Marine Genomics 23 (2015) 109–121

Contents lists available at ScienceDirect

Marine Genomics

journal homepage: www.elsevier.com/locate/margen

De novo transcriptome of the European brittle star *Amphiura filiformis* pluteus larvae

Jérôme Delroisse [a,*], Olga Ortega-Martinez [b], Sam Dupont [b], Jérôme Mallefet [c], Patrick Flammang [a,*]

[a] University of Mons — UMONS, Research Institute for Biosciences, Biology of Marine Organisms and Biomimetics, 23 Place du Parc, 7000 Mons, Belgium
[b] University of Gothenburg, Department of Biological and Environmental Science, The Sven Lovén Centre for Marine Sciences, Kristineberg, 45178 Fiskebäckskil, Sweden
[c] Catholic University of Louvain-La-Neuve, Marine Biology Laboratory, Place croix du Sud, Louvain-La-Neuve, Belgium

ARTICLE INFO

Article history:
Received 21 April 2015
Received in revised form 19 May 2015
Accepted 21 May 2015
Available online 2 June 2015

Keywords:
Transcriptome
Echinodermata
Larval stage
Paired-end Illumina Sequencing
Photoreception
RGR opsin

ABSTRACT

Background: In non-classical model species, Next Generation Sequencing increases the ability to analyze the expression of transcripts/genes. In this study, paired-end Illumina HiSeq sequencing technology has been employed to describe a larval transcriptome generated from 64 h post-fertilization pluteus larvae of the brittle star *Amphiura filiformis*. We focused our analysis on the detection of actors involved in the opsin based light perception, respectively the opsins and the phototransduction actors.
Methods & results: In this research, about 47 million high quality reads were generated and 86,572 total unigenes were predicted after de novo assembly. Of all the larval unigenes, 18% show significant matches with reference online databases. 46% of annotated larval unigenes were significantly similar to transcripts from the purple sea urchin. COG, GO and KEGG analyses were performed on predicted unigenes. Regarding the opsin-based photoreception process, even if possible actors of ciliary and rhabdomeric phototransduction cascades were detected, no ciliary or rhabdomeric opsin was identified in these larvae. Additionally, partial non-visual RGR (retinal G protein coupled receptor) opsin mRNAs were identified,possibly indicating the presence of visual cycle reaction in early pluteus larvae. The eye morphogene Pax 6 was also identified in the pluteus transcriptome.
Conclusions: Contrary to sea-urchin larvae, brittle star larvae appear to be characterized by an absence of visual-like opsins. These RNA-seq data also provide a useful resource for the echinoderm research community and researchers with an interest in larval biology.

1. Introduction

The initial draft of the genome of the sea urchin *Strongylocentrotus purpuratus* was completed in 2006 (Sodergren et al., 2006) and is still the only complete echinoderm genome available in GenBank databases even though multiple genome projects are currently under process or achievement (http://www.echinobase.org/Echinobase/). Echinoderm classes separated 500 million years ago (Smith et al., 1993) and for that reason molecular data limited to one echinoderm class or to a few model species are not representative of other echinoderm classes with various ways-of-life. For emerging model marine organisms, NGS technologies offer a great opportunity for rapid access to genetic information (Wang et al., 2009). Proof of this is that several echinoderm transcriptomes recently emerged in the literature, bringing crucial molecular information on various biological processes such as development, regeneration or adhesion (Zhou et al., 2014; Vaughn et al., 2012; Burns et al., 2013; Gillard et al., 2014; Mashanov et al., 2014; Wygoda et al., 2014; Dilly et al., 2014; Tu et al., 2014; Ma et al., 2014; Hennebert et al., 2014; Purushothaman et al., 2014). However, large-scale molecular data of non-model echinoderm species are still highly limited. *Amphiura filiformis* is a burrowing brittle star commonly found in muddy marine ecosystems of Europe (Rosenberg, 1995). In its environment, this easily accessible species is considered as a keystone species involved in sediment remodeling and aeration (Vopel et al., 2003) and is also an important food source for flatfishes and crayfishes (Duineveld and Van Noort, 1986; Baden et al., 1990). *A. filiformis* is characterized by astonishing biological capabilities such as, among other things, bioluminescence (Delroisse et al., 2014a; Herring, 1995), resistance to hypoxia (Vistisen and Vismann, 1997), and regeneration (Nilsson, 1999; Sköld and Rosenberg, 1996; Burns et al., 2011; Burns et al., 2012). The study of the high regeneration capabilities of *A. filiformis* involved the transcriptomic analysis of adult arm tissues (Purushothaman et al., 2014). *A. filiformis* has also the ability to react to environmental parameters such as food, current and light (Rosenberg and Lundberg, 2004) and, recently, genomic and transcriptomic data were used to investigate photoreception in this species (Delroisse et al., 2014b). Thirteen new opsin genes were identified in the draft genome of *A. filiformis* but only 3 of these opsins were

* Corresponding authors.
E-mail addresses: jerome.delroisse@umons.ac.be (J. Delroisse), olga.ortega-martinez@bioenv.gu.se (O. Ortega-Martinez), sam.dupont@bioenv.gu.se (S. Dupont), jerome.mallefet@uclouvain.be (J. Mallefet), patrick.flammang@umons.ac.be (P. Flammang).

http://dx.doi.org/10.1016/j.margen.2015.05.014
1874-7787/

retrieved from an adult arm transcriptome(Delroisse et al., 2014b). These findings raise the possibility that some of the other opsin genes could be specifically expressed during other developmental stages, in particular in pluteus larvae. In the sea urchins *S. purpuratus* and *Hemicentrotus pulcherrimus*, opsin expression has been detected in pluteus larvae and are proposed to be linked to photosensitive larval swimming vertical migration (Burke et al., 2006; Raible et al., 2006; Ooka et al., 2010).

In this study, paired-end Illumina HiSeq™ 2000 sequencing technology has been employed to generate a 64 h post-fertilization pluteus larvae transcriptome in the brittle star *A. filiformis* (see Fig. 1 for further explanations on the reproductive cycle of *A. filiformis*). In this indirect developer, the pluteus stage – the longest pelagic stage and only larval stage – was considered the most representative stage of the premetamorphic development (Dupont et al., 2009). The use of this early pluteus stage – not yet fed – moreover permitted to avoid algae contamination. To the best of our knowledge, it is the first brittle star pluteus transcriptome described in the literature. Considering the paucity of molecular data in brittle stars, RNA-seq of the ophiopluteus larvae would be valuable resources to better understand developmental and biological mechanisms and should greatly impact future molecular studies on *A. filiformis* but also on brittle stars in general. In the context of environment perception and, particularly, light perception processes, we searched for putative ciliary and rhabdomeric phototransduction actors in the larval transcriptome.

2. Material & methods

2.1. Animal collection and RNA extraction

Adult individuals of *A. filiformis* were collected in the vicinity of the Sven Lovén Centre for Marine Sciences — Kristineberg (Fiskebäckskil, Sweden) in summer 2011 at a depth of 30 m. The brittle stars were carefully rinsed out of the sediment, and intact specimens were kept

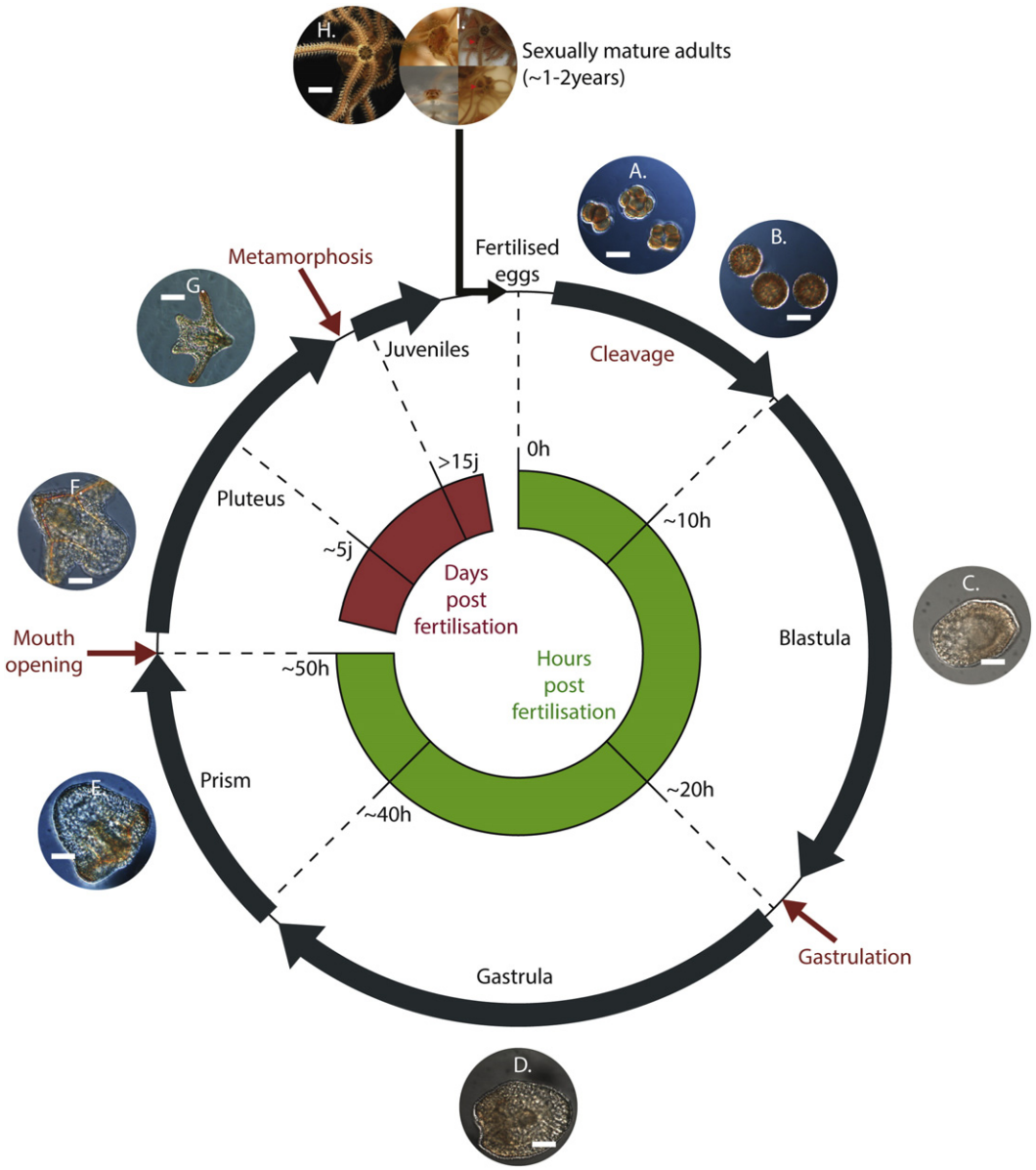

Fig. 1. Reproductive cycle of *Amphiura filiformis* at 14 °C. A–B: early embryos during cleavage, C: blastula, D: gastrula, E: prism, F: early pluteus, G: pluteus, H: adult, I: spawning (up: female, down: male, red arrow indicate bursal slits). The complete development of the brittle star *Amphiura filiformis* from the zygote to the juvenile take around 20 days at 14 °C however it is variable according to feeding and environmental conditions. The sexual maturity is reached after more than 1 year. Scale bars: A–F: 20 μm, G: 40 μm, H: 0.2 cm.

in sediment and running sea-water (14 °C, salinity 32, pH_T 8.0) pumped from 30 m water depth at a site close to where the sampling was performed. Following the protocol of Dupont et al. (2009), adult individuals with ripe gonads (white color for testes and orange color for ovaries) – visible through the extended wall of bursae – were used for fertilization experiments and larval culture (Fig. 1.I). Three males and 15 females were heat-shocked at 26 °C for 15 min and then kept in 1 l of filtered seawater in the dark until the release of sperm that induces the female to spawn. Fertilized eggs were rinsed in filtered seawater and at the two cell stage were transferred to 5 l aquaria at 14 °C, at a density of ten embryos per milliliter. Seawater was constantly aerated.

To harvest plutei (Fig. 1.F), water from the aquaria was delicately filtered, and the larvae were collected in a small water volume. This water sample containing high concentration of larvae was then centrifuged at low speed to pellet the larvae. After discarding the excess seawater, the pellet was treated with fresh TRIzol® solution and RNA extraction was directly performed using the RiboPure™ RNA extraction kit (Ambion AM1924). Extractions were performed according to the manufacturer's protocols. RNA extracts were quality checked by gel electrophoresis on a 1.2 M TAE agarose gel and by spectrophotometry using a Nanodrop spectrophotometer (LabTech International). The quality of the RNA was also assessed by size chromatography with an Agilent 2100 Bioanalyzer. Twenty microliters of the mRNA sample were used for the cDNA library preparation.

2.2. RNA-seq library preparation and sequencing

Libraries were prepared from the RNA pool by mRNA enrichment, fragment interruption, addition of adapters, size selection, PCR amplification, and RNA-Seq. The RNA sample was treated with DNase I to ensure that it was DNA-free. Subsequently the poly-(A) mRNAs were enriched using the oligo(dT) magnetic beads and then fragmented into short pieces (about 200 bp). Using these short fragments as templates, random hexamer-primers were used to synthesize first-strand cDNA. Second-strand cDNA was synthesized using DNA polymerase I. Double stranded cDNAs were purified and subjected to end reparation and 3′ single adenylation. Sequencing adaptors were then ligated to the adenylated fragments, which were then enriched by PCR amplification. High-throughput sequencing was then conducted using the Illumina HiSeq™ 2000 platform to generate 100-bp paired-end reads (Beijing Genomics Institute — BGI Shenzheng). Sequencing was performed according to the manufacturer's instructions (Illumina, San Diego, CA). The SRA (short read archive) raw reads have been deposited on GenBank public database under the accession number SRR1533125.

2.3. Data processing and de novo assembly

Transcriptome quality was checked using FastQC software (http://www.bioinformatics.babraham.ac.uk). Raw reads generally contain low quality reads and adaptor sequences. A preprocessing step of data cleaning is therefore required to obtain the “clean reads” that are used in the next steps of the analysis. The cleaning step, performed by BGI, include the adaptor removal as well as the application of a stringent filtering criterion to remove reads with more than 5% of unknown bases and low quality reads, i.e. reads that comprise more than 20% low-quality bases (base quality ≤ 10). Clean reads were used for the de novo assembly using Trinity software (Grabherr et al., 2011) (release-20121005) that comprises three independent software modules, namely Inchworm, Chrysalis and Butterfly. Inchworm assembles the reads into longer transcripts in order to form contigs, Chrysalis groups the contigs into clusters that represent the transcriptional complexity of the genes and Butterfly reports the full-length transcripts for alternatively spliced isoforms.

Unigenes can either form clusters in which the similarity among overlapping sequences is superior to 94%, or singletons that are unique unigenes. The family clustering was performed using the TIGR Gene Indices Clustering (TGICL) tools (Pertea et al., 2003) followed by Phrap assembler (http://www.phrap.org). Various quality assembly criteria were evaluated such as the size distribution of contigs and unigenes and the evaluation of the read distribution when realigned to unigenes (using SOAPaligner with the default settings, (Li et al., 2008)).

2.4. Unigene annotation and classification

Unigenes were subjected to database searches in NCBI non-redundant protein database (NR; http://www.ncbi.nlm.nih.gov), Swiss-Prot database (http://www.expasy.ch/sprot) (Bairoch and Boeckmann, 1991), Kyoto Encyclopedia of Genes and Genome (KEGG) (http://www.genome.jp/kegg) (Kanehisa and Goto, 2000; Kanehisa et al., 2008) and Cluster of Orthologous Groups (COG) of proteins (http://www.ncbi.nlm.nih.gov/COG) (Tatusov et al., 2001; Tatusov et al., 2003) using BLASTx with an E-value cutoff of 10^{-5} (Sotiriades and Dollas, 2007), and in the nucleic-acid database (NCBI NT) by BLASTn with the same E-value cutoff. When different databases returned inconsistent results, they were prioritized in the following order: NR, SwissProt, KEGG, COG. When a unigene did not align with any of the entries in these databases, ESTScan was used to predict potential coding regions and determine the direction of coding sequence in the unigene (Iseli et al., 1999).

Blast2GO (Conesa et al., 2005; Aparicio et al., 2006) was used to find GO annotations (http://www.geneontology.org) based on the NR annotation, and GO functional classification was calculated for all unigenes using WEGO software (Ye et al., 2006). This analysis mapped all the annotated unigenes to GO term under the categories of biological process, cellular component, and molecular function. The unigenes were also submitted to the COG database to predict and classify possible functions. We determined pathway annotations for unigenes using KEGG annotations.

2.5. Detection of opsin-based photoreception actors in the pluteus transcriptome of A. filiformis

Using a Blast approach, we searched for and identified putative homologs to proteins from the ciliary and rhabdomeric phototransduction pathways (Schnitzler et al., 2012; Tong et al., 2009; Fain et al., 2010; Yau and Hardie, 2009) in the pluteus transcriptome of *A. filiformis*. Phototransduction actor proteins (see Additional file 2) were used as queries in tBLASTn searches against the set of *A. filiformis* unigenes. A reciprocal best BLASTx search was then performed using the top unigenes against the NCBI non-redundant protein database. Blast hits with significant E-values strongly indicate homolog proteins. Additionally, Delroisse et al. (2014b) recently identified 13 opsin genes in a draft genome of *A. filiformis* (Af-opsin 1 [ciliary opsin, GenBank: KM276762], 2 [basal branch opsin, GenBank: KM276763], 3 [Go-opsin, GenBank: KM276764], 4.1 [rhabdomeric opsin, GenBank: KM276765], 4.2 [rhabdomeric opsin, GenBank: KM276766], 4.3 [rhabdomeric opsin, GenBank: KM276767], 4.4 [rhabdomeric opsin, GenBank: KM276768], 4.5 [rhabdomeric opsin, GenBank: KM276769], 4.6 [rhabdomeric opsin, GenBank: KM276770], 5 [basal branch opsin, GenBank: KM276771], 7.A [RGR opsin, GenBank: KM276772], 7.B [RGR opsin, GenBank: KM276773], 8.1 [neuropsin, GenBank: KM276774], 8.2 [neuropsin, GenBank: KM276775]). The genome of *A. filiformis* indeed contains non-rhabdomeric/ciliary opsins such as neuropsins, Go opsins and RGR opsins. Here, we specifically searched for these opsins in the pluteus transcriptome of *A. filiformis*. Local BLASTn and tBLASTn (2.2.26) searches were used to highlight the expression of these potential opsin genes.

The Pax6 gene, a transcription factor gene considered as a critical regulator of eye development in bilaterian animals (Gehring, 2002),

was here detected (mRNA) in the pluteus transcriptome using the previously mentioned BLAST approach.

2.6. Phylogenetic analyses of Opsins and Pax-6

In silico translation was performed on the opsin-like sequences retrieved from the pluteus transcriptome of *A. filiformis* using the online "translate tool" from Expasy (Gasteiger et al., 2005). A multiple amino-acid alignment of partial opsins sequences was performed on the predicted protein sequences using Seaview 4.2.12 (Gouy et al., 2010) and the MUSCLE algorithm (Edgar, 2004). For short sequences, the alignment was also corrected manually.

Putative opsin sequences retrieved from the transcriptome were included in phylogenetic analyses based on the previously mentioned MUSCLE alignment. The *A. filiformis* opsin sequences recently published (Delroisse et al., 2014b), were added to the analysis. Additionally, metazoan opsin sequence data were collected as references from open-access NCBI databases (http://www.ncbi.nlm.nih.gov) and are listed in the supplementary data (Additional file 2). The tree was constructed using truncated alignment/sequences mainly limited to the conserved 7TM core of the proteins. The phylogeny was constructed using the MrBayes software (Ronquist et al., 2012) based on (Feuda et al., 2012; Ullrich-Lüter et al., 2013; Feuda et al., 2014; Delroisse et al., 2014b). The GTR + G model – defined as the best-suited model for opsin phylogeny by (Feuda et al., 2012) – was used for the analyses. A non-opsin GPCR sequence was chosen as outgroup (i.e. melatonin receptor) following (Feuda et al., 2012). Three independent runs of 6,000,000 generations were performed reaching a standard deviation value inferior to 0.01. In parallel, a maximum likelihood opsin phylogeny was constructed using the PHYML tool (Guindon and Gascuel, 2003; Guindon et al., 2009) from SeaView 4.2.12 software (Gouy et al., 2010). A best-fit model analysis was performed using Mega v5.2.1 (following the Akaike information criteria) (Tamura et al., 2007; Kumar et al., 2008) and "Wheland and Goldman model of protein evolution" was found to be the best suited and was used for the analyses (WAG) (Whelan and Goldman, 2001). Branch support values were estimated as bootstrap proportions from 500 PhyML bootstrap replicates.

For Pax6 phylogeny estimation, in silico translation and alignment were performed following the same procedure. The tree was constructed using trimmed sequences mainly limited to the conserved paired domains and homeodomain of the protein based on (Omori et al., 2011) (see also (Blackburn et al., 2008)). Phylogenies were constructed using Maximum likelihood (PHYML tool, by default LG substitution model, (Le and Gascuel, 2008)), Parsimony and Distance analyses using SeaView 4.2.12 software (Gouy et al., 2010; Guindon and Gascuel, 2003; Guindon et al., 2009). Bootstrap values were calculated based on percentage of 100 (ML), 10,000 (MP) and 1,000,000 (NJ) iterations.

3. Results & discussion

3.1. cDNA sequence generation and de novo assembly

In order to generate the pluteus transcriptome of the brittle star *A. filiformis*, a cDNA library was produced from RNA isolated from 64 h larvae (see Fig. 1) and sequenced using paired-end Illumina HiSeq™ 2000 sequencing technology. In this research, 47,674,984 raw reads with the length of 100 bp were generated from a 200 bp insert library. After removing low-quality regions, adaptors and possible contaminants, the remaining 43,655,676 clean reads were used to assemble the pluteus transcriptome with Trinity software (Grabherr et al., 2011). According to the overlapping information of high-quality reads, 336,466 contigs were generated with an average length of 184 bp. The N50 (median contig size) was of 193 bp and the Q20 percentages (base quality more than 20) were superior to 97%. N percentage which indicates the percentage of nucleotides which could not be sequenced was estimated to 0%. The dataset of raw reads was deposited in NCBI SRA database under the accession number SRR1533125. Using paired-end joining and gap filling, contigs were further assembled into 86,572 unique sequences (non redundant sequences or unigenes) with a mean length of 422 bp including 27,305 clusters and 59,267 singletons. Numerical data are summarized in Table 1. To evaluate the quality of the assembled unigenes, all the usable sequencing reads were realigned to the unigenes and more than 85% of larval transcriptome unigenes were realigned by more than 5 reads (Fig. 2). Length distribution of contigs and unigenes are presented in Fig. 3.

Table 1
Data description of the *Amphiura filiformis* pluteus transcriptome A. Description of the sequencing output data for *A. filiformis* pluteus larvae. "Q20 percentage" is proportion of nucleotides with quality value larger than 20 in reads. "N percentage" is the proportion of unknown nucleotides in clean reads. "GC percentage" is the proportion of guanidine and cytosine nucleotides among total nucleotides. B. Summary statistics of transcriptome assembly for *A. filiformis* larvae.

A.		
	Total raw reads	47,674,984
	Total clean reads	43,655,676
	Total clean nucleotides (nt)	4,365,567,600
	Q20 percentage	97.28%
	N percentage	0.00%
	GC percentage	38.35%

B.	Contigs	Unigenes
Total number	336,466	86,572
Total length (bp)	62,072,294	36,501,163
Mean length (bp)	184	422
N50	193	451
Distinct clusters		27,305
Distinct singletons	/	59,267

3.2. General gene annotation

Unigene sequences were annotated using BLASTx to NCBI protein databases (NR), Swiss-Prot, KEGG and COG with a cut-off E-value of $1e^{-05}$. Unigenes were also annotated using BLASTn to NCBI nucleotide databases (NT) with a cut-off E-value of $1e^{-05}$. On the 86,572 unigenes, 15,941 show significant matches (18.4%; Fig. 4): 13,722 to NR (15.8%), 5894 to NT (6.8%), 10,601 to Swiss-Prot (12.2%), 9529 to KEGG (11%), 4357 to COG (5%) and to 5862 GO (6.8%) (see Fig. 4.A). This level of sequence similarity matching is low but comparable to those found in other studies (Vera et al., 2008; Hahn et al., 2009; Wang et al., 2010; Franchini et al., 2011; Clark et al., 2011; Rismani-Yazdi et al., 2011; Tao et al., 2012), including on echinoderm species (Zhou et al., 2014; Burns et al., 2013; Mashanov et al., 2014; Du et al., 2012), in which high throughput sequencing technology was used for the de novo transcriptome assembly of non-model species. The main reason for this is probably the lack of large-scale genomic resources for the genus *Amphiura* and other evolutionary related ophiuroids. Moreover short-sized unigenes can increase the difficulty of gene identification.

The first three hits in terms of species were *S. purpuratus*, *Saccoglossus kowalevskii* and *Branchiostoma floridae*, as expected from their phylogenetical proximity with *A. filiformis* (Fig. 4.B). On the 13,722 matches, 6377 unigenes (46%) were significantly similar to transcripts from the purple sea urchin *S. purpuratus*. The annotation success was estimated by ranking the annotation E-value and identity results with the NR database comparison. The E-value distribution is presented in Fig. 4.C. This E-value distribution of the top matches in the NR database showed that more than 35% of the mapped sequences have strong homology (E-value smaller than $1.0e^{-30}$), whereas 65% of the homologous sequences presented E-values ranging from $1.0e^{-05}$ to $1.0e^{-30}$ (Fig. 4.C). The sequence similarity distribution indicates that 8.8% of the sequences have a similarity higher than 80% and 31.9% of the sequences have a similarity higher than 60% (Fig. 4.D).

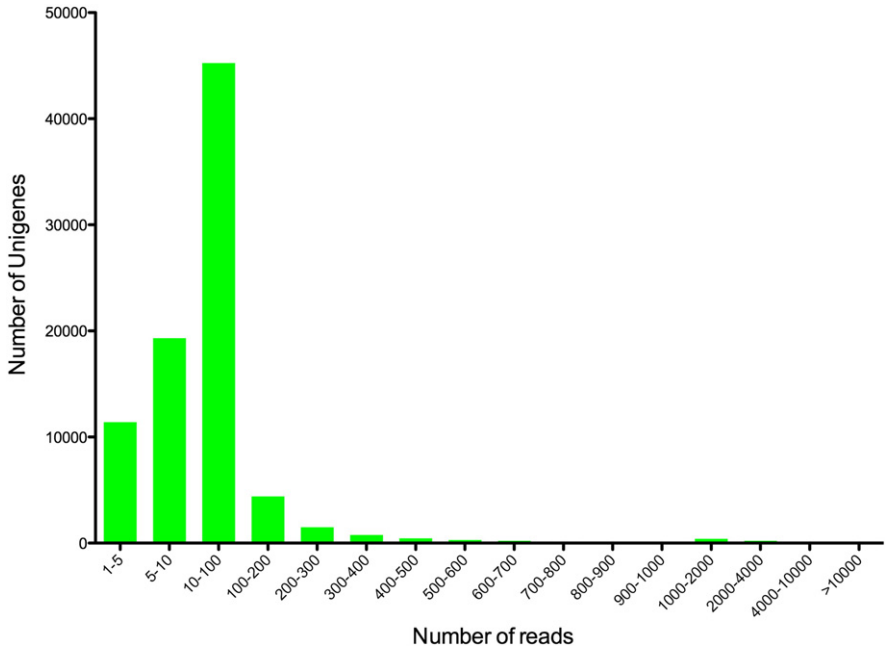

Fig. 2. Assessment of assembly quality. Distribution of unique mapped reads on the assembled unigenes. The x axis represents the "number of reads" classes. The y-axis represents the number of unigenes.

3.3. Gene Ontology annotation

On the basis of the NR annotation, GO functional classifications of the unigenes were performed. In total, 5862 unigenes with BLAST matches to known proteins were assigned to GO classes with a total of 48,078 functional terms (Fig. 5). The assignments to biological processes made up the majority (24,111; 50.1%) followed by cellular components (16,903; 35.2%) and molecular functions (7064; 14.7%). Under the category of biological processes, cellular processes (3807; 15.8% of the biological process total) and metabolic processes (3083; 13% of the biological process total) were prominently represented, indicating that some important metabolic activities and cell processes occurred in *A. filiformis* larvae. These observations are similar to previous studies (Fuenzalida et al., 2014; Zhang et al., 2014) including echinoderm studies (Du et al., 2012). Under the classification of molecular functions, binding and catalytic activity were respectively the first and second

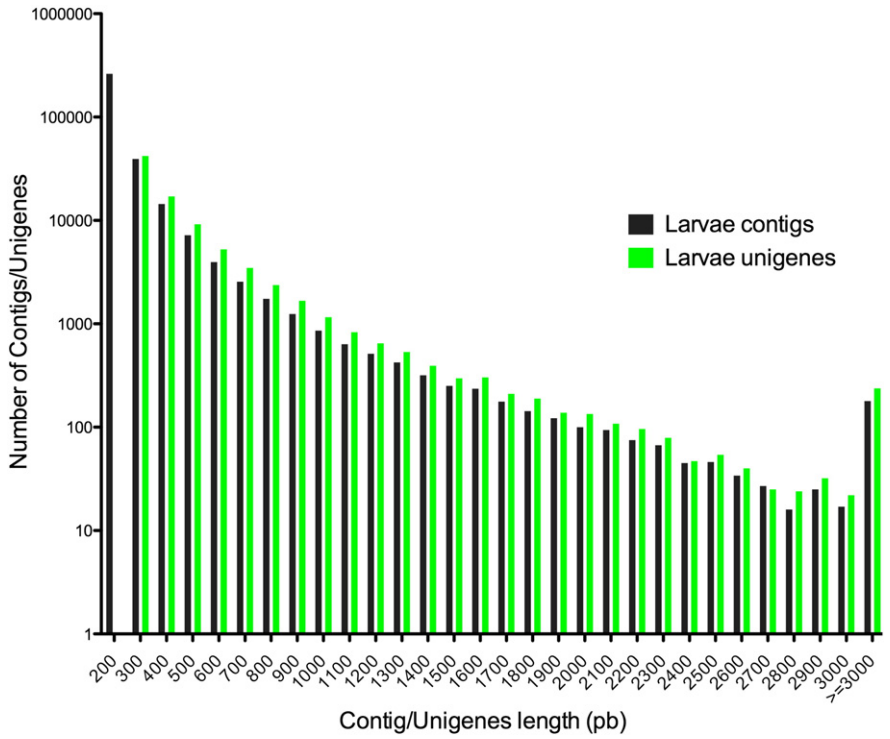

Fig. 3. Distribution of contigs and unigenes in *Amphiura filiformis* pluteus transcriptome. The length of contigs and unigenes ranged from 200 bp to more than 3000 bp. Each range is defined as follows: sequences within the range of X are longer than X bp but shorter than Y bp.

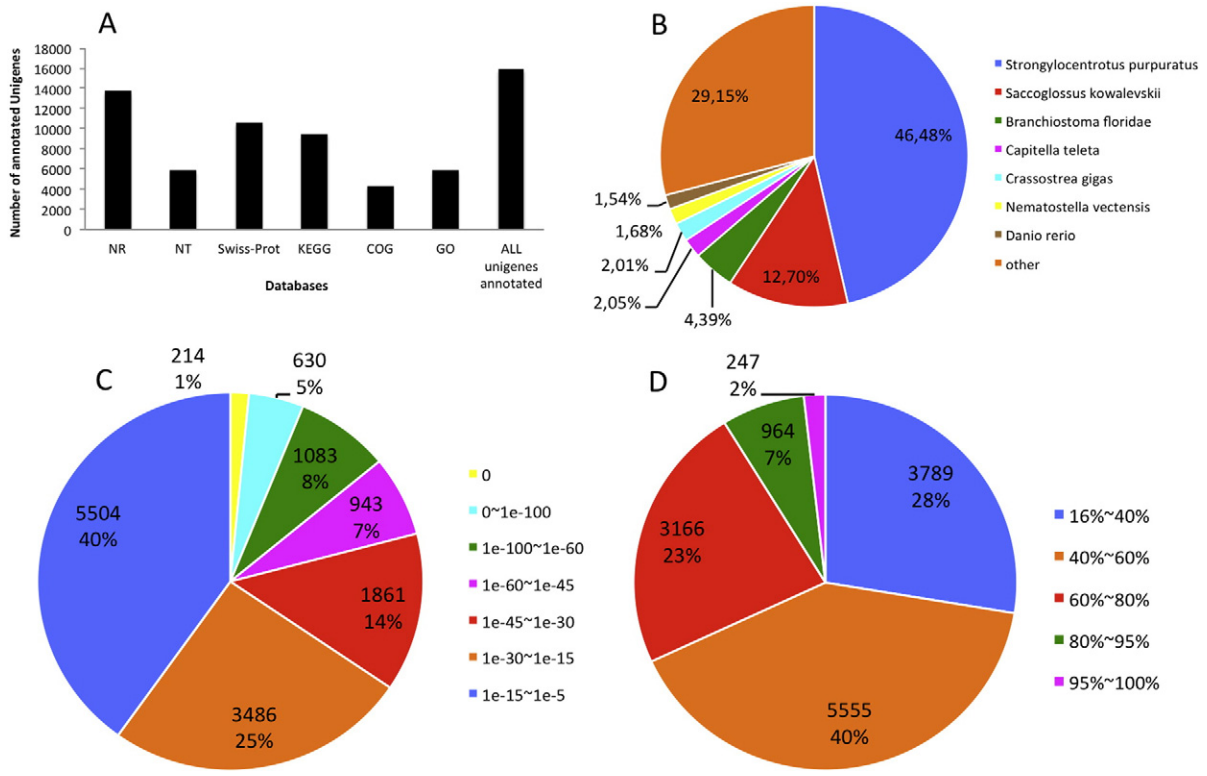

Fig. 4. Annotation statistics of the pluteus transcriptome from *Amphiura filiformis*. (A) Summary of functional annotation of assembled unigenes for classical databases: NR, NT, Swiss-Prot, KEGG, COG, GO. (B) Species distribution of the top BLAST hits for all homologous sequences. (C) E-value distribution of the top BLAST hits for unigenes with cut-off E-values lower than e^{-5}. (D) Similarity distribution of BLAST hits of each unigene in the NR database.

largest categories in the pluteus transcriptome. Other categories such as transporter activity, receptor activity, and enzyme regulator activity for example are less represented. For the cellular components, four categories, cell, cell part, organelle, and organelle part, represented approximately 70% of cellular components whereas few unigenes were assigned to extracellular region, membrane or synapse, for example.

3.4. Clusters of orthologous groups annotation

On 13,722 unigenes with significant similarity to NR proteins, 4357 sequences were assigned to COG annotations (Fig. 6). Among the 25 COG categories, the cluster "General prediction only" was the largest group (2404 unigenes). This first cluster is followed by the "Translation,

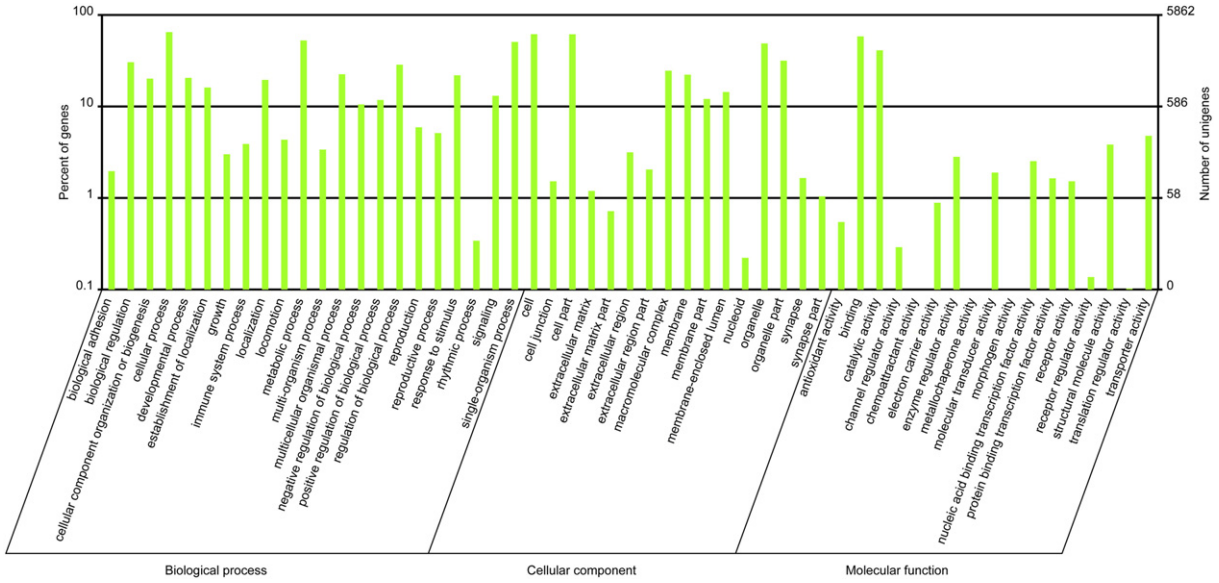

Fig. 5. Gene Ontology classification for the pluteus transcriptome from *Amphiura filiformis*. GO categories are shown on the x-axis. GO categories are grouped into three main categories: biological processes, cellular components and molecular functions. The y-axis indicates the percentage of total genes in each category.

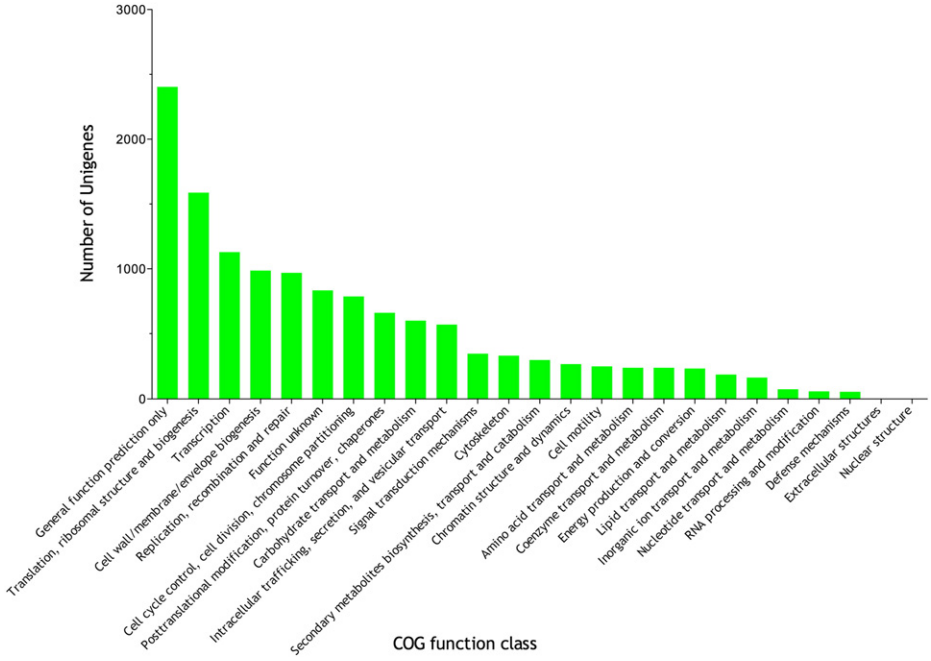

Fig. 6. Clusters of Orthologous Groups (COG) classification of unigene sequences. 4357 unigenes were assigned to 25 categories in the COG classification. The x-axis indicates the number of unigenes in a specific function cluster. The y-axis represents the 25 function categories.

ribosomal structure and biogenesis" (1588 unigenes), the "Transcription" (1129 unigenes) and the "Cell wall/membrane/envelope biogenesis" (987 unigenes) classes.

3.5. *Kyoto* Encyclopedia of Genes and Genomes annotation and pathway mapping

Based on a comparison against the KEGG database using BLASTx with an E-value threshold of $1e^{-5}$, 9529 unigenes have significant matches and were assigned to 5 main categories including 251 KEGG pathways. Among the 5 main categories, metabolism is the largest represented class and the top hit pathway in this category is the "metabolic pathways" with 1281 unigenes (13.44% of KEGG annotated unigenes). The 15 KEGG pathway top hits found in larval transcriptome are listed in Table 2. The functional classification of KEGG provided a valuable resource for investigating specific processes, functions and pathways taking place in the larvae of *A. filiformis*. Several "anthropocentric" pathways (bile secretion, vascular smooth muscle contraction, cardiac muscle contraction) are present but are linked to classical general functions such as digestion, absorption or muscle contraction.

Table 2
Top hit KEGG pathways in the pluteus transcriptome from *Amphiura filiformis*.

	15 top hit pathways	Larvae unigenes	KEGG classification
1	Metabolic pathways	1281 (13.4%)	Metabolism
2	Spliceosome	440 (4.6%)	Genetic information processing
3	Regulation of actin cytoskeleton	431 (4.5%)	Cellular processes
4	RNA transport	387 (4.0%)	Genetic information processing
5	Vascular smooth muscle contraction	350 (3.7%)	Organismal systems
6	Tight junction	301 (3.2%)	Cellular processes
7	Focal adhesion	293 (3.1%)	
8	Purine metabolism	287 (3.0%)	Metabolism
9	Bile secretion	240 (2.5%)	Organismal systems
10	mRNA surveillance pathway	232 (2.4%)	Genetic information processing
11	Pyrimidine metabolism	229 (2.4%)	Metabolism
12	Endocytosis	216 (2.3%)	Cellular processes
13	Cardiac muscle contraction	206 (2.1%)	Organismal systems
14	Protein processing in endoplasmic reticulum	193 (2.0%)	Genetic information processing
15	Adherens junction	189 (2.0%)	Cellular processes

3.6. Transcripts related to opsin-based photoreception in the pluteus transcriptome of A. filiformis

Phototransduction is a biochemical process by which the photoreceptor cells generate electrical signals in response to captured photons. Two main phototransduction cascades characterize classical rhabdomeric and ciliary photoreceptors of metazoans (Yau and Hardie, 2009). Phototransduction-associated genes were used as queries in BLASTx searches against the set of *Amphiura* predicted gene models. Targeted BLAST searches using a dataset of protein sequences from phototransduction pathways revealed transcripts encoding proteins with high similarities to the key components of both the Gq (rhabdomeric phototransduction, classically associated with protostome eyes) and the Gt (Gi, Go) (ciliary phototransduction, classically associated with deuterostomes eyes) phototransduction pathways (Table 3). For both pathways, we selected – in total – 17 classical phototransduction actors (including the genes encoding the visual opsins themselves) (Table 3). On the 6 genes selected for the "rhabdomeric phototransduction" cascade, we found 3 putative orthologous genes in *A. filiformis*. Considering the "ciliary phototransduction" cascade, on 7 selected actors, 6 putative orthologues were found in the pluteus transcriptome. For both phototransduction pathways, specific r-opsins (Gq-coupled opsins) and c-opsins/Go-opsins, defined as primary phototransduction actors, were not found in the larval transcriptome. Exceptions also included the drosophila-specific InaD scaffold proteins which gave a hit to "Lin-7 protein homolog". Several transcripts similar to classical phototransduction actors are expressed in the brittle star pluteus but neither the ciliary nor the rhabdomeric phototransduction pathway is fully covered. Indeed, several important phototransduction actors were not identified in *A. filiformis* plutei such as opsins (Table 3). Considering the common

Table 3
Search for key phototransduction genes in the pluteus transcriptome of *Amphiura filiformis*. Homologs to Gq and Gi/o/t phototransduction cascade components and their reciprocal best BLAST hit. For each query protein, the corresponding *A. filiformis* unigene or contig is listed with the E-value of the top blast result. The *A. filiformis* proteins were then used as query in a reciprocal best BLAST search of the non-redundant protein database (NCBI) and the top result is listed along with the E-value of the blast result. Complete table with protein query accession number and complete names and description is presented in supplementary file 2. Reciprocal best BLAST hits were to proteins with annotations that closely correspond to the query proteins are in bold and indicate good candidates of phototransduction actors.

Protein name	Accession number of top *Amphiura pluteus* result	E-value	Reciprocal best BLASTx query result [species]	E-value
Gq components				
Gq-opsin	CL12642.Cont2	2.00E−06	Probable G-protein coupled receptor No9-like [*S. purpuratus*]	8.00E−17
Gα-q SU	Uni39183	0.0	**Guanine nucleotide-binding protein G(q) α SU [*Lytechinus variegatus*]**	0.0
InaD	Uni10173	1.00E-06	Protein lin-7 homolog B-like isoform 2 [*S. purpuratus*]	5.00E−52
PLC β	CL2351.Cont1	6.00E-68	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase β-4 [*S. purpuratus*]	3.00E−130
TRP	Uni21737	9.00E-04	**Short transient receptor potential channel 4-like [*S. purpuratus*]**	2.00E−10
TRPL	Uni58537	0.001		2.00E−04
Gi/t/o components				
cGMP gated channel α	CL6685.Cont1	3.00E−33	**Tetrameric potassium-selective cyclic nucleotide gated channel [*S. purpuratus*]**	3.00E−115
cGMP gated channel β		5.00E−26		
GC	CL7992.Cont1	1.00E−28	**Guanylate cyclase 2G-like [*S. kowalevskii*]**	3.00E−34
Gα-i SU	Uni28507	5.00E−176	**Guanine nucleotide-binding protein G(i) subunit α [*Patiria pectinifera*]**	0.0
Gα-o SU		1.00E−148		
Gα-t SU		9.00E−140		
Go-opsin	Uni28504	0.029	Peropsin [*Hasarius adansoni*]	3.00E−05
Gt-opsin	Uni54165	0.031	Histamine H2 receptor-like [*S. kowalevskii*]	6.00E−08
PDE α	CL98.Cont2	7.00E−57	**Probable 3′,5′ - cyclic PDE-5-like [*S. purpuratus*]**	8.00E−149
PDE β		2.00E−57		
RGS9	CL975.Cont2	6.00E−25	**RGS 12-like [*S. purpuratus*]**	1.00E−66
NCKX	Uni3676	3.00E−24	**Na/K/Ca exchanger 2-like [*S. purpuratus*]**	3.00E−175
Both Gq/Gi				
Arrestin 2	/			
β Arrestin	CL15539.Contig1	1.7	Histone acetyltransferase 2B [*S. purpuratus*]	1.00E−19
G-protein β SU	Uni34513	0.0	**Guanine nucleotide-binding protein SU β -like isoform 2 [*S. purpuratus*]**	0.0
G-protein γ SU	Uni36409	7.00E−05	**Guanine nucleotide-binding protein G(I/S/O) SU γ − 10 [Danio rerio]**	2.00E−13
		2.00E−07		
RK	Uni33370	4E−95	**GPCR kinase 5-like [*S. kowalevskii*]**	0.0
		2E−145		

actors of both phototransduction pathways (4 selected actors). Only G protein β and γ subunits and a « GPCR kinase-like sequence » were identified.

Delroisse et al. (2014b) recently identified 13 opsin genes in a draft genome of *A. filiformis*. Three partial opsin mRNAs were also detected in an adult transcriptome, including one neuropsin [Af-opsin 8.2, GenBank: KM276774], one basal branch opsin [Af-opsin 2, GenBank: KM276763] and one rhabdomeric opsin [Af-opsin 4.3, GenBank: KM276767] (Fig. 7). In order to identify the potential expression of opsin genes in the pluteus transcriptome, the Af-opsin gene sequences were searched in the larval transcriptome using a «Blast/Alignment» strategy. Local BLASTn and tBLASTn searches did not permit to retrieve any typical visual opsin (r-opsin, c-opsin, Go-opsins). This was unexpected because, in sea urchins, a ciliary opsin (Sp-opsin 1) is expressed in the 72 h pluteus of *S. purpuratus* (Burke et al., 2006; Raible et al., 2006). Ooka et al. (2010) showed the presence of Hp-encephalopsin (homologous to Sp-opsin 1 and Af-opsin 1) mRNA in the embryos and larvae of *H. pulcherrimus*, starting at the swimming blastula stage

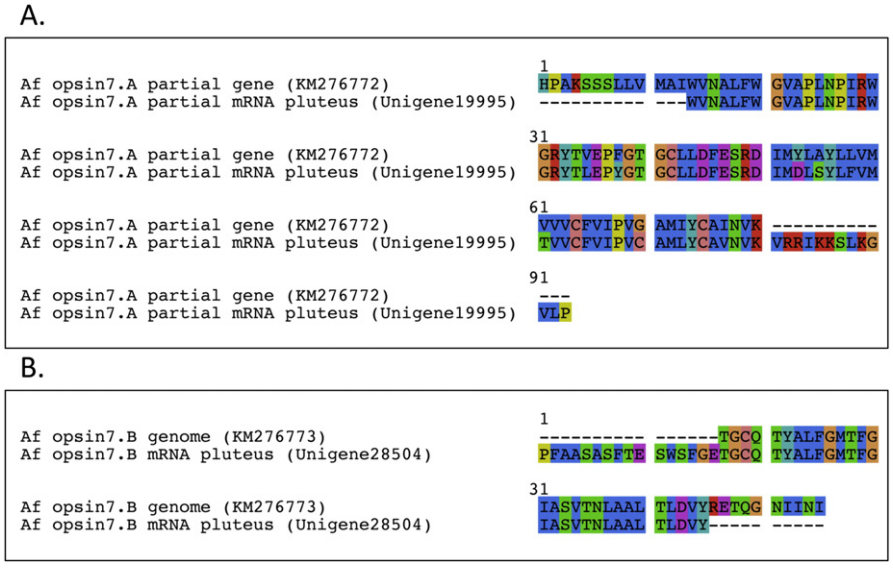

Fig. 7. *Amphiura filiformis* RGR-opsin predicted peptide alignments. (A) Af Opsin 7.A; (B) Af Opsin 7.B.

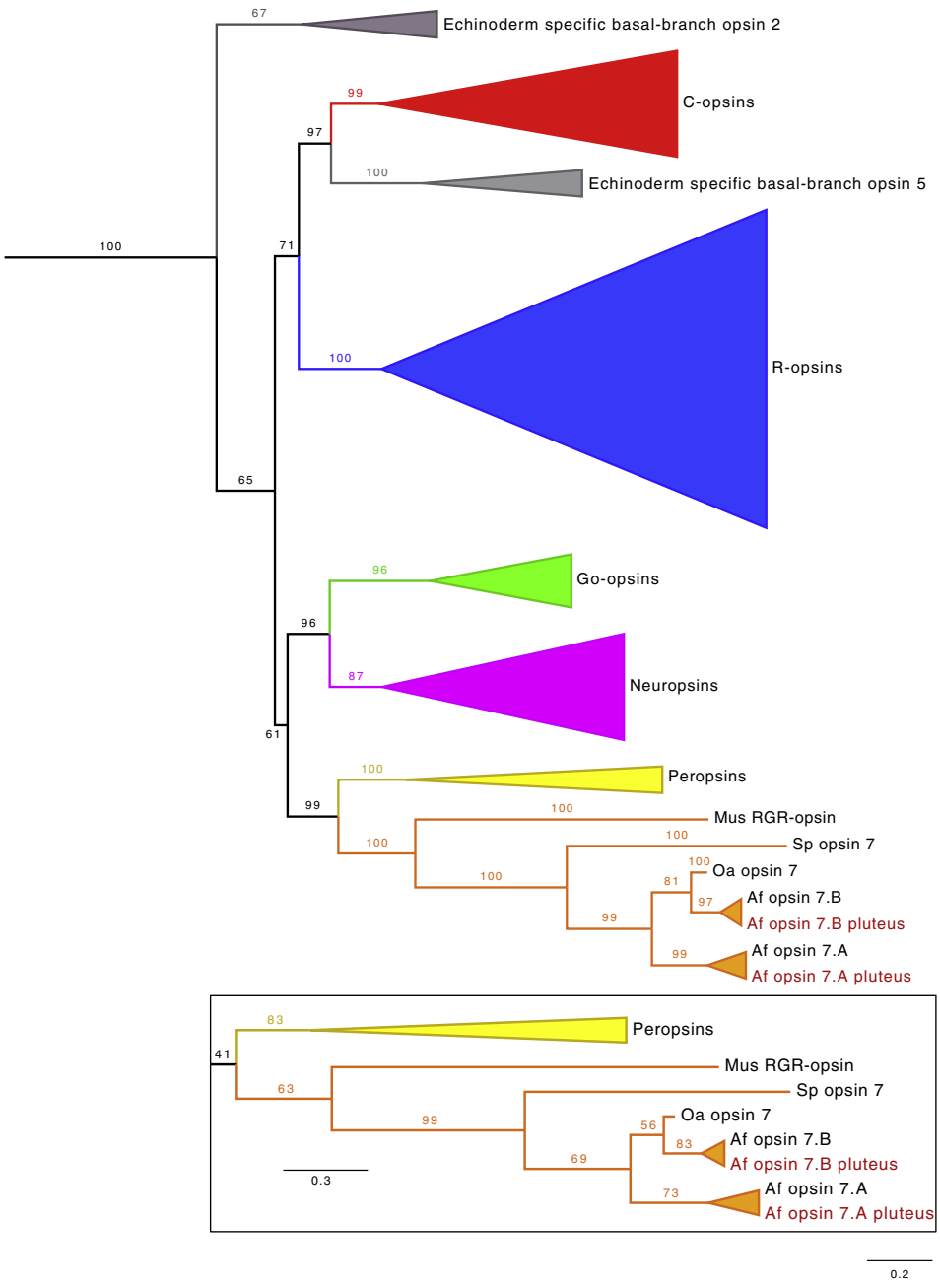

Fig. 8. Phylogenetic tree of opsins based on Bayesian analysis. RGR/Peropsin group for which two sequences were found in the pluteus transcriptome is detailed (posterior probabilities values are indicate on each branch). Maximum likelihood analyses relative to the Peropsin/RGR groups is also presented in the frame (bootstrap values indicated on the branches). See methods for more details. One unpublished echinoderm RGR-Opsin 7 from the brittle star *Ophiopsila aranea* was added to the analysis (SRX791869).

(14 h post-fertilization). Conversely to sea urchin larvae, therefore, early brittle star plutei do not seem to express ciliary opsins at this stage. However, two partial non-visual opsin mRNAs belonging to retinal G protein-coupled receptors (RGR) and corresponding to Af-opsin 7.A (Unigene19995) and Af-opsin 7.B (Unigene28504) (Delroisse et al., 2014b) (see Additional file 3) were detected in the pluteus transcriptome of *A. filiformis*. These two mRNA fragments presumably originate from a single gene, Af-opsin 7 (see Delroisse et al., 2014b). In vertebrates, RGR opsins function as photoisomerases, harvesting light energy to catalyze the recovery of the bleached retinaldehyde released from classical opsins (Hao and Fong, 1996; Hao and Fong, 1999; Chen et al., 2001). RGR opsins are similar in amino acid sequence to retinochromes, photoisomerases that catalyze the conversion of all-*trans*- to 11-*cis*-retinal in squid photoreceptors (Terakita et al., 2000; Tsuda et al., 2003). Due to the short size of the potential Af-opsin 7 fragments, the characteristic opsin "schiff base" residue cannot be identified. However, the phylogenetic position of these partial sequences in the RGR opsin/peropsins cluster is highly supported (Fig. 8). Af-Opsin 7 is therefore supposed to bind to all-*trans*-retinal and should be capable of operating as a stereospecific photoisomerase that would generate 11-*cis*-retinal, as squid retinochrome and mammalian RGR do (Tsuda et al., 2003). Alternatively, Af-opsin 7 may also have a physiological role in visible or UV light perception, as proposed for the urochordate homologous RGR opsin (Ci-opsin 3, (Nakashima et al., 2003)).

3.7. *Pax6 expression in the pluteus transcriptome of A. filiformis*

During the last decade, regulatory genes have been identified encoding transcription factors and signaling molecules that coordinate

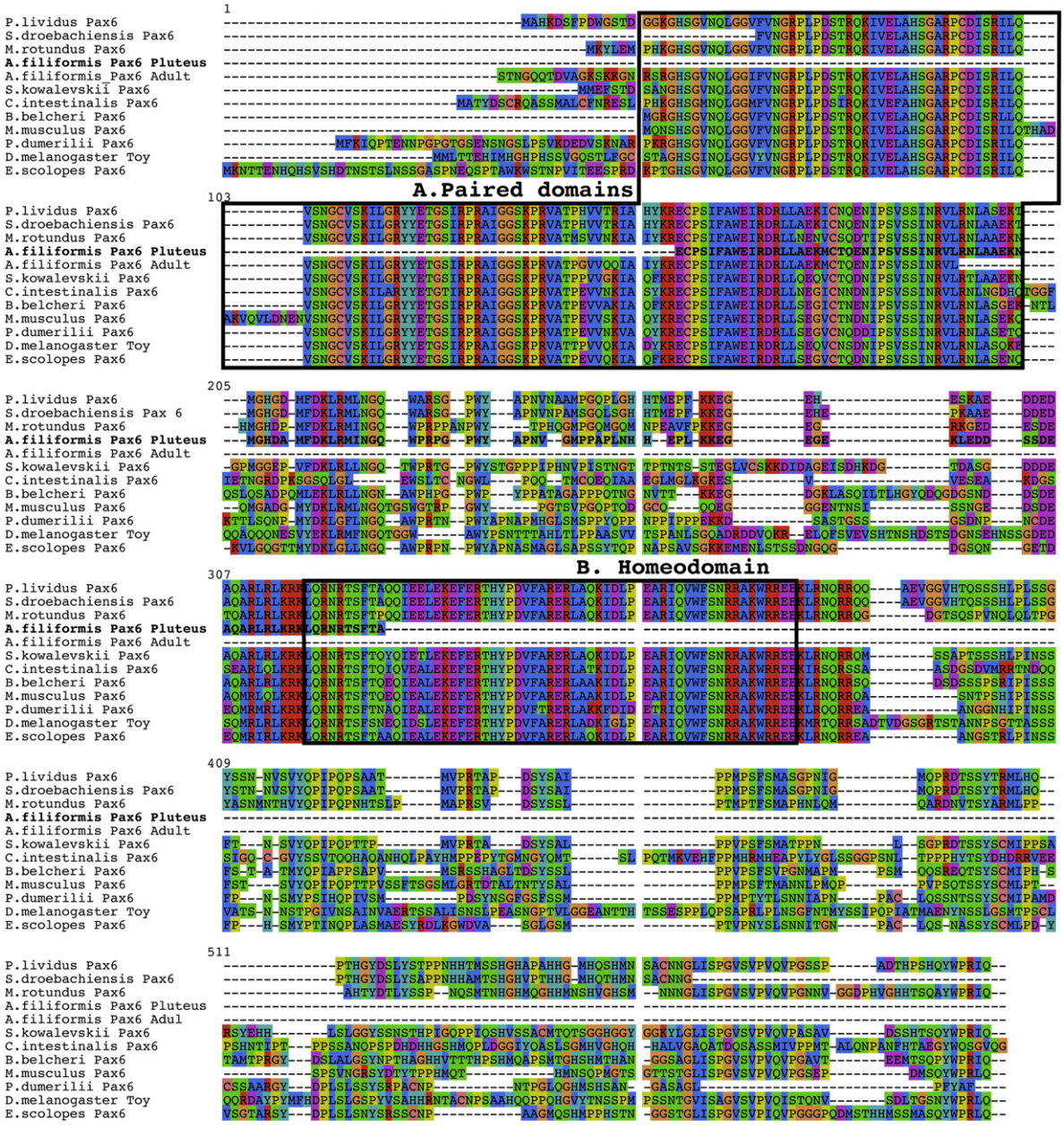

Fig. 9. Multiple sequence alignment of *Amphiura filiformis* Pax6 predicted protein. Pax6 paired-domains and homeodomain are framed. Accession: *Branchiostoma belcheri* Pax6 (ABK54278.1), *Ciona intestinalis* Pax6 (BAB85207.1), *Drosophila melanogaster* Toy (AAD31712.1), *Euprymna scolopes* Pax6 (AAM74161.1), *Mus musculus* Pax6 (NP_038655.1), *Metacrinus rotundus* Pax6 (ADE59459.1), *Paracentrotus lividus* Pax6 (AAA75363.1), *Platynereis dumerilii* Pax6 (CAJ40659.1), *Strongylocentrotus droebachiensis* Pax6 (ABB52751.1), *Saccoglossus kowalevskii* Pax6 (NP_001158383.1).

eye development (Callaerts et al., 1997; Tomarev et al., 1997; Gehring and Ikeo, 1999; Gehring, 2014). These genes are structurally and functionally conserved across wide phylogenetic distances (Gehring, 2014). The genome sequencing of the purple sea urchin *S. purpuratus* led to the discovery of a large number of these typical "eye" genes such as Pax6. In sea urchins, Pax6 is expressed in the tube feet which are defined as the main light sensitive organs (Czerny and Busslinger, 1995; Ullrich-Lüter et al., 2011). In the pluteus transcriptome of *A. filiformis*, two fragments of Pax6 cDNA consisting of a coding region of 142 (Contig 13598, see Additional file 3) and 120 (Unigene 27357, see Additional file 3) amino acids, respectively, were detected. Amino acid sequence alignment of the paired domains and the homeodomain of Af-Pax6 (Unigene27357) with homologs from other species reveal very similar features and, particularly, a 76% identity with the sea-urchin (*S. purpuratus*) sequence (Fig. 9). Positions of the specific amino acids in the conserved domain of Pax6 (Ullrich-Lüter et al., 2011) are completely conserved in Af-Pax6. The phylogenetic analyses using entire amino acid sequences of the predicted Pax proteins derived from different phyla supports the inclusion of Af-Pax6 in the Pax6 subfamily (Fig. 10).

3.8. Data availability

The Illumina derived short read files (*A. filiformis* pluteus transcriptome) are available at the NCBI Sequence Read Archive (SRA) under the study accession number SRR1533125. The unigene dataset will be available online after publication. The unigene dataset is temporarily added as a supplementary file: Additional file 3.

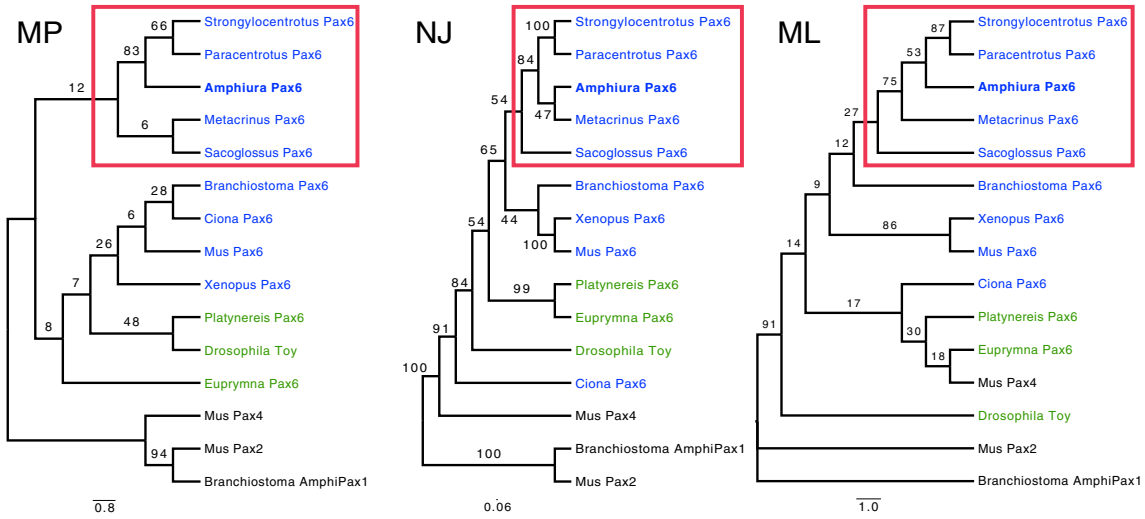

Fig. 10. Cladograms of evolutionary relationships between Pax6 from A. filiformis (bold) and other animals (deuterostomes in blue, protostomes in green). Phylogenetic analyses were performed using trimming alignment of the ORF of each Pax6 predicted proteins. MP. Tree based on maximum parsimony analysis. Bootstrap values are supported by 10,000 tests. NJ. Tree based on Neighbor-joining distance analysis. Bootstrap values are supported by 1,000,000 tests. ML. Consensus tree based on maximum likelihood analyses. Bootstrap values are supported by 100 tests. Based on MEGA4 prediction, LG substitution model and 4 gamma rate categories were used. Ambulacraria sequences are framed in red. Accessions: *Branchiostoma belcheri* Pax6 (ABK54278.1), *Branchiostoma floridae* Amphipax-1 (AAA81364.1), Ciona intestinalis Pax6 (BAB85207.1), *Drosophila melanogaster* Toy (AAD31712.1), Euprymna scolopes Pax6 (AAM74161.1), *Mus musculus* Pax6 (NP_038655.1), *Mus musculus* Pax4 (BAA24517.1), *Mus musculus* Pax2 (CAA39302.1), *Metacrinus rotundus* Pax6 (ADE59459.1), *Paracentrotus lividus* Pax6 (AAA75363.1), *Platynereis dumerilii* Pax6 (CAJ40659.1), *Strongylocentrotus droebachiensis* Pax6 (ABB52751.1), *Saccoglossus kowalevskii* Pax6 (NP_001158383.1), *Xenopus laevis* Pax6 (AAB36683.1).

4. Conclusions

In this work, we produced a developmental transcriptome dataset of the early pluteus larvae (64 h post-fertilization at 14 °C) of the brittle star *A. filiformis* using Illumina HiSeq™ technology. This dataset significantly increases the amount of *A. filiformis* sequence data in the public database, allowing gene discovery for future developmental studies and comparative analyses into a broad range of fundamental questions regarding the evolution of developmental mechanisms in deuterostome animals.

The generated transcriptome was used to complement the photoreception study of the brittle star *A. filiformis* initiated by Delroisse et al.(2014b) by the detection of phototransduction actors in the pluteus larvae. Potential secondary actors of the classical phototransduction pathways, including Gq and Gi proteins, were identified in the transcriptome. However, contrary to reports from sea-urchin larvae at a similar developmental stage, no visual-like opsin was detected in the brittle star early pluteus. Two partial retinal G-protein-coupled receptor (RGR) mRNAs are present in the transcriptome suggesting a "visual cycle mechanism" (Hao and Fong, 1996; Hao and Fong, 1999; Chen et al., 2001; Terakita et al., 2000; Tsuda et al., 2003; Nakashima et al., 2003) despite the apparent absence of classical opsins. Considering their estimated phylogenetic position, these opsins are supposed to be photoisomerases involved in retinoid regeneration (Maeda et al., 2003). Besides the classical phototransduction actors and the opsins, Pax 6 mRNA was also detected in the pluteus transcriptome. The fact that the brittle star pluteus transcriptome contains multiple genes encoding homologs of retinal transcription factor, opsins and phototransduction proteins provides indications for a light receptive system or at least for the development of such a system. Although we have not shown that RGR-opsins, pax6 or phototransduction actor expression is specific to photoreceptor-like cells, or that the genes are expressed together in the same cells, the results support the hypothesis that early pluteus larvae contain cells that express eye-specific transcripts. These results also confirm that Illumina paired-end sequencing is a fast and cost-effective approach for new gene discovery in non-model organism for which the complete genome is not available.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.margen.2015.05.014.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

J.D. performed the experiments and data analyses. O.O.M. and S.D. provide helpful assistance for larval culture. J.D., J.M, and P.F. conceived and designed the experiments. J.D., J.M. and P.F. wrote the manuscript. All authors revised the manuscript.

Acknowledgments

We thank the Beijing Genomic Institute for cDNA library preparation, sequencing and preliminary data analysis. We also thank Nathan Puozzo for his technical support. J.D., J.M. and P.F. are respectively Research Fellow, Research Associate, and Research Director of the Fund for Scientific Research of Belgium (F.R.S.-FNRS). O.O.M. and S.D. were financially supported by the Linnaeus Centre for Marine Evolutionary Biology (CeMEB) at the University of Gothenburg (http://www.cemeb.science.gu.se/) and a Linnaeus grant from the Swedish Research Councils VR and Formas. This work was supported in part by a FRFC Grant no. 2.4590.11. This study is a contribution from the "Centre Interuniversitaire de Biologie Marine" (CIBIM).

References

Aparicio, G., Gotz, S., Conesa, A., Segrelles, D., Blanquer, I., García, J.M., Hernandez, V., Robles, M., Talon, M., 2006. Blast2GO goes grid: developing a grid-enabled prototype for functional genomics analysis. Stud. Health Technol. Inform. 120, 194.

Baden, S.P., Pihl, L., Rosenberg, R., 1990. Effects of oxygen depletion on the ecology, blood physiology and fishery of the Norway lobster *Nephrops norvegicus*. Mar. Ecol. Prog. Ser. 67, 141–155.

Bairoch, A., Boeckmann, B., 1991. The SWISS-PROT protein sequence data bank. Nucleic Acids Res. 19, 2247.

Blackburn, D.C., Conley, K.W., Plachetzki, D.C., et al., 2008. Isolation and expression of Pax6 and atonal homologues in the American horseshoe crab, *Limulus polyphemus*. Dev. Dyn. 237, 2209–2219.

Burke, R.D., Angerer, L.M., Elphick, M.R., et al., 2006. A genomic view of the sea urchin nervous system. Dev. Biol. 300, 434–460. http://dx.doi.org/10.1016/j.ydbio.2006.08.007.

Burns, G., Ortega-Martinez, O., Thorndyke, M.C., et al., 2011. Dynamic gene expression profiles during arm regeneration in the brittle star *Amphiura filiformis*. J. Exp. Mar. Biol. Ecol. 407, 315–322.

Burns, G., Ortega-Martinez, O., Dupont, S., et al., 2012. Intrinsic gene expression during regeneration in arm explants of *Amphiura filiformis*. J. Exp. Mar. Biol. Ecol. 413, 106–112.
Burns, G., Thorndyke, M.C., Peck, L.S., Clark, M.S., 2013. Transcriptome pyrosequencing of the Antarctic brittle star *Ophionotus victoriae*. Mar. Genomics 9, 9–15.
Callaerts, P., Halder, G., Gehring, W.J., 1997. PAX-6 in development and evolution. Annu. Rev. Neurosci. 20, 483–532.
Chen, P., Hao, W., Rife, L., et al., 2001. A photic visual cycle of rhodopsin regeneration is dependent on Rgr. Nat. Genet. 28, 256–260.
Clark, M.S., Thorne, M.A., Toullec, J.-Y., et al., 2011. Antarctic krill 454 pyrosequencing reveals chaperone and stress transcriptome. PLoS One 6, e15919.
Conesa, A., Götz, S., García-Gómez, J.M., et al., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Comput. Appl. Biosci. CABIOS 21, 3674–3676.
Czerny, T., Busslinger, M., 1995. DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). Mol. Cell. Biol. 15, 2858–2871.
Delroisse, J., Flammang, P., Mallefet, J., 2014a. Marine luciferases: are they really taxon-specific? A putative luciferase evolved by co-option in an echinoderm lineage. LUMINESCENCE Vol. 29. Wiley-Blackwell, 111 River St, Hoboken 07030–5774, NJ USA, pp. 15–16.
Delroisse, J., Ullrich-Lüter, E., Ortega-Martinez, O., et al., 2014b. High opsin diversity in a non-visual infaunal brittle star. BMC Genomics 15, 1035.
Dilly, G.F., Gaitán-Espitia, J.D., Hofmann, G.E., 2014. Characterization of the Antarctic sea urchin (*Sterechinus neumayeri*) transcriptome and mitogenome: a molecular resource for phylogenetics, ecophysiology and global change biology. Mol. Ecol. Resour.
Du, H., Bao, Z., Hou, R., et al., 2012. Transcriptome sequencing and characterization for the sea cucumber *Apostichopus japonicus* (Selenka, 1867). PLoS One 7, e33311.
Duineveld, G., Van Noort, G.J., 1986. Observations on the population dynamics of *Amphiura filiformis* (Ophiuroidea: Echinodermata) in the southern North Sea and its exploitation by the dab, *Limanda limanda*. Neth. J. Sea Res. 20, 85–94.
Dupont, S., Thorndyke, W., Thorndyke, M.C., Burke, R.D., 2009. Neural development of the brittlestar *Amphiura filiformis*. Dev. Genes Evol. 219, 159–166.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32 (5), 1792–1797.
Fain, G.L., Hardie, R., Laughlin, S.B., 2010. Phototransduction and the evolution of photoreceptors. Curr. Biol. 20 (3), R114–R124.
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. U. S. A. 109, 18868–18872.
Feuda, R., Rota-Stabelli, O., Oakley, T.H., Pisani, D., 2014. The comb jelly opsins and the origins of animal phototransduction. Genome Biol. Evol. 6, 1964–1971.
Franchini, P., van der Merwe, M., Roodt-Wilding, R., 2011. Transcriptome characterization of the South African abalone *Haliotis midae* using sequencing-by-synthesis. BMC Res. Notes 4, 59.
Fuenzalida, G., Poulin, E., Gonzalez-Wevar, C., et al., 2014. Next-Generation Transcriptome Characterization in Three Nacella Species (Patellogastropoda: Nacellidae) from South America and Antarctica. Mar. Genomics.
Gasteiger, E., Hoogland, C., Gattiker, A., et al., 2005. Protein Identification and Analysis Tools on the ExPASy Server. pp. 571–607.
Gehring, W.J., 2002. The genetic control of eye development and its implications for the evolution of the various eye-types. Int. J. Dev. Biol. 46, 65–74.
Gehring, W.J., 2014. The evolution of vision. WIREs Dev. Biol. 3, 1–40.
Gehring, W.J., Ikeo, K., 1999. Pax 6: mastering eye morphogenesis and eye evolution. Trends Genet. 15 (9), 371–377.
Gillard, G.B., Garama, D.J., Brown, C.M., 2014. The transcriptome of the NZ endemic sea urchin Kina (*Evechinus chloroticus*). BMC Genomics 15 (1), 45.
Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221–224.
Grabherr, M.G., Haas, B.J., Yassour, M., et al., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.
Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.
Guindon, S., Delsuc, F., Dufayard, J.-F., Gascuel, O., 2009. Estimating Maximum Likelihood Phylogenies with PhyML. pp. 113–137.
Hahn, D.A., Ragland, G.J., Shoemaker, D.D., Denlinger, D.L., 2009. Gene discovery using massively parallel pyrosequencing to develop ESTs for the flesh fly *Sarcophaga crassipalpis*. BMC Genomics 10, 234.
Hao, W., Fong, H.K., 1996. Blue and ultraviolet light-absorbing opsin from the retinal pigment epithelium. Biochemistry 35 (20), 6251–6256.
Hao, W., Fong, H.K., 1999. The endogenous chromophore of retinal G protein-coupled receptor opsin from the pigment epithelium. J. Biol. Chem. 274 (10), 6085–6090.
Hennebert, E., Wattiez, R., Demeuldre, M., Ladurner, P., Hwang, D.S., Waite, J.H., Flammang, P., 2014. Sea star tenacity mediated by a protein that fragments, then aggregates. Proc. Natl. Acad. Sci. 111 (17), 6317–6322.
Herring, P.J., 1995. Bioluminescent echinoderms: unity of function in diversity of expression. Echinoderm Res. 1995, 9–17.
Iseli, C., Jongeneel, C.V., Bucher, P., 1999. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. 99, pp. 138–148.
Kanehisa, M., Goto, S., 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27–30.
Kanehisa, M., Araki, M., Goto, S., et al., 2008. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 36, D480–D484.
Kumar, S., Nei, M., Dudley, J., Tamura, K., 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinform. 9 (4), 299–306.
Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25 (7), 1307–1320.
Li, R., Li, Y., Kristiansen, K., Wang, J., 2008. SOAP: short oligonucleotide alignment program. Bioinformatics 24 (5), 713–714.
Ma, D., Yang, H., Sun, L., Chen, M., 2014. Transcription profiling using RNA-Seq demonstrates expression differences in the body walls of juvenile albino and normal sea cucumbers *Apostichopus japonicus*. Chin. J. Oceanol. Limnol. 32, 34–46.
Maeda, T., Van Hooser, J.P., Driessen, C.A., Filipek, S., Janssen, J.J., Palczewski, K., 2003. Evaluation of the role of the retinal G protein-coupled receptor (RGR) in the vertebrate retina in vivo. J. Neurochem. 85 (4), 944–956.
Mashanov, V.S., Zueva, O.R., García-Arrarás, J.E., 2014. Transcriptomic changes during regeneration of the central nervous system in an echinoderm. BMC Genomics 15 (1), 357.
Nakashima, Y., Kusakabe, T., Kusakabe, R., Terakita, A., Shichida, Y., Tsuda, M., 2003. Origin of the vertebrate visual cycle: genes encoding retinal photoisomerase and two putative visual cycle proteins are expressed in whole brain of a primitive chordate. J. Comp. Neurol. 460 (2), 180–190.
Nilsson, H.C., 1999. Effects of hypoxia and organic enrichment on growth of the brittle stars *Amphiura filiformis* (OF Müller) and *Amphiura chiajei* Forbes. J. Exp. Mar. Biol. Ecol. 237, 11–30.
Omori, A., Akasaka, K., Kurokawa, D., Amemiya, S., 2011. Gene expression analysis of Six3, Pax6, and Otx in the early development of the stalked crinoid *Metacrinus rotundus*. Gene Expr. Patterns 11, 48–56.
Ooka, S., Katow, T., Yaguchi, S., et al., 2010. Spatiotemporal expression pattern of an encephalopsin orthologue of the sea urchin *Hemicentrotus pulcherrimus* during early development, and its potential role in larval vertical migration. Develop. Growth Differ. 52, 195–207.
Pertea, G., Huang, X., Liang, F., et al., 2003. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Comput. Appl. Biosci. CABIOS 19, 651–652.
Purushothaman, S., Saxena, S., Meghah, V., et al., 2014. Transcriptomic and proteomic analyses of *Amphiura filiformis* arm tissue-undergoing regeneration. J. Proteome 112, 113–124.
Raible, F., Tessmar-Raible, K., Arboleda, E., et al., 2006. Opsins and clusters of sensory G-protein-coupled receptors in the sea urchin genome. Dev. Biol. 300, 461–475.
Rismani-Yazdi, H., Haznedaroglu, B.Z., Bibby, K., Peccia, J., 2011. Transcriptome sequencing and annotation of the microalgae *Dunaliella tertiolecta*: pathway description and gene discovery for production of next-generation biofuels. BMC Genomics 12, 148.
Ronquist, F., Teslenko, M., van der Mark, P., et al., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Rosenberg, R., 1995. Benthic marine fauna structured by hydrodynamic processes and food availability. Neth. J. Sea Res. 34, 303–317.
Rosenberg, R., Lundberg, L., 2004. Photoperiodic activity pattern in the brittle star *Amphiura filiformis*. Mar. Biol. 145, 651–656.
Schnitzler, C.E., Pang, K., Powers, M.L., et al., 2012. Genomic organization, evolution, and expression of photoprotein and opsin genes in *Mnemiopsis leidyi*: a new view of ctenophore photocytes. BMC Biol. 10, 107.
Sköld, M., Rosenberg, R., 1996. Arm regeneration frequency in eight species of Ophiuroidea (Echinodermata) from European sea areas. J. Sea Res. 35, 353–362.
Smith, M.J., Arndt, A., Gorski, S., Fajber, E., 1993. The phylogeny of echinoderm classes based on mitochondrial gene arrangements. J. Mol. Evol. 36, 545–554.
Sodergren, E., Weinstock, G.M., Davidson, E.H., et al., 2006. The genome of the sea urchin *Strongylocentrotus purpuratus*. Science 314, 941–952.
Sotiriades, E., Dollas, A., 2007. A general reconfigurable architecture for the BLAST algorithm. J. VLSI Signal Process. Syst. Signal Image Video Technol. 48, 189–208.
Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24 (8), 1596–1599.
Tao, X., Gu, Y.-H., Wang, H.-Y., et al., 2012. Digital gene expression analysis based on integrated de novo transcriptome assembly of sweet potato [*Ipomoea batatas* (L.) Lam.]. PLoS One 7, e36234.
Tatusov, R.L., Natale, D.A., Garkavtsev, I.V., et al., 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 29, 22–28.
Tatusov, R.L., Fedorova, N.D., Jackson, J.D., et al., 2003. The COG database: an updated version includes eukaryotes. BMC Bioinforma. 4, 41.
Terakita, A., Yamashita, T., Shichida, Y., 2000. Highly conserved glutamic acid in the extracellular IV–V loop in rhodopsins acts as the counterion in retinochrome, a member of the rhodopsin family. Proc. Natl. Acad. Sci. 97 (26), 14263–14267.
Tomarev, S.I., Callaerts, P., Kos, L., Zinovieva, R., Halder, G., Gehring, W., Piatigorsky, J., 1997. Squid Pax-6 and eye development. Proc. Natl. Acad. Sci. 94 (6), 2421–2426.
Tong, D., Rozas, N.S., Oakley, T.H., Mitchell, J., Colley, N.J., McFall-Ngai, M.J., 2009. Evidence for light perception in a bioluminescent organ. Proc. Natl. Acad. Sci. 106 (24), 9836–9841.
Tsuda, M., Kusakabe, T., Iwamoto, H., Horie, T., Nakashima, Y., Nakagawa, M., Okunou, K., 2003. Origin of the vertebrate visual cycle: II. Visual cycle proteins are localized in whole brain including photoreceptor cells of a primitive chordate. Vis. Res. 43 (28), 3045–3053.
Tu, Q., Cameron, R.A., Davidson, E.H., 2014. Quantitative developmental transcriptomes of the sea urchin *Strongylocentrotus purpuratus*. Dev. Biol. 385, 160–167.

Ullrich-Lüter, E.M., Dupont, S., Arboleda, E., Hausen, H., Arnone, M.I., 2011. Unique system of photoreceptors in sea urchin tube feet. Proc. Natl. Acad. Sci. 108 (20), 8367–8372.
Ullrich-Lüter, E.M., D'Aniello, S., Arnone, M.I., 2013. C-opsin Expressing Photoreceptors in Echinoderms. Am. Zool. 53, 27–38.
Vaughn, R., Garnhart, N., Garey, J.R., Thomas, W.K., Livingston, B.T., 2012. Sequencing and analysis of the gastrula transcriptome of the brittle star *Ophiocoma wendtii*. EvoDevo 3 (19), 1–16.
Vera, J.C., Wheat, C.W., Fescemyer, H.W., et al., 2008. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Mol. Ecol. 17, 1636–1647.
Vistisen, B., Vismann, B., 1997. Tolerance to low oxygen and sulfide in *Amphiura filiformis* and *Ophiura albida* (Echinodermata: Ophiuroidea). Mar. Biol. 128 (2), 241–246.
Vopel, K., Thistle, D., Rosenberg, R., 2003. Effect of the Brittle Star *Amphiura filiformis* (Amphiuridae, Echinodermata) on Oxygen Flux into the Sediment.
Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10, 57–63.
Wang, X.W., Luan, J.B., Li, J.M., Bao, Y.Y., Zhang, C.X., Liu, S.S., 2010. De novo characterization of a whitefly transcriptome and analysis of its gene expression during development. BMC Genomics 11 (1), 400.
Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18, 691–699.
Wygoda, J.A., Yang, Y., Byrne, M., Wray, G.A., 2014. Transcriptomic analysis of the highly derived radial body plan of a sea urchin. Genome Biol. Evol. 6, 964–973.
Yau, K.W., Hardie, R.C., 2009. Phototransduction motifs and variations. Cell 139 (2), 246–264.
Ye, J., Fang, L., Zheng, H., et al., 2006. WEGO: a web tool for plotting GO annotations. Nucleic Acids Res. 34, W293–W297.
Zhang, S., Sui, Z., Chang, L., et al., 2014. Transcriptome de novo assembly sequencing and analysis of the toxic dinoflagellate *Alexandrium catenella* using the Illumina platform. Gene 537, 285–293.
Zhou, Z.C., Dong, Y., Sun, H.J., et al., 2014. Transcriptome sequencing of sea cucumber (*Apostichopus japonicus*) and the identification of gene-associated markers. Mol. Ecol. Resour. 14, 127–138.