

The opsin repertoire of the European lancelet: a window into light detection in a basal chordate

CHRYSOULA N. PANTZARTZI^{1,#}, JIRI PERGNER^{2,#}, IRYNA KOZMIKOVA² and ZBYNEK KOZMIK^{*,1, 2}.

¹Laboratory of Eye Biology, Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic, Division BIOCEV, Vestec, Czech Republic and ²Laboratory of Transcriptional Regulation, Institute of Molecular Genetics of the ASCR, Prague, Czech Republic

ABSTRACT Light detection in animals is predominantly based on the photopigment composed of a protein moiety, the opsin, and the chromophore retinal. Animal opsins originated very early in metazoan evolution from within the G-Protein Coupled Receptor (GPCR) gene superfamily and diversified into several distinct branches prior to the cnidarian-bilaterian split. The origin of opsin diversity, opsin classification and interfamily relationships have been the matter of long-standing debate. Comparative studies of opsins from various Metazoa provide key insight into the evolutionary history of opsins and the visual perception in animals. Here, we have analyzed the genome assembly of the cephalochordate *Branchiostoma lanceolatum*, applying BLAST, gene prediction tools and manual curation in order to predict *de novo* its complete opsin repertoire. We investigated the structure of predicted opsin genes, encoded proteins, their phylogenetic placement, and expression. We identified a total of 22 opsin genes in *B. lanceolatum*, of which 21 are expressed and the remaining one appears to be a pseudogene. According to our phylogenetic analysis, representatives from the three major opsin groups, namely C-type, the R-type and the Group 4, can be identified in *B. lanceolatum*. Most of the *B. lanceolatum* opsins exhibit a stage-specific, but not a tissue-specific, expression pattern. The large number of opsins detected in *B. lanceolatum*, the observed similarities and differences in terms of sequence characteristics and expression patterns lead us to conclude that there may be a fine tuning in opsin utilization in order to facilitate visually-guided behavior of European amphioxus under various environmental settings.

KEY WORDS: *Branchiostoma*, *amphioxus*, *opsin*, *expression*

Introduction

Light sensing systems have evolved to be uniquely suited to the environment and behavior of any given species. Animals detect light using sensory cells known as photoreceptors, present in the eyes or, in case of extraocular photoreceptors, outside of the eyes. Although other systems of light detection exist in the animal kingdom, such as cryptochromes (Rivera *et al.*, 2012) or LITE-1 (Gong *et al.*, 2016), opsins are dominantly utilized as visual pigments among Metazoa. Opsins are seven transmembrane domain proteins that belong to the G-Protein Coupled Receptor (GPCR) superfamily and are canonically distinguished from other GPCRs by a highly conserved lysine in the seventh helix. The number of

opsin genes differs significantly among species studied so far and does not generally correlate with the overall level of body plan sophistication. Cnidaria for example, despite having a relatively simple body plan and limited number of cell types, are known to possess a large number of opsins originating by species-specific gene duplications (Koyanagi *et al.*, 2008; Kozmik *et al.*, 2008; Suga *et al.*, 2008). Opsin classification, interfamily relationships and evolution of animal vision have been the debate of numerous studies so far (D'Aniello *et al.*, 2015; Feuda *et al.*, 2012; Liegertova *et al.*,

Abbreviations used in this paper: CNS, Central Nervous System, GPCR, G-protein coupled receptor, TM, transmembrane.

*Address correspondence to: Zbynek Kozmik. Institute of Molecular Genetics of the Czech Academy of Sciences, v.v.i., Videnska 1083, 14220, Prague, Czech Republic. Tel: +420 241062110. E-mail: kozmik@img.cas.cz -  <http://orcid.org/0000-0002-5850-2105>

#Note: These authors contributed equally.

Supplementary Material for this paper is available at: <http://dx.doi.org/10.1387/ijdb.170139zk>

Submitted: 21 June, 2017; Accepted: 11 July, 2017.

2015; Peirson *et al.*, 2009; Plachetzki *et al.*, 2007; Porter *et al.*, 2012; Ramirez *et al.*, 2016; Shichida and Matsuyama, 2009; Suga *et al.*, 2008; Terakita, 2005). Opsins can be roughly clustered into four major groups, namely the ciliary opsins expressed in ciliary photoreceptors (C-type), the rhabdomeric opsins expressed in rhabdomeric photoreceptors (R-type), the Group 4 opsins, and the Cnidarian opsins. A further subdivision of the first three groups into seven subfamilies was suggested (Lamb *et al.*, 2007; Terakita, 2005), based on their sequence and on the type of the G-protein to which they are coupled: 1) the vertebrate visual and non-visual, 2) the encephalopsin/TMT, 3) the Gq-coupled/melanopsin, 4) the neuropsin, 5) the Go-coupled, 6) the peropsin, and 7) the retinal photoisomerase subfamilies. Ramirez and colleagues have proposed that a repertoire of at least nine opsin paralogs was present in the bilaterian ancestor (Ramirez *et al.*, 2016). In regard to the cnidarian opsins, there are conflicting results regarding their position in the phylogenetic tree (Feuda *et al.*, 2012; Liegertova *et al.*, 2015; Plachetzki *et al.*, 2007; Porter *et al.*, 2012; Suga *et al.*, 2008). The function of most opsins consists of two steps: light absorption and G-protein activation in both visual and non-visual systems. Isomerization of the chromophore 11-*cis*-retinaldehyde to all-*trans*-retinaldehyde due to light absorption, changes the conformation of the opsin and triggers a signal transduction cascade, the type of which is dependent on the G-protein to which the opsin binds

Fig. 1. Molecular phylogenetic analysis of opsins by Maximum Likelihood method.

The evolutionary history of opsin proteins was inferred by using the Maximum Likelihood method based on the Le_Gascuel_2008 model. The tree with the highest log likelihood is shown. Bootstrap values are shown (only values >50) either at the nodes or above the branches in the case of collapsed subgroups (e.g. neuropsin). A discrete Gamma distribution was used to model evolutionary rate differences among sites (2 categories). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 787 amino acid sequences. There were a total of 419 positions in the final dataset (third cytoplasmic loop is excluded). Numbers in yellow boxes correspond to the bilaterian opsin paralogs identified by Ramirez *et al.*, 2016.

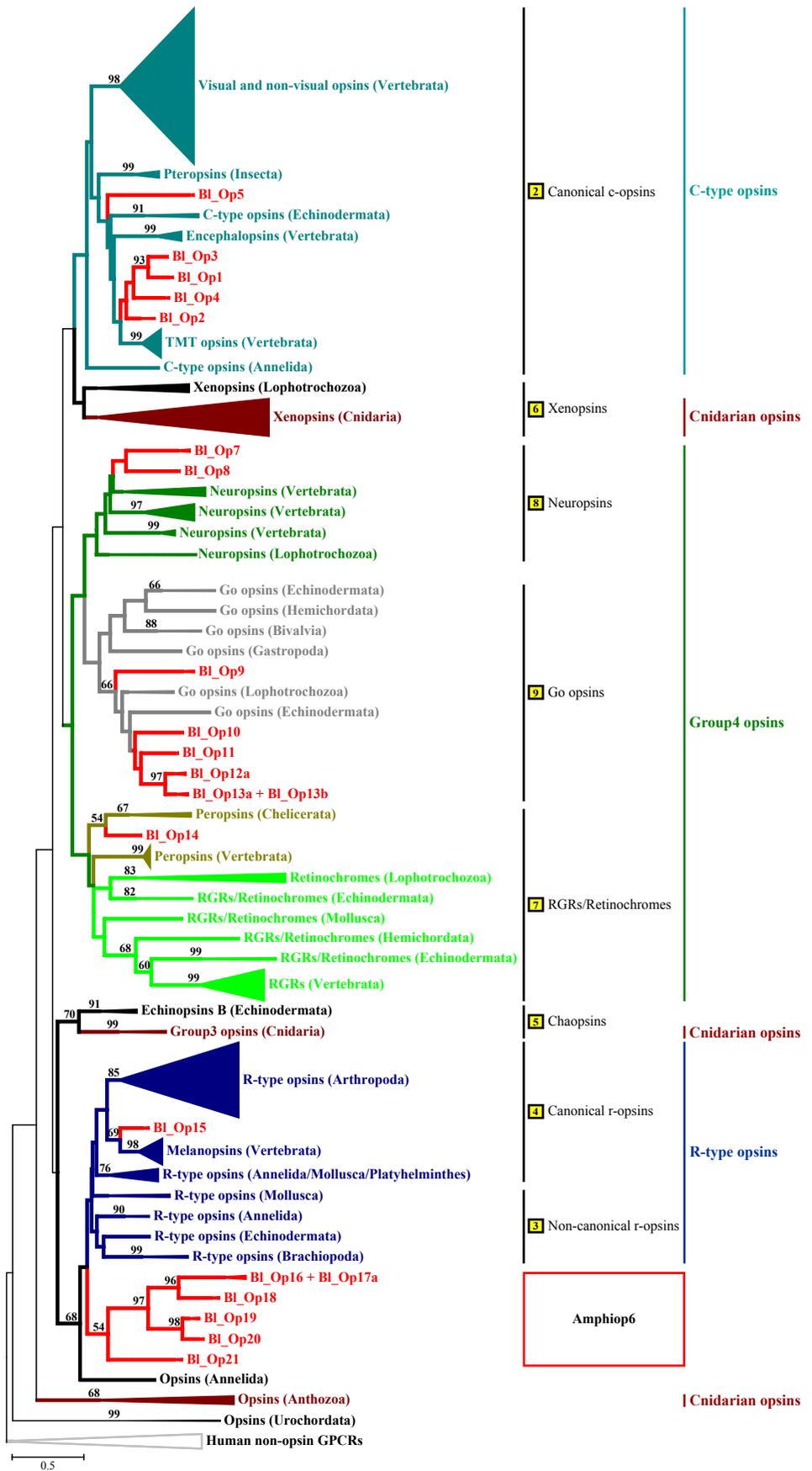


TABLE 1

NOMENCLATURE FOR BRANCHIOSTOMA OPSIN GENES

Group	This study	<i>Branchiostoma floridae</i> models	<i>B. belcheri</i> mRNA sequences	References
C-type opsins	<i>Bl_op1</i> (MF464463)	124039	AB050610	c-opsin1 (Vopalensky et al., 2012), <i>Amphiop5</i> (Koyanagi et al., 2002)
	<i>Bl_op2</i> (MF464464)	205982		
	<i>Bl_op3</i> (MF464465)	84894	AB050609	c-opsin3 (Vopalensky et al., 2012), <i>Amphiop4</i> (Koyanagi et al., 2002)
	<i>Bl_op4</i> (MF464466)	70447		
	<i>Bl_op5</i> (MF464467)	84890		
Neuroopsins	Not identified	210643		
	<i>Bl_op7</i> (MF464468)	65045		
	<i>Bl_op8</i> (MF464469)	94083		
Go opsins	<i>Bl_op9</i> (MF464470)	71561	AB050607	<i>Amphiop2</i> (Koyanagi et al., 2002)
	<i>Bl_op10</i> (MF464471)	215180		
	<i>Bl_op11</i> (MF464472)	84844		
	<i>Bl_op12a</i> (MF464473)	91094	AB050606	<i>Amphiop1</i> (Koyanagi et al., 2002)
	<i>Bl_op13a</i> (MF464474)	91095		
	<i>Bl_op13b</i> (MF464475)	Not identified		
Peropsins	<i>Bl_op14</i> (MF464476)	90832	AB050608	<i>Amphiop3</i> (Koyanagi et al., 2002)
Melanopsins	<i>Bl_op15</i> (MF464477)	65960	AB205400	<i>AmphMop</i> (Koyanagi et al., 2005)
Amphiop6	<i>Bl_op16</i> (MF464478)	86640		
	<i>Bl_op17a</i> (MF464479)	201585		
	<i>Bl_op17b</i> (MF464480)	Not identified		
	<i>Bl_op18</i> (MF464481)	Not identified		
	<i>Bl_op19</i> (MF464482)	87094	AB050611	<i>Amphiop6</i> (Koyanagi et al., 2002)
	<i>Bl_op20</i> (MF464483)	110003		
	<i>Bl_op21</i> (MF464484)	86195		

* Models from JGI, genome version v1.0 (<http://genome.jgi.doe.gov/Brafl1/Brafl1.home.html>)

(Koyanagi *et al.*, 2008; Liegertova *et al.*, 2015; Yarfitz and Hurley, 1994). The conserved lysine in the seventh helix is used to form a Schiff-base bond to the retinal chromophore (Bownds, 1967; Wang *et al.*, 1980) and in its protonated form it is stabilized by a negatively charged amino acid, called a counterion, whose position varies among different opsin subfamilies (reviewed in Porter *et al.*, 2012; Shichida and Matsuyama, 2009).

Comparative studies of opsins provide valuable insight not only into the origins of opsin diversity but also into the evolution of visual organs and light perception in animals. To help answer questions about vertebrate evolution at the key invertebrate chordate-vertebrate transition, such as “How did the vertebrate eye evolved?”, one can examine opsins and visual organs of the extant and most basally divergent chordates, the cephalochordates. The subphylum Cephalochordata, a.k.a. amphioxus or lancelet, consists of approximately 29 extant species, with a worldwide distribution (Poss and Boschung, 1996). Amphioxus possesses four types of photoreceptive systems – the dorsal ocelli and the Joseph cells are rhabdomeric receptors, while the frontal eye and the lamellar body contain ciliary photoreceptors (Lacalli, 2004). Function of the first two receptors is still not clear, nevertheless, rhabdomeric receptors similar to Joseph cells are present in the cerebral eyes of tunicates (salps). Amphioxus frontal eye is considered homologous to the paired eyes of vertebrates, since its photoreceptors and pigment cells co-express a combination of transcription factors and opsins typical of the vertebrate eye photoreceptors (Vopalensky *et al.*, 2012). The lamellar body, on the other hand, is proposed as the amphioxus homolog of the pineal gland (Lacalli, 2004). Seven mRNA (*Amphiop1*–*Amphiop6* and *AmphMop*) and 20 opsin genes have been previously identified in *B. belcheri* and *B. floridae*, respectively (Holland *et al.*, 2008; Koyanagi *et al.*, 2005; Koyanagi *et al.*, 2002). There are representatives from the encephalopsin/TMT, neuropsin, Go, peropsin and melanopsin subfamilies; a divergent subfamily, *Amphiop6* appears to be specifically duplicated in amphioxus (Holland *et al.*, 2008). No homologs of the vertebrate visual/nonvisual opsin subfamily were detected (Holland *et al.*, 2008). Studies on

expression pattern and function of the *Branchiostoma* opsins are rather limited; the biochemical data support *AmphiOp1* as a typical visual pigment and *AmphiOp3* as a photoisomerase (Koyanagi *et al.*, 2002), while the properties of the *B. belcheri* melanopsin homolog were found to resemble those of the visual opsins present in the intrinsically photosensitive rhabdomeric photoreceptor cells of vertebrates (Koyanagi *et al.*, 2005). In agreement with their phylogenetic placement, antibody staining revealed specific expression of two ciliary-type opsins in the ciliary photoreceptor cells of the *B. floridae* frontal eye (Vopalensky *et al.*, 2012). Here, we report for the first time the opsin repertoire of *B. lanceolatum* (Pallas, 1774) and provide information on the expression patterns of opsin genes across multiple tissues and developmental stages.

Results

Identification, classification and genome organization of opsin genes in *Branchiostoma lanceolatum*

To identify the opsin gene repertoire in *B. lanceolatum* we used the available genome assembly provided by the *Branchiostoma lanceolatum* genome consortium. Applying BLAST searches and *de novo* gene prediction we were able to identify 22 opsin genes in *B. lanceolatum*, out of which *Bl_op17b* is a putative non-functional gene (pseudogene), since it bears a stop codon in the first exon. Predicted transcripts and encoded proteins for newly characterized opsins in *B. lanceolatum*, details on gene organization and genomic location are provided in Supplementary file 1. In general, all predicted opsins have seven transmembrane helices and a lysine in the seventh helix. We did not detect any intronless opsins in *B. lanceolatum*.

The large scale phylogenetic relations of opsins has been the object of many studies (Albalat, 2012; Liegertova *et al.*, 2015; Porter *et al.*, 2012; Ramirez *et al.*, 2016). We wanted to classify the *B. lanceolatum* opsins and investigate their distribution over the major opsin groups. We combined the available datasets from the most recent analyses and enriched them with the 21 newly

identified homologs from *B. lanceolatum*; BI_Op17b was omitted, for being the product of a putative pseudogene. Two different alignments were used for our phylogenetic analyses, excluding (Fig. 1 and Supplementary Fig. S1) and including (Suppl. Fig. S2) the variable third cytoplasmic loop. In both datasets, the four traditional opsin groups, i.e. the C-type, the R-type, Group4 and the Cnidarian opsins, were recovered along with some other groups, albeit supported by low bootstrap values in many cases. Clustering of *B. lanceolatum* opsins is the same in both cases. Few differences were observed between the two datasets; two of the most striking are the relative positioning of R-type and Group4 opsins and the placement of the clade containing the *CiNut* homolog (Etani and Nishikata, 2002) (Fig. 1 and Supplementary Fig. S1 and Fig. S2). In Fig. 1, the major cnidarian group together with a small set of lophotrochozoan opsins (the Xenopsin according to Ramirez *et al.*, 2016) cluster as a sister group to C-type opsins. Group4 is closer to C-type and the Xenopsin group, than the R-type group. Chaopsin (Ramirez *et al.*, 2016), consisting of Echinopsin B (D'Aniello *et al.*, 2015) and a small clade of cnidarian opsins namely Group3 (Mason *et al.*, 2012), an anthozoan and a *Ciona*-specific clade are identified in our phylogenetic trees (Fig. 1 and Supplementary Fig. S1 and Fig. S2).

Members from almost all opsin subfamilies have been identified for *B. lanceolatum* (Fig. 1 and Suppl. Fig. S1 and Fig. S2). Five belong to the C-type (BI_Op1-5), two are in a clade sister to neuropsins (BI_Op7 and BI_Op8), six cluster with the Go opsins (Op9-Op13b), one with peropsins (Op14), one with melanopsins (BI_Op15) and six within the Amphiop6 clade (BI_Op16-BI_Op21). Based on sequence similarity and genomic location, BI_Op17b should be a member of the Amphiop6 group. No homologs of the vertebrate visual/non visual opsins have been detected. *B. lanceolatum* genes identified in this study, arranged in groups according to their phylogenetic placement, and their relation to previously identified homologs from two other *Branchiostoma* species are provided in Table 1. Next, we wanted to visualize how opsin genes are arranged in the genome of *B. lanceolatum* and whether there is some genetic linkage between opsin gene paralogues. The *B. lanceolatum* opsin-containing loci identified during our *in silico* analysis are depicted in Fig. 2. In particular, opsin genes are spread over 16 genomic regions (scaffolds) in *B. lanceolatum*, whereas in some cases, members of the same group are clustered in the same locus (scaffold), for example the Amphiop6 BI_op19 and BI_op20 (Fig. 2).

Table 2 summarizes the residues found in the three putative counterion positions, as well as the tripeptide associated with the binding to G-proteins (Arendt *et al.*, 2004; Marin *et al.*, 2000). Aspartic acid is present at position 83 of not only the melanopsin ortholog (BI_Op15) but of most of the rest of the opsins. In almost all cases,

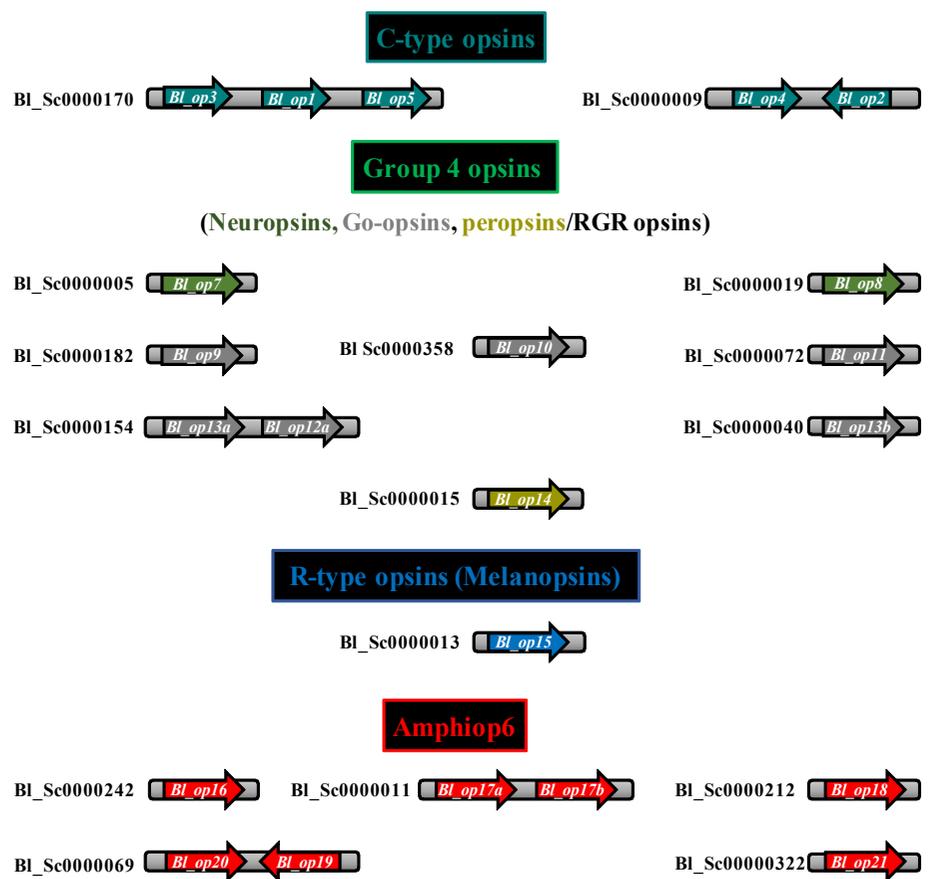


Fig. 2. Opsin containing genomic loci from *Branchiostoma lanceolatum*. The genomic scaffolds from *Branchiostoma lanceolatum* (BI) containing opsin genes. Groups are colored based on their phylogenetic position (see Fig. 1).

position 113 is occupied by neutral or non-polar amino acids with the striking exception of BI_Op11, BI_Op19 and BI_Op20, for which a negatively charged amino acid is present, similarly to vertebrate visual and non-visual opsins. In fact, BI_Op11 bears aspartic acid in both positions 83 and 113, but no negatively charged amino acid residue at position 181. In the case of BI_Op5, a negatively charged amino acid is encountered only at position 83. A really interesting case is that of BI_Op21, in the sense that no negative amino acid is present at any of the three putative counterion positions. All of the C-type opsins, except for BI_Op5, possess the N(N/S)Q tripeptide; the encountered motif in BI_Op5 is EKE. Melanopsins nicely bear the R-type tripeptide HPK (Arendt *et al.*, 2004). This is not the case for BI_Op16-BI_Op22 (AmphiOp6); they may closely cluster with the R-type opsins but they only have the central proline, while the first position is occupied by a non-polar amino acid.

B. lanceolatum opsin gene expression patterns

Opsin genes are used by many animals not only for visual, but also for non-visual tasks. Their tissue specificity thus can significantly vary. Recent analyses of opsin gene expression in Cnidaria or Arthropoda documented a wide range of tissues where opsins can be detected (Liegertova *et al.*, 2015; Battelle *et al.*, 2016). Our study examined the expression pattern of opsins across multiple developmental stages (Fig. 3A and B), as well as in multiple tissue types of *B. lanceolatum* adult body (Fig. 3C and D). To achieve this

we performed quantitative real-time PCR (qRT-PCR). We scanned opsin expression in developmental stages starting from late neurula (N3), where photoreceptive 1st Hesse cell is present, to adult, where frontal eye, rudiments of lamellar body, Joseph cells and series of dorsal ocelli along the neural tube are fully developed (Fig. 3A). All *B. lanceolatum* examined genes were detected at mRNA level, in at least one developmental stage (Fig. 3B and Supplementary Fig. S3A), with the sole exception of *Bl_op17b*. This is in accordance with the *in silico* prediction of this gene as a pseudogene due to a premature stop codon at the beginning of the coding region. The neuropsin *Bl_op7* and the Amphio6 *Bl_op16* show their highest expression in N3 neurula stage; in fact *Bl_op7* is among the dominating opsins at this stage (Supplementary Fig. S3A). However, their expression is significantly reduced at later stages (Fig. 3B). Onset of several other opsin gene expression starts at L1 stage, in which frontal eye and lamellar body (ciliary photoreceptive organs) start to develop. The maximal expression of two Go opsins (*Bl_op9* and *Bl_op10*) and one Amphio6 opsin (*Bl_op20*) was observed at this stage. We should point out that actin was used as a reference gene for normalization, a fact that could lead to a false underrepresentation of opsin genes in later developmental stages (L2/3 or adult), when compared to neurula stage, where lower total number of cells are present. Despite this fact, the majority of the opsins show most predominant expression

in L2/3 stage, where most of the known amphioxus photoreceptor organs - the frontal eye, the lamellar body and the Hesse cell - are differentiated. As expected from the ciliary nature of photoreceptor cells in the frontal eye and the lamellar body, all C-type opsins reach a peak in their expression in L2/3. On the other hand, *B. lanceolatum* opsins belonging to Go coupled opsin group show broad expression across various stages. Most of the Amphio6 opsins show elevated expression in L2/3 stage, which is true also for the single peropsin (*Bl_op14*). All of the 21 examined *B. lanceolatum* opsins could be detected in various adult tissues (Fig. 3C and D). All *B. lanceolatum* opsins but *Bl_op2*, exhibit some specificity for either the cerebral vesicle (10 opsins – *Bl_op1*, *Bl_op4*, *Bl_op5*, *Bl_op7*, *Bl_op9*, *Bl_op11*, *Bl_op12a*, *Bl_op13b*, *Bl_op17a*, *Bl_op20*) or the neural tube (7 opsins – *Bl_op3*, *Bl_op8*, *Bl_op13a*, *Bl_op14*, *Bl_op15*, *Bl_op18*, *Bl_op21*); on the contrary, *Bl_op5*, *Bl_op17a*, *Bl_op21* are expressed in both tissues (Fig. 3D and Supplement Fig. S3B). These findings show that most of the photoreceptive cells in amphioxus reside in the central nervous system, however, the cerebral vesicle and the neural tube probably exhibit a strong specialization for various photoreceptive tasks. Expression of one neuropsin (*Bl_op7*) was documented in female gonads, whereas a Go and an Amphio6 (*Bl_op13a* and *Bl_op16*, respectively) showed noticeable expression in male gonads (Fig. 3D). Interestingly, one of the C-type opsins (*Bl_op2*) displays its highest level of expression in the most posterior part of the tail. Expression of four other opsins was significantly increased in tail. In contrast to analysis of opsin expression across various developmental stages (Fig. 3B), we were not able to observe any preference in tissue-specific usage of opsins belonging to different groups (Fig. 3D). The only exceptions were the peropsin (*Bl_op14*) and the melanopsin (*Bl_op15*), being highly expressed only in neural tube.

To investigate expression during larval development (prior to metamorphosis), we performed whole mount *in situ* hybridization. We analyzed expression of *Bl_op1*, *Bl_op2*, *Bl_op8*, *Bl_op10*, *Bl_op11*, *Bl_op12a*, *Bl_op14*, *Bl_op15*, *Bl_op17a*, *Bl_op19* and *Bl_op20*. Of these, we have observed specific expression patterns only for *Bl_op11*, *Bl_op12* and *Bl_op15* (Fig. 4). *Bl_op11* is expressed in L1 stage in the area of developing preoral pit and 1st gill slit. In L2/3 the signal was detected in pharyngeal region and tail fin (Fig. 4 A,B). For *Bl_op12a* no specific signal was detected in L1 stage. In L2/3 stage the signal was detected in preoral pit, oral papilla, cells around mouth, 1st gill slit and tail fin (Fig. 4C, D). *In situ* hybridization of *Bl_op15* in L1 and L2/3 stage identified specific expression in 1st dorsal ocelli, in agreement with previous findings in *B. belcheri* (Koyanagi *et al.*, 2005).

In summary, our analysis confirms expression of all but one *B. lanceolatum* opsin genes and documents their stage and/or tissue specificity.

Discussion

We sought to characterize the opsin gene family in the genome assembly of the cephalochordate *Branchiostoma lanceolatum* and study the expression patterns of opsin genes from this species in different tissue types and across various developmental stages. We identified a total of 22 opsin genes in *B. lanceolatum*, one of which is a putative pseudogene. There is less than 20% amino acid similarity between vertebrate opsin subfamilies but more than 40% among members of a single family (Peirson *et al.*, 2009; Shichida

TABLE 2

AMINO ACIDS OBSERVED AT PUTATIVE COUNTERION POSITIONS AND THE TRIPEPTIDE AT THE FORTH CYTOPLASMIC LOOP FOR *B. LANCEOLATUM* OPSINS

		Counterion			Tripeptide		
		83	113	181			
C-type opsins	Op1	D	Y	E	N	S	Q
	Op2	D	Y	E	N	S	Q
	Op3	D	Y	E	N	N	Q
	Op4	G	Y	E	N	N	Q
	Op5	D	Y	V	E	K	E
Neuropsins	Op7	D	Y	E	N	N	R
	Op8	D	Y	E	N	D	S
Go opsins	Op9	D	Y	E	S	E	V
	Op10	D	Y	E	H	K	K
	Op11	D	D	Q	N	Q	R
	Op12a	D	Y	E	S	K	A
	Op13a	D	Y	E	N	S	K
	Op13b	D	F	E	N	S	K
Peropsins	Op14	D	Y	E	N	R	R
Melanopsins	Op15	D	Y	E	H	P	K
Amphio6	Op16	G	F	D	L	P	V
	Op17a	D	F	D	L	P	A
	Op18	Q	A	D	L	P	A
	Op19	D	E	D	I	P	S
	Op20	D	E	D	I	P	S
	Op21	N	T	Q	M	P	D

Polar negative	Polar positive	Polar neutral
Non-Polar aromatic	Non-polar aliphatic	Special cases

Amino acids are colored based on their physicochemical properties

and Matsuyama, 2009; Terakita, 2005). Given these low similarity levels, *de novo* prediction of opsin genes could be largely hampered. Therefore, BLAST results should be carefully filtered and used in combination with synteny analyses since true positive results could be obscured by low similarity scores. In addition, manual curation of the genome assembly was needed in some cases.

Discrepancies between different phylogenetic studies have been noted before and could be attributed to the dataset used, the alignment method, the substitution model, and the tree constructing method applied. In our study, we show how sensitive the outcome of the phylogenetic analysis can be to the exclusion or inclusion of the highly variable third cytoplasmic loop (Fig. 1 and Supplementary Fig. S2, respectively). Among the observed differences, two of the most striking are the relative positioning of R-type and Group4 opsins and the placement of the Urochordate opsins

clade containing the *CiNut* homolog (Etani and Nishikata, 2002), for which inconsistencies were previously observed (Albalat, 2012; Porter et al., 2012; Ramirez et al., 2016). A new clade – “bathyopsin” – was recently introduced (Ramirez et al., 2016), consisting of one brachiopod and three echinoderm opsins. We excluded these sequences from our analysis either due to an incorrect number of transmembrane domains, based on TOPCONS and HMMTOP predictions (4 TM in the case of *Strongylocentrotus purpuratus* and 8 in the case of *Lingula anatina*) or because of their extremely small size (as in the case of *Eucidaris tribuloides*). It is obvious that there are still obstacles to reconstructing the complete evolutionary history of opsins. Sampling from specific taxonomic groups is still poor, mainly due to lack of data at the level of whole genome and missing functional data that could greatly facilitate the opsin classification. Collectively, our phylogenetic analysis ascribes *B.*

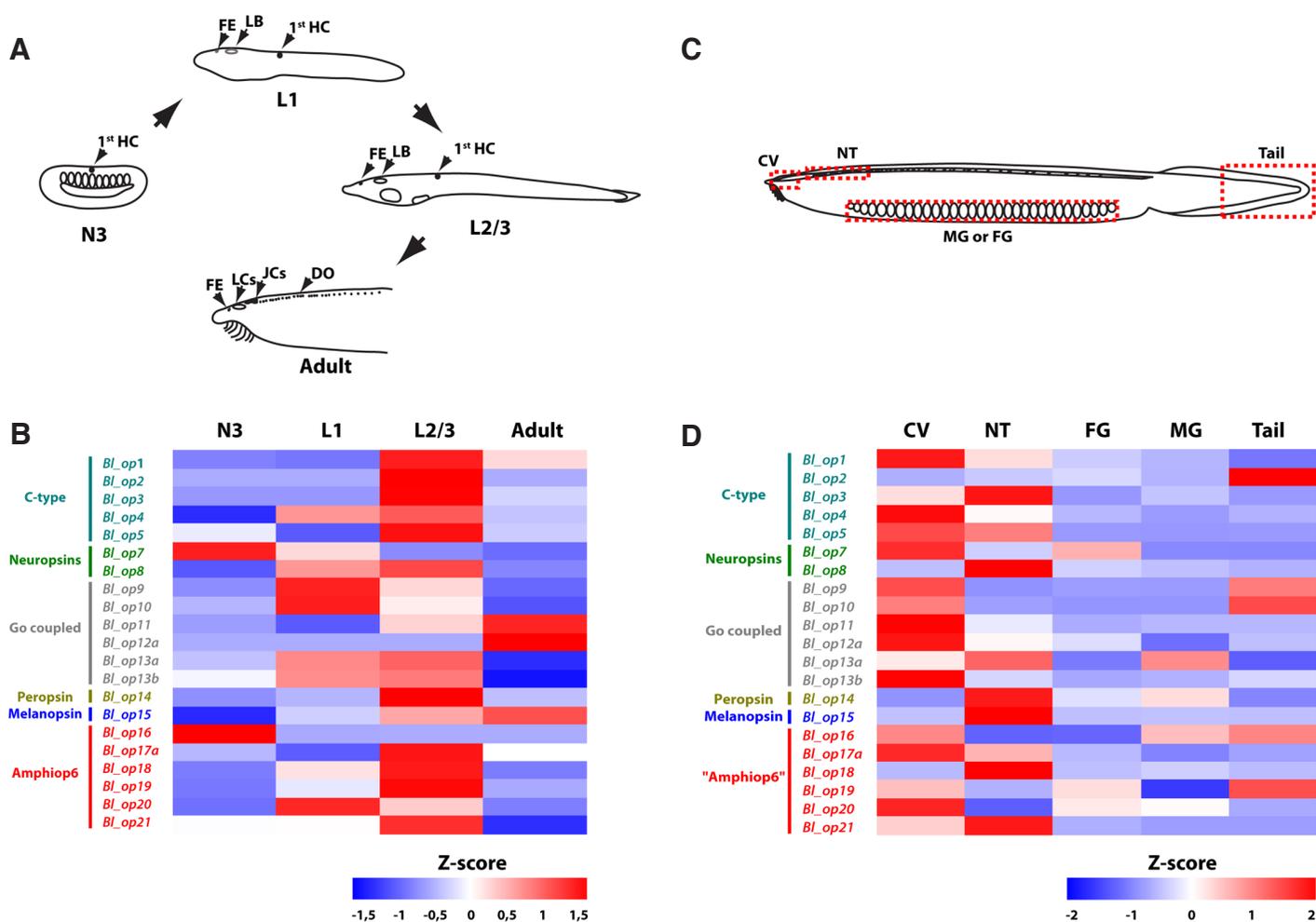


Fig. 3. mRNA expression levels of individual *B. lanceolatum* opsins across different developmental stages and various tissues of the adult body. (A) Schematic drawing of developmental stages (N3, L1, L2/3, adult), in which detection of opsin genes expression was performed. Staging was determined according to Hirakow and Kajita (1994) (see Materials and Methods). 1st HC: 1st Hesse cell, FE: frontal eye, LB: lamellar body, DO: dorsal ocelli, JCs: Joseph cells, LCs: lamellate cells (rudiment of larval lamellar body). (B) Heat map displaying expression of opsin genes across different developmental stages. Each row represents particular opsin gene expression in various developmental stages. (C) Schematic drawing of amphioxus adult body parts in which detection of opsin genes expression was performed. CV: cerebral vesicle, NT: the most anterior third of neural tube including Joseph cells and dense clusters of dorsal ocelli, FG: female gonads, MG: male gonads, Tail: most posterior part of adult tail without dorsal ocelli. (D) Heat map displaying expression of opsin genes in various parts of *B. lanceolatum* adult body. Opsin expression was detected by qRT-PCR and normalized to expression of actin. Each row represents particular opsin gene expression in various parts of amphioxus adult body. Blue color represents expression below row average, white color represents average row expression, red color expression above row average.

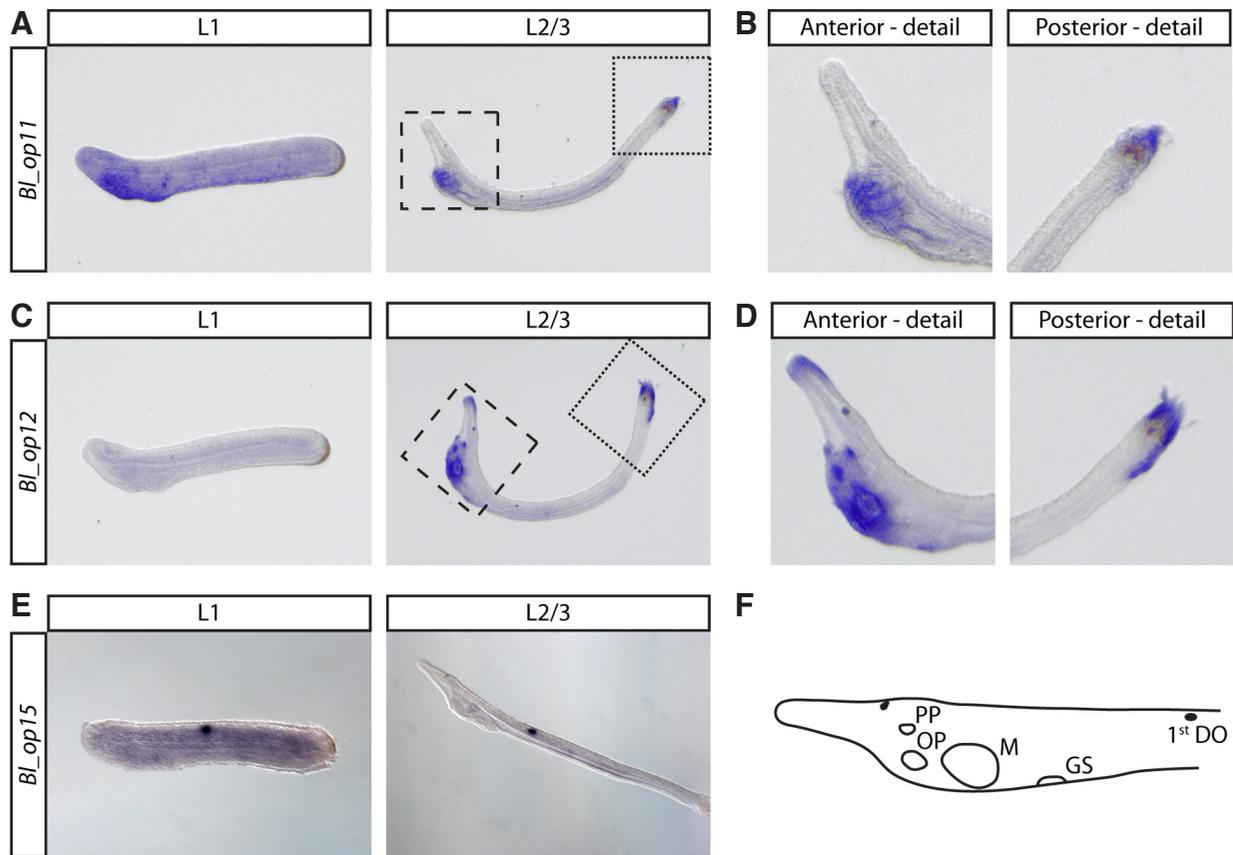


Fig. 4. *In situ* hybridization analysis of *B. lanceolatum* *BL_op11*, *BL_op12a* and *BL_op15* in developing larvae (L1 and L2/3 stages). **(A)** Expression of *BL_op11*. In stage L1 the signal was detected in the area of developing preoral pit and 1st gill slit. In L2/3 the signal was detected in pharyngeal region and tail fin. **(B)** Detail of anterior and posterior region of the larvae (dashed and dotted framed regions in (A)). **(C)** Expression of *BL_op12a*. In L1 stage no specific signal was detected. In L2/3 the signal was detected in preoral pit, oral papilla, cells around mouth, 1st gill slit and tail fin. **(D)** Detail of anterior and posterior region of the larvae (dashed and dotted framed regions shown in (C)). **(E)** Expression of *BL_op15*. Specific signal was detected in 1st dorsal ocelli. **(F)** Scheme of anterior part of L2/3 larvae with marked landmarks. PP, preoral pit; OP, oral papilla; M, mouth; GS, gill slit; 1st DO, first dorsal ocelli.

lanceolatum opsins in the three of the four traditionally recognized opsin lineages, in particular C-type, R-type and Group4 (neuropeptides, RGRs and peropsins), with no differences between the two sequence datasets used in this study. The number of genes in each subfamily, including a putative pseudogene, varies from one (melanopsins and peropsins) to a maximum of seven (Amphiop6). The expansion of three opsin subfamilies in the cephalochordate lineage (Amphiop6, C-type and Go-opsins) was previously (Holland *et al.*, 2008) correlated to the diversity of photoreceptor organs in amphioxus, including both ciliary and rhabdomeric photoreceptors (Lacalli, 2004). The large size of the Go group suggests extensive redundancy and/or could be an indication of possible fine tuning of molecular properties among these opsins in order to achieve distinct photoreceptive properties.

In general, *B. lanceolatum* Go opsins exhibit a preferential expression in either the cerebral vesicle or the neural tube (parts of the Central Nervous System - CNS), with two interesting exceptions: *BL_op13a* in male gonads and *BL_op9-op10* in tail. A CNS-specific pattern is observed for the C-type opsins as well, with *BL_op2*, however, being expressed almost solely in the tail. Similar is the expression in the horseshoe crab *Limulus polyphemus*, where various opsins have been detected in the tail (Battelle *et al.*, 2016), proposed to be a circadian photoreceptor organ. Photoreceptors

located in the tail mediate light avoidance in larval lampreys (Binder and McDonald, 2008). Light sensitivity in the tail of amphioxus has been reported as early as 1908 (Parker, 1908), yet no study so far has documented any photoreceptor cells there. In regard to opsin expression in gonads, it has also been observed in the gonads of the cnidarian species *Cladonema radiatum* (Suga *et al.*, 2008) and *Tripedalia cystophora* (Liegertova *et al.*, 2015), and the oyster *Crassostrea gigas* (Porath-Krause *et al.*, 2016). The role in opsin-mediated light-controlled release of gametes was proposed (Liegertova *et al.*, 2015; Suga *et al.*, 2008). In agreement with data regarding the *B. belcheri* melanopsin (Koyanagi *et al.*, 2005), an orthologous *BL_op15* is expressed in the neural tube where rhabdomeric dorsal ocelli are located. Nevertheless, it is not the unique opsin expressed in this tissue. The *B. lanceolatum* peropsin (*BL_op14*), two C-type opsins, namely *BL_op3* (ortholog of *B. belcheri* Amphiop4) and *BL_op5* (not previously analyzed in *B. belcheri*), neuropeptide (*BL_op8*) and two Amphiop6 opsins (*BL_op18* and *BL_op21*) were all found to be expressed in the neural tube of adult *B. lanceolatum*.

Examination of amino acid residues at key positions (namely, the counterion and the tripeptide on the fourth cytoplasmic loop) may provide insight into molecular function of individual opsins. Nevertheless, there are instances of a non-typical pattern in re-

gard to these positions (Plachetzki *et al.*, 2007). The presence of negative residues at position 113, typical for vertebrate opsins, is also evident in three *Branchiostoma* (this study) as well as in four *Tripedalia* (Liegertova *et al.*, 2015) opsins. Among them is BI_Op11, which further lacks a negative residue at position 181, thus raising questions about its role and expression domain. This opsin could also provide some insight in regard to the timepoint that counterion replacement occurred during the molecular evolution of vertebrate opsins (Terakita *et al.*, 2004). Slightly differentiated from the rest of C-type opsins is BI_Op5, for which the sole negatively charged residue is located at position 83. Another thing that should be highlighted is the absence of a negatively charged residue in any of the three putative counterion positions of the Amphip6 BI_Op21. The type of signaling cascade activated by various opsins is summarized by Porter *et al.*, 2012 and an *in silico* attempt has been made to correlate the motif present in the N-terminal of the fourth cytoplasmic loop (a.k.a. the tripeptide) with the target G-proteins. As shown before, the NKQ motif of the rhodopsin holds an important role in the activation of the G-protein transducin (Marin *et al.*, 2000). In regard to this position, some of our results are rather conventional, but some should attract more attention in the future. For example, BI_Op5 again stands out from the rest of the C-type opsins, since it contains the rather unique EKE motif. Amphip6 genes possess only the central proline of the R-opsin HPK fingerprint, therefore their ability to couple to any downstream phototransduction cascades remains an open question. Apart from the conserved proline, there is no clear pattern for the tripeptide in Amphip6 subgroup, except for the fact that a non-polar aliphatic residue occupies the first position. This group clusters closely with Gq opsins, however its members also differ in regard to motifs characteristic of Gq opsins, necessary for structural integrity maintenance and binding to the chromophore (Porath-Krause *et al.*, 2016).

In summary, our genome-wide analysis identifies the complement of opsin genes in *B. lanceolatum*, confirms expression of all but one genes and documents their stage and/or tissue specificity. Studies of opsin diversity can offer clues to how separate lineages of animals have repurposed different opsin paralogs for different light-detecting functions. To gain a deeper insight into the function of amphioxus photoreceptive organs, more detailed expression analysis of individual opsins (e.g. by *in situ* hybridization or immunohistochemistry staining) in conjunction with light-mediated behavioral tests of animals is of key importance. In addition, studies aimed to dissect the biochemical properties of individual amphioxus opsins and the nature of the downstream phototransduction cascade are highly warranted. Such further studies may provide evidence for fine tuning of molecular properties within the pool of available opsins that were necessary to adapt visually guided behavior of amphioxus to changing habitats.

Materials and Methods

Genome assembly and de novo gene prediction

B. lanceolatum opsin genomic loci were detected through tBLASTn searches against the European amphioxus (v. BI71nemh 20/11/13) Assembly (Bralan2). The previously characterized (Holland *et al.*, 2008; Koyanagi *et al.*, 2005; Koyanagi *et al.*, 2002) *Branchiostoma* opsin homologs were used as queries. Additional BLAST searches were performed using various visual and non-visual homologs from vertebrate and protostome species. *De novo* prediction of *B. lanceolatum* opsin genes was accomplished combining

results from Genscan (Burge and Karlin, 1997) and exon-intron borders predictions by SpliceView (Rogozin and Milanese, 1997). In the case of discrepancies between database gene models and our *in silico* analysis, PCR amplification of the “suspicious” regions was performed, followed by cloning and sequencing (see paragraph “Cloning and Sequencing of Opsin Gene Fragments/Transcripts”).

Prediction of membrane protein topology and functional domains

Newly identified *B. lanceolatum* opsin homologs were run in ScanProsite (de Castro *et al.*, 2006), in order to identify protein family domains/motifs, more specifically “the G Protein Coupled Receptor (GPCR) signature/profile” and the “Visual pigments (opsins) retinal binding site”. In addition, TOPCONS (Tsirigos *et al.*, 2015) and CCTOP (Dobson *et al.*, 2015) were used to detect the protein topology in general and the exact position of the seven transmembrane (TM) helices, characteristic of the GPCRs in particular. In order for the proteins to be considered reliable opsin homologs they had to meet the following three criteria: 1) exhibit similarity to known opsins, 2) bear seven TM domains and 3) possess a lysine residue at the seventh TM domain.

Sequence collection, alignments and phylogenetic analysis

Given the fact that the datasets used in the most recent large scale analyses (D’Aniello *et al.*, 2015; Liegertova *et al.*, 2015; Porter *et al.*, 2012; Ramirez *et al.*, 2016) had included significant number of opsins representative of a large number of taxonomic groups, we first tried to combine the available datasets from them. Sequences of poor quality (large gaps, missing the K296, not bearing all seven TM domains) were omitted, predicted *de novo* or replaced by orthologs from relative species. We then enriched this dataset with the 21 newly identified homologs from *B. lanceolatum*.

Multiple sequence alignments were produced with the Clustal algorithm, incorporated in MEGA v7 (Kumar *et al.*, 2016), and PROMALS3D (Pei *et al.*, 2008). The latter constructs alignments for multiple protein sequences and/or structures using information from sequence database searches, secondary structure prediction, available homologs with 3D structures and user-defined constraints (Pei *et al.*, 2008), therefore, it should be more reliable in the case of such a diverse group as opsins. Ambiguously aligned regions of the sequences, i.e. parts of the N-terminal and C-terminal ends were trimmed off in the MEGA7 alignment editor, leaving only the TM domains and the connecting extra-cellular and cytoplasmic loops. TOPCONS and CCTOP predictions were taken into consideration during the trimming process. We then created another subset in which the third cytoplasmic loop was removed from the alignment; the size of this loop is rather conserved within members of the same subfamily but it is highly variable between different families; it can fluctuate from 18aa in vertebrate RGRs, to 20aa in molluscan retinochromes, 30aa in human short-wave opsins and up to 73aa in echinopsins B. High degree of sequence dissimilarity is also observed, which renders alignment rather problematic.

Maximum Likelihood (ML) trees were constructed for both datasets. LG, gamma distributed (Le and Gascuel, 2008) was indicated as the best substitution model by the “Find Best DNA/Protein Models” tool incorporated in MEGA7 (Kumar *et al.*, 2016). ML trees were built in MEGA7 and tree topologies were evaluated with the bootstrap test (100 pseudoreplicates). *B. lanceolatum* sequences are included in Supplementary file 1, accession numbers (ACs) for the rest of the protein sequences (or genomic scaffold in case of newly predicted sequences) used in the phylogenetic analysis are included in Supplementary Fig. 1. All trees were rooted with the 22 human non-opsin GPCRs used in previous studies as well (Liegertova *et al.*, 2015).

Animal collection

B. lanceolatum adult animals were collected in Argeles-sur-Mer (France) and preserved in the lab in a day/night cycle of 14h/10h until spawning, which was induced by a shift in temperature (Fuentes *et al.*, 2007). Staging of all collected embryos was performed according to Hirakow and Kajita (1994). Specimens from late neurula (N3), larvae (L1-L2/3) and adult stage were collected and frozen in RNAlater® Stabilization Solution (ThermoFisher

Scientific), under light conditions. In N3, 1st Hesse cell is present, in L1 the developing frontal eye and lamellar body are present, in L2/3 the 1st gill slit and the mouth are open, frontal eye and lamellar body have differentiated. In adult, frontal eye, rudiments of lamellar body, Joseph cells and series of dorsal ocelli are present.

B. lanceolatum adults were anesthetized with Tricaine methane sulfonate and dissected in order to obtain the required tissue types, specifically the cerebral vesicle, the most anterior third of neural tube including Joseph cells and dense clusters of dorsal ocelli, the female and the male gonads and finally the most posterior part of tail without dorsal ocelli. Tissues were stored in RNAlater® Stabilization Solution (ThermoFisher Scientific) and frozen.

RNA isolation / cDNA preparation

Total RNA was isolated from *B. lanceolatum* embryos or adult tissues stored in RNAlater® Stabilization Solution using the Trizol reagent (Ambion). To avoid genomic DNA contamination, isolated RNA was treated with DNaseI and purified on RNeasy Mini Kit (Qiagen) column. Random-primed cDNA was prepared from 250 ng or 150 ng of RNA (for analysis of the different developmental stages or adult tissue types, respectively) in a 20 µl reaction using SuperScript VILO cDNA Synthesis kit (Invitrogen).

Cloning and sequencing of opsin gene fragments/transcripts

For validation of the *in silico* predicted gene models, cloning of opsin gene fragments and complete transcripts from *B. lanceolatum* was performed. Primers were designed in Primer3 software (Supplementary Table 1). PCR was performed on cDNA from neurula N3, larvae L2/3 or adult stages, using DreamTaq polymerase (ThermoFisher Scientific) according to the manufacturer's protocol. Amplified fragments were blunted and cloned into pCR™-Blunt II-TOPO® (ThermoFisher Scientific). Plasmid DNA was isolated using GeneJET Plasmid Miniprep Kit (ThermoFisher Scientific). Sequencing was performed by the standard Sanger sequencing procedure (GATC Biotech).

qRT-PCR

Primers were designed in Primer3 software (Supplementary Table 1). The qRT-PCR was performed in the LightCycler 2.0 System using the LightCycler® 480 DNA SYBR Green I Master kit (Roche Diagnostics, Germany), according to the manufacturer's standard protocol. For each cDNA sample, both target and housekeeping genes were measured under the same conditions. Results were analyzed using the LightCycler software. Crossing point values (Cp) were determined as an average of triplicates for each gene and normalized by Cp values of the housekeeping gene (actin). Results were analyzed in R software and plotted in the form of a Z-score heat map.

Whole-mount in situ hybridization in amphioxus embryos

Whole-mount *in situ* hybridization with digoxigenin-labeled RNA probes was done as previously described (Kozmikova *et al.*, 2013). For better interpretation of the signal, *in situ* hybridization with Vector Blue substrate was performed in some cases, followed by subsequent confocal microscopic analysis. Primers used to generate probes are summarized in Supplementary Table 1.

Acknowledgments

This work was funded by the Czech Science Foundation (GACR, 17-15374S); Ministry of Education, Youth and Sports (LO1220 CZ-OPENSREEN, LQ1604 NPU II, BIOCEV-CZ.1.05/1.1.00/02.0109) and the Institute of Molecular Genetics institutional support (RVO 6878050). This work was supported by the Branchiostoma lanceolatum genome consortium that provided access to the *B. lanceolatum* genome sequence.

References

ALBALAT R (2012). Evolution of the genetic machinery of the visual cycle: a novelty of the vertebrate eye? *Mol Biol Evol* 29: 1461-1469.

ARENDR 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: 869-871.

BATTELLE B A, RYAN J F, KEMPLER K E, SARAF S R, MARTEN C E, WARREN W C, MINX P J, MONTAGUE M J, GREEN P J, SCHMIDT S A *et al.*, (2016). Opsin repertoire and expression patterns in horseshoe crabs: evidence from the genome of *Limulus polyphemus* (Arthropoda: Chelicerata). *Genome Biol Evol* 8: 1571-1589.

BINDER T R, MCDONALD D G (2008). The role of dermal photoreceptors during the sea lamprey (*Petromyzon marinus*) spawning migration. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 194: 921-928.

BOWNDSD D (1967). Site of attachment of retinal in rhodopsin. *Nature* 216: 1178-1181.

BURGE C, KARLIN S (1997). Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 268: 78-94.

D'ANIELLO S, DELROISSE J, VALERO-GRACIA A, LOWE E K, BYRNE M, CANNON J T, HALANYCH K M, ELPHICK M R, MALLEFET J, KAUL-STREHLOW S *et al.*, (2015). Opsin evolution in the Ambulacraria. *Mar Genomics* 24, Part 2: 177-183.

DE CASTRO E, SIGRIST C J A, GATTIKER A, BULLIARD V, LANGENDIJK-GENEVAUX P S, GASTEIGER E, BAIROCHA, HULO N (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* 34: W362-W365.

DOBSON L, REMENYI I, TUSNADY G E (2015). CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Res* 43: W408-412.

ETANI K, NISHIKATA T (2002). Novel G-protein-coupled receptor gene expressed specifically in the entire neural tube of the ascidian *Ciona intestinalis*. *Dev Genes Evol* 212: 447-451.

FEUDA R, HAMILTON S C, MCINERNEY J O, PISANI D (2012). Metazoan opsin evolution reveals a simple route to animal vision. *Proc Natl Acad Sci USA* 109: 18868-18872.

FUENTES M, BENITO E, BERTRAND S, PARIS M, MIGNARDOT A, GODOY L, JIMENEZ-DELGADO S, OLIVERI D, CANDIANI S, HIRSINGER E *et al.*, (2007). Insights into spawning behavior and development of the European amphioxus (*Branchiostoma lanceolatum*). *J Exp Zool B Mol Dev Evol* 308: 484-493.

GONG J, YUAN Y, WARD A, KANG L, ZHANG B, WU Z, PENG J, FENG Z, LIU J, XU X Z (2016). The *C. elegans* Taste Receptor Homolog LITE-1 Is a Photoreceptor. *Cell* 167: 1252-1263 e1210.

HIRAKOW R, KAJITA N (1994). Electron microscopic study of the development of amphioxus, *Branchiostoma belcheri tsingtauense*: the neurula and larva. *Kaibogaku Zasshi* 69: 1-13.

HOLLAND L Z, ALBALAT R, AZUMIK, BENITO-GUTIERREZ E, BLOW M J, BRONNER-FRASER M, BRUNET F, BUTTS T, CANDIANI S, DISHAW L J *et al.*, (2008). The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res* 18: 1100-1111.

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: 1065-1069.

KOYANAGI M, TAKANO K, TSUKAMOTO H, OHTSU K, TOKUNAGA F, TERAKITA A (2008). Jellyfish vision starts with cAMP signaling mediated by opsin-G(s) cascade. *Proc Natl Acad Sci USA* 105: 15576-15580.

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: 525-528.

KOZMIK Z, RUZICKOVA J, JONASOVA K, MATSUMOTO Y, VOPALENSKY P, KOZMIKOVA I, STRNAD H, KAWAMURA S, PIATIGORSKY J, PACES V *et al.*, (2008). Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proc Natl Acad Sci USA* 105: 8989-8993.

KOZMIKOVA I, CANDIANI S, FABIAN P, GURSKA D, KOZMIK Z (2013). Essential role of Bmp signaling and its positive feedback loop in the early cell fate evolution of chordates. *Dev Biol* 382: 538-554.

KUMAR S, STECHER G, TAMURA K (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33: 1870-1874.

LACALLI T C (2004). Sensory systems in amphioxus: a window on the ancestral chordate condition. *Brain Behav Evol* 64: 148-162.

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: 960-976.

LE S Q, GASCUEL O (2008). An improved general amino acid replacement matrix.

- Mol Biol Evol* 25: 1307-1320.
- LIEGERTOVAM, PERGNER J, KOZMIKOVAI, FABIAN P, POMBINHOAR, STRNAD H, PACES J, VLCEK C, BARTUNEK P, KOZMIK Z (2015). Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution. *Sci Rep* 5: 11885.
- MARIN E P, KRISHNA A G, ZVYAGA T A, ISELE J, SIEBERT F, SAKMART T P (2000). The amino terminus of the fourth cytoplasmic loop of rhodopsin modulates rhodopsin-transducin interaction. *J Biol Chem* 275: 1930-1936.
- MASON B, SCHMALE M, GIBBS P, MILLER M W, WANG Q, LEVAY K, SHESTOPALOV V, SLEPAK V Z (2012). Evidence for Multiple Phototransduction Pathways in a Reef-Building Coral. *PLoS One* 7: e50371.
- PARKER G H (1908). The sensory reactions of amphioxus. *Proc Am Acad Arts Sci* 43: 415-455.
- PEI J, KIM B H, GRISHIN N V (2008). PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* 36: 2295-2300.
- PEIRSON S N, HALFORD S, FOSTER R G (2009). The evolution of irradiance detection: melanopsin and the non-visual opsins. *Phil Trans R Soc B* 364: 2849-2865.
- PLACHETZKI D C, DEGNAN B M, OAKLEY T H (2007). The origins of novel protein interactions during animal opsin evolution. *PLoS One* 2: e1054.
- PORATH-KRAUSE A J, PAIRETT A N, FAGGIONATO D, BIRLA B S, SANKAR K, SERB J M (2016). Structural differences and differential expression among rhabdomeric opsins reveal functional change after gene duplication in the bay scallop, *Argopecten irradians* (Pectinidae). *BMC Evol Biol* 16: 250.
- PORTER M L, BLASIC J R, BOK M J, CAMERON E G, PRINGLE T, CRONIN T W, ROBINSON P R (2012). Shedding new light on opsin evolution. *Proc Biol Sci* 279: 3-14.
- POSS S G, BOSCHUNG H T (1996). Lancelets (Cephalochordata: Branchiostomata): How many species are valid? *Isr J Zool* 42: S13-S66.
- RAMIREZ M D, PAIRETT A N, PANKEY M S, SERB J M, SPEISER D I, SWAFFORD A J, OAKLEY T H (2016). The last common ancestor of most bilaterian animals possessed at least 9 opsins. *Genome Biol Evol* 8: 3640-3652.
- RIVERA A S, OZTURK N, FAHEY B, PLACHETZKI D C, DEGNAN B M, SANCAR A, OAKLEY T H (2012). Blue-light-receptive cryptochrome is expressed in a sponge eye lacking neurons and opsin. *J Exp Biol* 215: 1278-1286.
- ROGOZIN I B, MILANESIL (1997). Analysis of donor splice sites in different eukaryotic organisms. *J Mol Evol* 45: 50-59.
- SHICHIDA Y, MATSUYAMAT (2009). Evolution of opsins and phototransduction. *Phil Trans R Soc B* 364: 2881-2895.
- SUGA H, SCHMID V, GEHRING W J (2008). Evolution and functional diversity of jellyfish opsins. *Curr Biol* 18: 51-55.
- TERAKITAA (2005). The opsins. *Genome Biol* 6: 213.
- TERAKITAA, KOYANAGI M, TSUKAMOTO H, YAMASHITA T, MIYATA T, SHICHIDA Y (2004). Counterion displacement in the molecular evolution of the rhodopsin family. *Nat Struct Mol Biol* 11: 284-289.
- TSIRIGOS K D, PETERS C, SHU N, KALL L, ELOFSSON A (2015). The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res* 43: W401-W407.
- VOPALENSKY P, PERGNER J, LIEGERTOVAM, BENITO-GUTIERREZ E, ARENDT D, KOZMIK Z (2012). Molecular analysis of the amphioxus frontal eye unravels the evolutionary origin of the retina and pigment cells of the vertebrate eye. *Proc Natl Acad Sci USA* 109: 15383-15388.
- WANG J K, MCDOWELL J H, HARGRAVE P A (1980). Site of attachment of 11-cis-retinal in bovine rhodopsin. *Biochemistry* 19: 5111-5117.
- YARFITZ S, HURLEY J B (1994). Transduction mechanisms of vertebrate and invertebrate photoreceptors. *J Biol Chem* 269: 14329-14332.

Further Related Reading, published previously in the *Int. J. Dev. Biol.*

From the American to the European amphioxus: towards experimental Evo-Devo at the origin of chordates

Jordi Garcia-Fernández, Senda Jiménez-Delgado, Juan Pascual-Anaya, Ignacio Maeso, Manuel Irimia, Carolina Minguillón, Èlia Benito-Gutiérrez, Josep Gardenyes, Stéphanie Bertrand and Salvatore D'Aniello
Int. J. Dev. Biol. (2009) 53: 1359-1366
<https://doi.org/10.1387/ijdb.072436jg>

Evolution of CUT class homeobox genes: insights from the genome of the amphioxus, *Branchiostoma floridae*

Naohito Takatori and Hidetoshi Saiga
Int. J. Dev. Biol. (2008) 52: 969-977
<https://doi.org/10.1387/ijdb.072541nt>

Peter Holland, homeobox genes and the developmental basis of animal diversity

Sebastian M. Shimeld
Int. J. Dev. Biol. (2008) 52: 3-7
<https://doi.org/10.1387/ijdb.072394ss>

Developmental expression of the High Mobility Group B gene in the amphioxus, *Branchiostoma belcheri tsingtauense*

Xiangwei Huang, Lifeng Wang and Hongwei Zhang
Int. J. Dev. Biol. (2005) 49: 49-46
<http://www.intjdevbiol.com/web/paper/041915xh>

Cell morphology in amphioxus nerve cord may reflect the time course of cell differentiation

T C Lacalli
Int. J. Dev. Biol. (2000) 44: 903-906
<http://www.intjdevbiol.com/web/paper/11206331>

Embryonic development of heads, skeletons and amphioxus: Edwin S. Goodrich revisited

P W Holland
Int. J. Dev. Biol. (2000) 44: 29-34
<http://www.intjdevbiol.com/web/paper/10761843>

Amphioxus Hox genes: insights into evolution and development

J Garcia-Fernández and P W Holland
Int. J. Dev. Biol. (1996) 40: S71-S72
<http://www.intjdevbiol.com/web/paper/9087701>

5 yr ISI Impact Factor (2013) = 2.879

