB
B. floridae	AmphiOtx	F	64-PV	CAGGATCCCAGGCCAAACCCCGTCCCAAG	C-Otx AB
B. floridae	AmphiOtx	R	44-PV	GCAAGCTTCAGTCGGCTTCACAGAACCTG	C-Otx AB
B. floridae	AmphiRx	F	ZK834A	GAGGATCCATGAACGGCCAGAGCGACAG	N-Rx AB
B. floridae	AmphiRx	R	66-PV	GCAAGCTTTCAGTCTTCTCCCACCTGCGTCTTC	N-Rx AB
B. floridae	AmphiRx	F	67-PV	CAGGATCCATAGAAGCCCAGACCATGAAG	C-Rx AB
B. floridae	AmphiRx	R	ZK834B	CGTCAAGCTTCCTGAGTTGCAAACTAACTCATG	C-Rx AB
B. floridae	AmphiProx	F	69-PV	CAGGATCCAAGTTTGCTCGCCAGGCCATGAG	C-Prox AB
B. floridae	AmphiProx	R	49-PV	CGAAGCTTAGGTCTCACTCTCCAAGGAGCTC	C-Prox AB
P. dumerilii	Pdu-Prox	F	163-PV	gctgAGATCTGAGAAATACTCCCGCCAGGCCTTAG	C-Prox AB
P. dumerilii	Pdu-Prox	R	164-PV	AAGCTTACATATCGTTCAACGTTTCCATCC	C-Prox AB

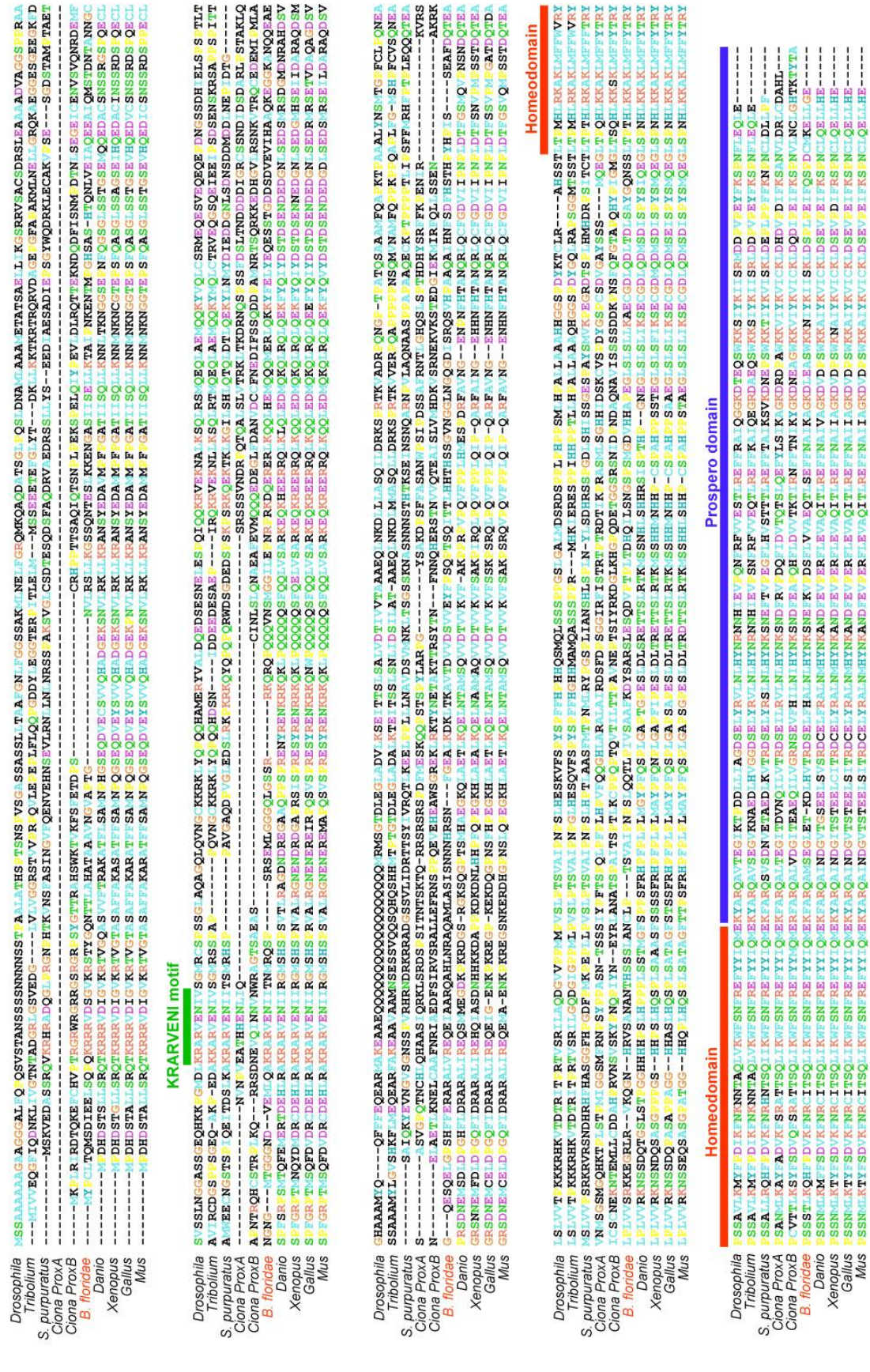
Luciferase reporter assays

B. floridae	Otx HD mut Q50K	F	127-PV	GAGTACAGGTGTGGTTCcAGAACAGAAGGGCCAAG
B. floridae	Otx HD mut Q50K	R	128-PV	CTTGGCCCTTCTGTTCTGGAACCACACCTGTACTC
3x prospero DNA-binding motif		F	138-PV	CTAGCGGCTAAGACGATTTCTGTGGCGGCTAAGACGATTTCTGTGGCGGCTAAGACGC
3x prospero DNA-binding motif		R	139-PV	TCGAGCGTCTTAGCCGCCACAGAAATCGTCTTAGCCGCCACAGAAATCGTCTTAGCCG
D. melanogast Rh6 promoter		F	140-PV	GCTGGATCCGATCAGACGGAGCAGCTGC
D. melanogast Rh6 promoter		R	141-PV	GCTGGATCCGCAGGACGGCTGACTCTCG
3x TAATT		F	ZK901A	CTAGCCTGCCTAATTAAATTAGCGTGTCGTAATTAAATTATGCTTCGGTAATTAAATTACCGTCC
3x TAATT		R	ZK901B	TCGGGACGAGTAATTTAATTACCGAAGCATAATTTAATTACGACACGCTAATTTAATTAGGCAGG
3x TAATC		F	ZK902A	CTAGCTAATCAGATTAGCGTGTCGTAATTCAGTTATGCTTCGGTAATCAGATTAC
3x TAATC		R	ZK902B	TCGAGTAATCTGATTACCGAAGCATAATCTGATTACGACACGCTAATCTGATTAG

8.3 Sequence alignment of Rx proteins



8.4 Sequence alignment of Prox proteins



8.5 Accession numbers

8.5.1 Opsin alignment

Gallus adenosine receptor NP_990418.1, *Mus* serotonine receptor NP_766400.1, *Branchiostoma belcheri* opsin 6 BAC76024.1, *Drosophila* Rh6 NP_524368.3, *Drosophila* Rh2 NP_524398.1, *Drosophila* Rh3 NP_524411.1, *Apis* blue rhodopsin NP_001011606.1, *Mizuhopecten* Gq O15973, *Platynereis* r-opsin CAC86665.1, *Octopus* opsin P09241, *B. belcheri* Mop Q4R1I4, *Xenopus* melanopsin AAC41235.1, *Danio* melanopsin NP_840074.1, *Homo* melanopsin NP_150598.1, *Strongylocentrotus* Sp1 GLEAN3_05569, *Homo* rhodopsin NP_000530.1, *Danio* extra-ocular NP_571287.1, *Latimeria* Rh2 AAD30520.1, *Danio* SW opsin NP_571394.1, *Xenopus* violet P51473, *Xenopus* green AAO38746.1, *Salmo* VAL opsin O13018, *Danio* VAL opsin NP_571661.1, *Ciona* opsin NP_001027727.1, *Xen*. parapinopsin NP_998830.1, *Uta* parietopsin AAZ79904.1, *Xen*. parietopsin NP_001039256.1, *Mus* encephalopsin NP_034228.1, *Homo* encephalopsin NP_055137.2, *Platynereis* c-opsin AAV63834.1, *Takifugu* TMT NP_001027778.1, *B. belcheri* opsin 4 BAC76021.1, *B. belcheri* opsin 5 BAC76022.1, *Apis* pteropsin NP_001035057.1, *Aedes* opsin EAT43163.1, *Anopheles* GPRop11 XP_312503.3, *Anopheles* GPRop12 XP_312502.2, *Mizuhopecten* Go O15974, *B. belcheri* opsin 2 BAC76020.1, *B. belcheri* opsin 1 BAC76019.1, *Strongylocentrotus* Sp3.2 GLEAN3_27633, *Strongylocentrotus* Sp3.1 GLEAN3_27634, *Rattus* Opn5 NP_861437.1, *Homo* Opn5 NP_859528.1, *Homo* peropsin NP_006574.1, *Mus* peropsin AAC53344.1, *B. belcheri* opsin 3 BAC76023.1, *Gallus* RGR NP_001026387.1, *Mus* RGR NP_067315.1, *Todarodes* retinochrome CAA40422.1.

8.5.2 Gα subunits alignment

Giardia ‘transducin’ - XP_001709656.1; *Drosophila* melanogaster GNAF - Q05337.1; *Drosophila* simulans GNAF - XP_002085174.1; *Tribolium* GNAF - XP_970742.1; *Brugia* GNAS - XP_001901198.1; *Danio* GNAS - XP_971664.2; *Homarus* GNAS - O16118.1; *Drosophila* GNAS - P20354.1; *Lymnaea* GNAS - CAA78808.1; *S. purpuratus* GNAS - NP_001001474.1; *Lytechinus* GNAS - AAS38583.1; *Capitella* GNAV - JGI Capca1:227716; *Oryzias* GNAV - (Oka et al, 2009); *Danio* GNAV - XP_699972.2; *Tribolium* GNAV - (Oka et al, 2009); *C. elegans* GPA2 - P22454.1; *C. briggsae* GPA2 - Q4VT35.1; *Drosophila* GNAI - P20353.2; *Homarus* GNAI - P41776.2; *Helisoma* GNAI -

P51876.2; *Lymnaea* GNAI - P30682.3; *Patiria* GNAI - P30676.3; *Helisoma* GNAO - AAC41539.1; *Mizuhopecten* GNAO - O15976.3; *Drosophila* GNAO - P16378.1; *C. elegans* GNAO - P51875.3; *Danio* GNAT - AAL05601.1; *Sparus* GNAT - AAB41887.1; *Ambystoma* GNAT - AAC67569.1; *Canis* GNAT - NP_001003068.1; *Bos* GNAT - P04695.3; *X. laevis* GNAT - NP_001084030.1; *Homo* GNA13 - NP_006563.2; *Mus* GNA13 - NP_034433.3; *Danio* GNA13 - AAR25617.1; *Danio* GNA12 - NP_001013295.1; *Mus* GNA12 - NP_034432.1; *S. purpuratus* GNA12 - NP_001001476.1; *Nasonia* GNA12/13 - XP_001600076.1; *Drosophila* GNAQ - P23625.2; *Loligo* GNAQ - P38412.1; *Homarus* GNAQ - AAB49314.1; *Lymnaea* GNAQ - P38411.1; *Mizuhopecten* GNAQ - O15975.1; *S. purpuratus* GNAQ - NP_999835.1; *Danio* GNAQ - CAK04448.1; *Homo* GNAQ - NP_002063.2; *Gallus* GNAQ - NP_001026598.1; *X. laevis* GNAQ - AAH81126.1;

8.5.3 Rx alignment

Drosophila Rx - CAA11241.1; *Tribolium* Rx XP_973468.1; *Platynereis* Rx - AAU20320.1; *S. purpuratus* Rx - XP_782307.1; *B. floridae* JGI ID 78608; *Zebrafish* Rx1 - NP_571300.2; *X. laevis* Rx1 - AAB70267.1; *Mus Rax* - NP_038861.2; *Homo Rax* - NP_038463.1;

8.5.4 Prox alignment

Drosophila prospero - BAA01464.1; *Tribolium* Prospero - XP_971664.2; *S. purpuratus* Prox - XP_001189091.1; *Ciona* ProxA - BAE06658.1; *Ciona* ProxB - BAE06659.1; *Danio* Prox1 - NP_571480.2; *X. laevis* Prox1 - BAB17310.1; *Gallus* Prox1 - NP_001005616.1; *Mus* Prox1 - NP_032963.1; *Homo* Prox1 - NP_002754.2; *Capitella* – JGI 2449; *Lottia* – JGI 107690;

8.6 Publications

8.6.1 L. Z. Holland *et al.*, *Genome Res* 18, 1100 (Jul, 2008)

Holland LZ, Albalat R, Azumi K, Benito-Gutiérrez E, Blow MJ, Bronner-Fraser M, Brunet F, Butts T, Candiani S, Dishaw LJ, Ferrier DE, Garcia-Fernàndez J, Gibson-Brown JJ, Gissi C, Godzik A, Hallböök F, Hirose D, Hosomichi K, Ikuta T, Inoko H, Kasahara M, Kasamatsu J, Kawashima T, Kimura A, Kobayashi M, Kozmik Z, Kubokawa K, Laudet V, Litman GW, McHardy AC, Meulemans D, Nonaka M, Olinski RP, Pancer Z, Pennacchio LA, Pestarino M, Rast JP, Rigoutsos I, Robinson-Rechavi M, Roch G, Saiga H, Sasakura Y, Satake M, Satou Y, Schubert M, Sherwood N, Shiina T, Takatori N, Tello J, Vopalensky P, Wada S, Xu A, Ye Y, Yoshida K, Yoshizaki F, Yu JK, Zhang Q, Zmasek CM, de Jong PJ, Osoegawa K, Putnam NH, Rokhsar DS, Satoh N, Holland PW.
The amphioxus genome illuminates vertebrate origins and cephalochordate biology.
Genome Res. 2008 Jul;18(7):1100-11
IF = ~ 10.2

Contribution:
In this paper associated with the first *Branchiostoma floridae* genome release, I was responsible for the opsin part: I performed the bioinformatic screen of amphioxus genome for opsin genes, manually currated gene models, performed the phylogenetic analysis, prepared figures and together with the supervisor wrote the corresponding text in the manuscript.

This appendix contains the main text and the supplementary material pages concerning the opsins (16 pages in total).

8.6.2 Z. Kozmik *et al.*, *Proc Natl Acad Sci U S A* 105, 8989 (Jul 1, 2008)

Kozmik Z, Ruzickova J, Jonasova K, Matsumoto Y, Vopalensky P, Kozmikova I, Strnad H, Kawamura S, Piatigorsky J, Paces V, Vlcek C.
Assembly of the cnidarian camera-type eye from vertebrate-like components.
Proc Natl Acad Sci U S A. 2008 Jul 1;105(26):8989-93
IF = ~ 9.4

Contribution:
I prepared the normalized *Tripedalia cystophora* larva cDNA library for deep EST sequencing; I performed the phylogenetic analysis of opsin, phosducin and MiTF.

This appendix contains the main text and the supplementary material (11 pages in total).

8.6.3 P. Vopalensky, Z. Kozmik, *Philos Trans R Soc Lond B Biol Sci* 364, 2819 (Oct 12, 2009)

Vopalensky P and Kozmik Z.
Eye evolution: common use and independent recruitment of genetic components.
Philos Trans R Soc Lond B Biol Sci. 2009 Oct 12;364(1531):2819-32.
IF = ~ 5.6

Contribution:
This review paper was written together with my supervisor. I took the main part responsibility for the chapters concerning photoreceptor, opsin, dark pigments, Otx, Rx and MiTF.

This appendix contains the main text and the supplementary material (20 pages in total).