RESEARCH ARTICLE

Comparative Transcriptomics of Alternative Developmental Phenotypes in a Marine Gastropod

JOURNAL OF EXPERIMENTAL ZOOLOGY PART B MOLECULAR AND DEVELOPMENTAL EVOLUTION

MARYNA PAULINE LESOWAY[1,2], EHAB ABOUHEIF[1]*, AND RACHEL COLLIN[2]*

[1]Department of Biology, McGill University, Montréal, Quebec, Canada
[2]Smithsonian Tropical Research Institute, Balboa, Ancon, Panama

ABSTRACT

Alternative phenotypes are discrete phenotypic differences that develop in response to both genetic and environmental cues. Nutritive embryos, which arrest their development to serve as nutrition for their viable siblings, are an example of an alternative developmental phenotype found in many animal groups. Females of the marine snail, *Crepidula navicella*, produce broods that consist mainly of nutritive embryos and a small number of viable embryos. In order to better understand the genetic mechanisms that lead to the development of alternative phenotypes in this species, we compared the transcriptomes of viable and nutritive embryos at the earliest stage that we were able to distinguish visually between the two. Using high-throughput Illumina sequencing, we assembled and annotated a de novo transcriptome and compared transcript levels in viable and nutritive embryos. Viable embryos express high levels of transcripts associated with known developmental events, while nutritive embryos express high levels of apoptosis-related transcripts. Gene Ontology term enrichment with GOSeq found that these are associated with the negative regulation of apoptotic processes. This enrichment, combined with morphological evidence, suggests that apoptosis is important in the formation of gastrula-like nutritive embryos. Apoptosis has been implicated in the development of alternative phenotypes in other animal groups, raising the possibility that this mechanism's role in alternative phenotypes is conserved in gastropod development. We suggest possible alternative mechanisms of nutritive embryo development. Most importantly, we contribute further evidence to the hypothesis that nutritive embryos are an alternative developmental phenotype. *J. Exp. Zool. (Mol. Dev. Evol.) 00:1–17, 2016.* © 2016 Wiley Periodicals, Inc.

J. Exp. Zool. (Mol. Dev. Evol.) 00:1–17, 2016

How to cite this article: Lesoway MP, Abouheif E, Collin R. 2016. Comparative transcriptomics of alternative developmental phenotypes in a marine gastropod. J. Exp. Zool. (Mol. Dev. Evol.) 00:1–10.

Grant sponsor: STRI, NSF; Grant number: IOS-1019727; Grant sponsor: Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery; Grant sponsor: FRQNT.

Conflict of interest: None.

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Ehab Abouheif, Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montreal, QC H3A-1B1, Canada.
E-mail: ehab.abouheif@mcgill.ca;
Rachel Collin, Smithsonian Tropical Research Institute, Apartado Postal 0843-03092, Balboa, Ancon, Panama.
E-mail: collinr@si.edu

Received 12 November 2015; Revised 8 April 2016; Accepted 11 April 2016
DOI: 10.1002/jez.b.22674
Published online in Wiley Online Library (wileyonlinelibrary.com).

BACKGROUND

Alternative phenotypes are discrete morphologies produced by individuals within the same species at the same life stage (West- Eberhard, '86, 2003). They are widespread and can be primarily influenced by environmental factors (polyphenisms) or genetic factors (polymorphisms). The developmental causes underlying alternative phenotypes have been traced in several systems including the presence or absence of wings in different castes of ants (Abouheif and Wray, 2002), sexually dimorphic horns in beetles (Moczek and Nijhout, 2002), predator-induced polyphenisms in *Daphnia* (Stabell et al., 2003; Miyakawa et al., 2010), and environmentally induced seasonal polyphenisms in butterflies (Brakefield et al., '96). Although there is a tendency to focus on the impact of alternative phenotypes on adult morphology, they also impact early developmental stages, as in the tadpoles of spadefoot toads that exhibit a resource-based polyphenism (Pfennig, '92). Environmental factors, including the presence of an animal food source, produce either an omnivorous tadpole that feeds on detritus, or a carnivorous tadpole that feeds on small insects and other tadpoles (Ledón-Rettig and Pfennig, 2011). The ubiquity of alternative phenotypes and their complexity has led to the suggestion that they can play an important role in the development of novel characters by allowing the existence of an alternative trait without disrupting the development of an existing one (West-Eberhard, 2003, 2005).

Nutritive embryos, the focus of this study, represent an alternative phenotype in early development. Also known as nurse or trophic embryos, nutritive embryos arrest development and serve as a source of nutrition for their developing siblings. This developmental strategy is found in diverse animal groups, from fish and frogs, to ants and spiders, to polychaete annelids, and marine gastropods (Elgar and Crespi, '92a). In these species, ingestion of either noncleaved embryos (oophagy) or cleaved embryos (adelphophagy), depending on the stage of nutritive embryo developmental arrest, is thought to allow mothers to produce larger offspring without changing initial egg size (Elgar and Crespi, '92b). Nutritive embryos may also increase variation in hatching size (Rivest, '83), reduce intersibling cannibalism (Crespi, '92), or provision offspring in low-food environments (Perry and Roitberg, 2006). Nutritive embryos have evolved numerous times independently in different animal groups, and several times within groups such as the calyptraeid gastropods, buccinid gastropods, and spionid polychaetes (Spight, '76; Gibson, '97; Collin, 2004, 2012; Smith and Thatje, 2013).

Understanding the developmental mechanisms underlying alternative phenotypes can provide insights into how normal developmental pathways can be modified to produce alternate developmental phenotypes. Although nutritive embryos may appear to result from degeneration or simple disruptions of the normal developmental trajectory, the evolution of nutritive embryos as alternative phenotypes implies the gain of novel regulation of developmental control networks. For example, changes in developmental networks produce loss of wings in worker ants. Different wingless castes in the same species show disruptions at different points in the gene network that controls wing development (Abouheif and Wray, 2002). Similarly, the production of nutritive embryos is likely due to a novel developmental pathway.

Despite the repeated evolution of nutritive embryos, the mechanism of development has been investigated in only a few cases. In spionid polychates, the development of nutritive (nurse) embryos occurs via active recruitment of apoptosis to early stages of development (Smith and Gibson, '99; Gibson et al., 2012). DNA degradation is evident in nutritive embryos of *Boccarida proboscoidea* as early as the eight-cell stage (Smith and Gibson, '99). Similar degradation is seen in nutritive embryos of *Polydora cornuta*. Additionally, the apoptotic marker Capsase-3 is active in nutritive embryos of *P. cornuta* at stages equivalent to blastula and trochophore in viable embryos (Gibson et al., 2012). In ants, both queens and workers of several species are able to produce nutritive (trophic) embryos (Crespi, '92). Queens produce either viable embryos that concentrate and localize maternally provided transcripts and proteins, or nutritive embryos that show diffuse maternal determinants (Khila and Abouheif, 2008). Workers also produce nutritive embryos showing diffuse maternal determinants that do not localize (Khila and Abouheif, 2008). These examples suggest at least two mechanisms that may be active during development of nutritive embryos: early mislocalization of maternal determinants in oocytes or early embryos, and the precocious activation of developmental pathways, such as apoptosis, which promote nutritive embryo development.

The current study focuses on calyptraeid gastropods, which are an excellent model for the evolution and development of the nutritive embryo as an alternative developmental phenotype. Approximately 20% of species with documented development produce nutritive embryos (Collin, 2003). Phylogenetic analysis has shown that nutritive embryos have evolved at least eight times and more likely 10 or 11 times independently within the group (Collin, unpublished data), allowing for comparative study (Collin, 2004). The distribution of nutritive embryos across the phylogeny is not different from random, and there is little evidence of phylogenetic constraint on the evolution of nutritive embryo production or other transitions in modes of development (Collin, 2004). Further, in cases where a feeding, swimming (planktotrophic) larva is potentially gained following a loss, it is from clades with nutritive embryos (Collin, 2004; Collin et al., 2007), suggesting that development with nutritive embryos may be unique in fostering evolutionary reversals.

To date, nutritive embryo development has been described in detail for only a few species of calyptraeids. In all cases where nutritive embryos are produced, only a small subset of embryos become viable embryos (Fig. 1D), while the majority of embryos arrest development to produce nutritive embryos (Fig. 1E). Females of *Crepipatella dilatata* produce eggs that do

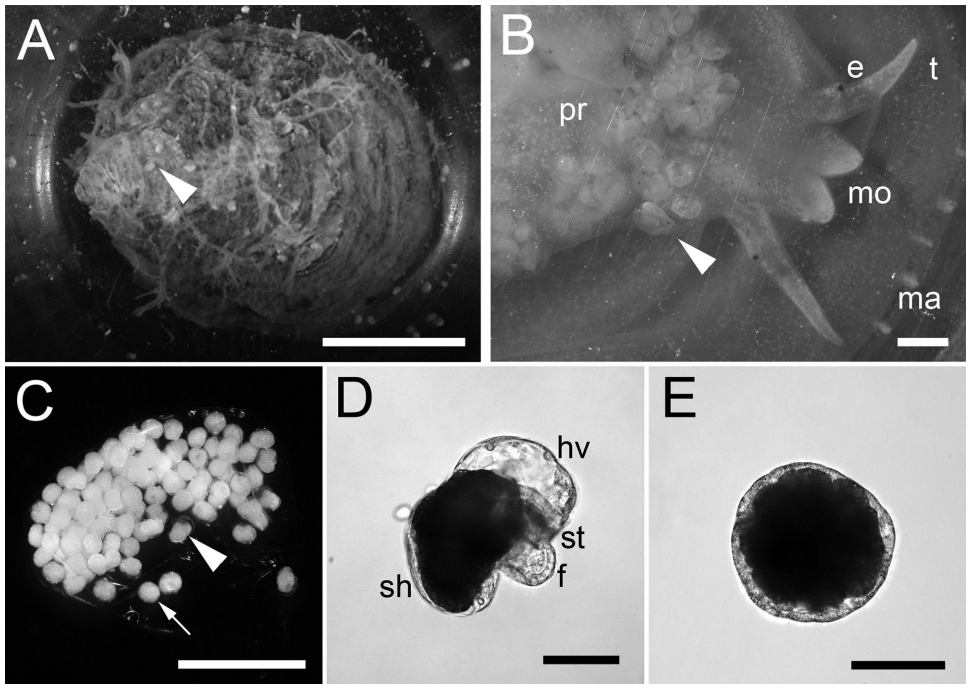

Figure 1. Direct development with nutritive embryos in the calyptraeid gastropod, *Crepidula navicella*. (A) Adult females (dorsal view) brood embryos that hatch as crawl-away juveniles (arrowhead). Scale = 1 cm. (B) Females (ventral view) hold multiple capsules in the space between the neck and the substrate using the propodium. Scale = 1 mm. (C) Individual capsules contain approximately 150 embryos, which develop into either viable (7%, arrowhead) or nutritive (93%, arrow) embryos. Scale = 1 mm. (D) Viable embryo at earliest stage morphologically distinguishable from nutritive embryos. Scale = 100 μm. (E) Gastrula-like nutritive embryo from same capsule as previous. Scale = 100 μm. e, eyespot; f, foot; hv, head vesicle; ma, mantle; mo, mouth; pr, propodium; sh, shell; st, stomodeum; t, tentacle.

not develop prior to being ingested; these are likely unfertilized (Gallardo, '77; Gallardo and Garrido, '87). In contrast, *Crepipatella occulta* and *Crepidula coquimbensis* produce nutritive embryos that cleave and gastrulate (Véliz et al., 2003, 2012). Similarly, all embryos produced by *Crepidula navicella* (Fig. 1) begin development normally, with no variation in initial egg size (Collin and Spangler, 2012). All embryos cleave and develop in the same way until gastrulation, although timing of early cleavage is variable (Lesoway et al., 2014).

We asked what developmental pathways were involved in the elaboration of viable and nutritive embryo phenotypes in the calyptraeid gastropod, *C. navicella*. Our current knowledge of nutritive embryo development suggests that apoptosis and early mislocalization of developmental patterning genes play an important role in producing nutritive embryos (Smith and Gibson, '99; Khila and Abouheif, 2008; Gibson et al., 2012). As we are unable to visually determine the identity of nutritive embryos until late in development, we are limited in our ability to identify maternal determinants in early development. We focus here instead on identifying genes that are active during nutritive and viable embryo elaboration. Using RNASeq, we make a first approach to understanding the developmental mechanisms that produce the nutritive embryo phenotype. Here, we identify and compare developmental transcripts associated with differentiation events in viable and nutritive embryos, investigate the potential role of apoptosis during elaboration of the nutritive embryo phenotype in *C. navicella*, and contribute further evidence to the hypothesis that nutritive embryos are an alternative developmental phenotype.

RESULTS

Sequencing and Assembly

Three pooled samples each of viable and gastrula-like nutritive embryos from four different adult females were sequenced across two lanes of Illumina HiSeq 2000, producing a total of 290,359,312 raw reads. Quality control and digital normalization reduced the number of reads assembled to 60,765,730. De novo assembly produced a total of 290,014 "transcripts" (set of overlapping transcripts, contigs) and 168,839 "genes" (cluster of contigs with shared sequence similarity, also known as unigenes) (Table 1). Most contigs were shorter than 1,000 bp (Fig. 2),

Table 1. Sequencing and assembly statistics of *Crepidula navicella* de novo transcriptome

Sequencing	Total reads	290,359,312
	Total nucleotides	58,071,862,400
Assembly	Total trinity "transcripts"	290,014
	Total trinity "genes"	168,839
	Percent GC	41.75
	Percent reads mapped	65.96
All transcript contigs	N50	1,270
	Maximum length	116,840
	Mean length	782.90
	Median length	429
	Assembled bases	227,052,880
Longest isoform per "gene"	N50	968
	Mean length	635.40
	Median length	346
	Assembled bases	107,280,713
Annotation	Transcripts with BLASTx result: Uniprot Swiss-Prot	17,885
	Transcripts with BLASTx result: *Aplysia*	10,570
Decontamination	Contaminated transcripts (DeconSeq)	10,234
	Clean transcripts (DeconSeq)	279,780

Three replicates each of viable and nutritive embryos from four females were sequenced, 100 bp paired end reads using Illumina HiSeq2000. All reads were pooled to produce a de novo transcriptome assembly with the Trinity software package.

and the average contig length was 782 bp (Table 1). Approximately 65% of the reads used to produce the assembly map correctly as proper pairs, while roughly 27% map as left or right reads only. The remaining reads, approximately 5%, map incorrectly (Supplementary File 1, Table S1). The assembly was subsequently assessed for contamination using DeconSeq (Schmieder and Edwards, 2011), which compared assembled transcripts to RefSeq databases of bacteria, protozoa, virus, plastids, and human sequence (grCH38), removing 10,234 of the assembled contigs from the assembly. This decontaminated assembly was used for downstream differential expression analyses.

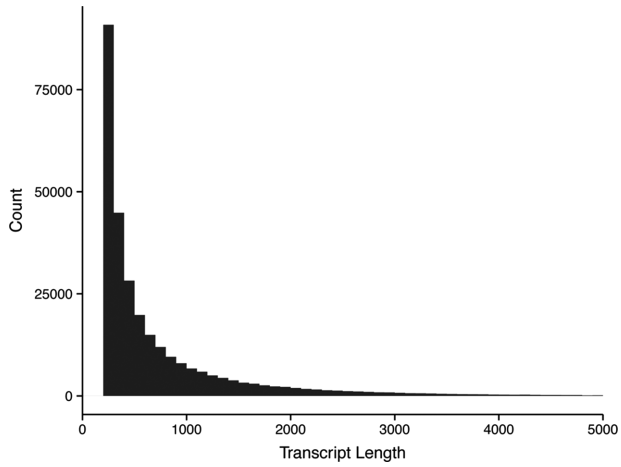

Figure 2. Distribution of Trinity transcript (contig) length for all de novo assembled transcripts. Read lengths over 5,000 bp are pooled.

Annotation

All transcripts were compared to existing databases to suggest transcript identity. A BLASTx search against the Uniprot Swiss-Prot database found hits for 17,885 transcripts with coverage of at least 20% of the transcript length. A similar search against an annotated protein database of *Aplysia californica*, a gastropod with a sequenced genome, found 10,570 hits with coverage of at least 20% transcript length. Candidate coding regions were identified using Transdecoder, which produced 134,692 potential protein translations that were compared to the Uniprot Swiss-Prot database to augment putative transcript identifications. These results were incorporated into an sql template database in Trinotate, which also incorporated protein domain identification results from Pfam and HMMER, and added Gene Ontology (GO) terms to identified transcripts to suggest putative function. Level 2 GO terms found in the complete transcriptome are related to basic cellular processes including biological regulation, cellular process, developmental process, death, and metabolic processes (Fig. 3; Supplementary File 2, Table S2).

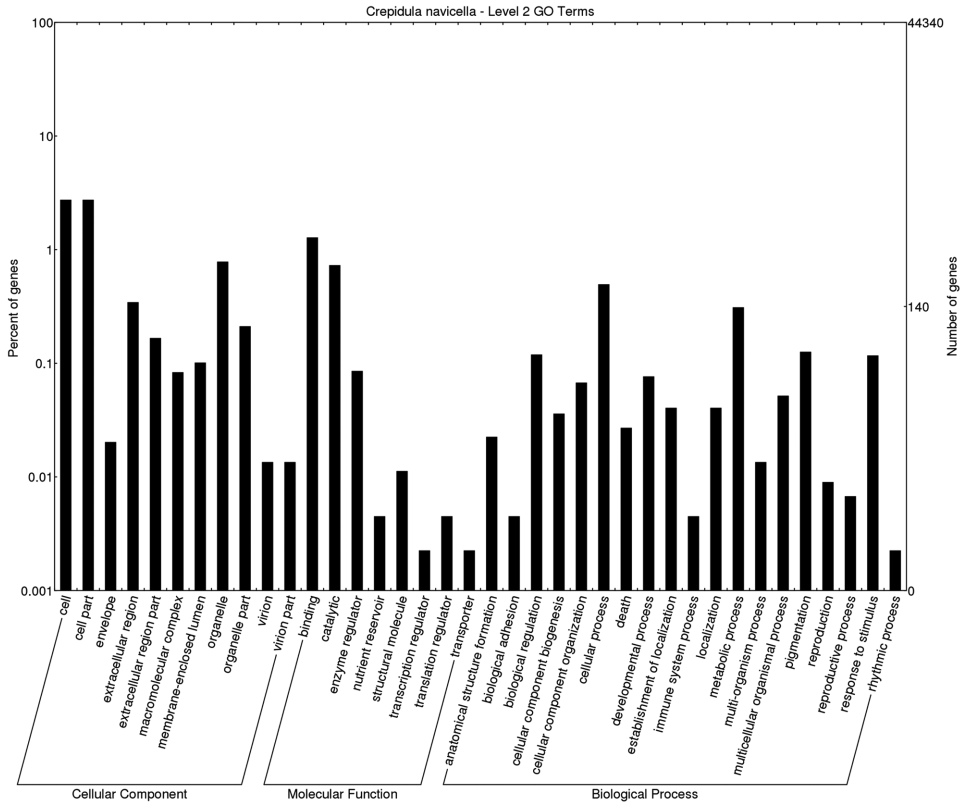

Figure 3. Level 2 GO terms of de novo transcriptome assembly of differentiation stage *Crepidula navicella* embryos.

Differential Expression and GO Enrichment

Developmental differences between viable and nutritive embryos could be due to several factors, including differences in timing and pattern of expression, lack of expression, or changes in interactions within a gene network. Here, we focused on relative differences in the level of expression between the two types of embryos examined. Differential expression analysis found 1,005 differentially expressed transcripts (contigs) between nutritive and viable embryos with at least a fourfold change in expression and a P-value < 0.001 (Fig. 4; Supplementary File 3, Fig. S1 and File 4, Table S3). Of these, 215 transcripts were highly differentially expressed, with at least a 64-fold change in expression and a P-value < 0.00001 (highlighted in Supplementary File 4, Table S3). Principal components analysis (Supplementary File 5, Fig. S2) and a sample correlation matrix (Supplementary File 6, Fig. S3) confirmed that samples of nutritive embryos clustered closely together and were distinct from viable embryos. Of the 1,005 differentially expressed transcripts, 548 transcripts were highly expressed in the viable embryos, and 457 were highly expressed in the nurse embryos (Fig. 4). Differential expression values are relative to overall expression within the pooled samples; increased expression in one type of embryo does not imply zero expression in the other (Fig. 5). Trinotate was able to assign a putative identity based on BLASTx and BLASTp results and assigned GO terms to 27.9% (153/548) of the upregulated transcripts from viable embryos and 59.9% (274/457) of the upregulated transcripts from nutritive embryos.

Highly Expressed Transcripts in Viable Embryos. The majority of differentially expressed transcripts identified in viable embryos can be linked to developmental events occurring in viable embryos at the time of differentiation (Supplementary File 4, Table S3). Numerous transcripts are associated with the cytoskeleton, cell movements, and muscular development including actin, numerous types of myosin (myosin essential light chain, unconventional myosin, and myosin regulatory light chain LC-2: mantle muscle), collagen (alpha-1-II, alpha-1-IV, alpha-1-XV, alpha-2-I), tropomyosin, troponin, and nebulin, a giant muscle associated protein. Transcripts associated with nervous system development were also evident including ELAV-like 1 and CLN8, both previously reported in the nervous system of *A. californica* (Moroz et al., 2006). Other transcripts associated with neural development include neurolignin-4 and repulsive

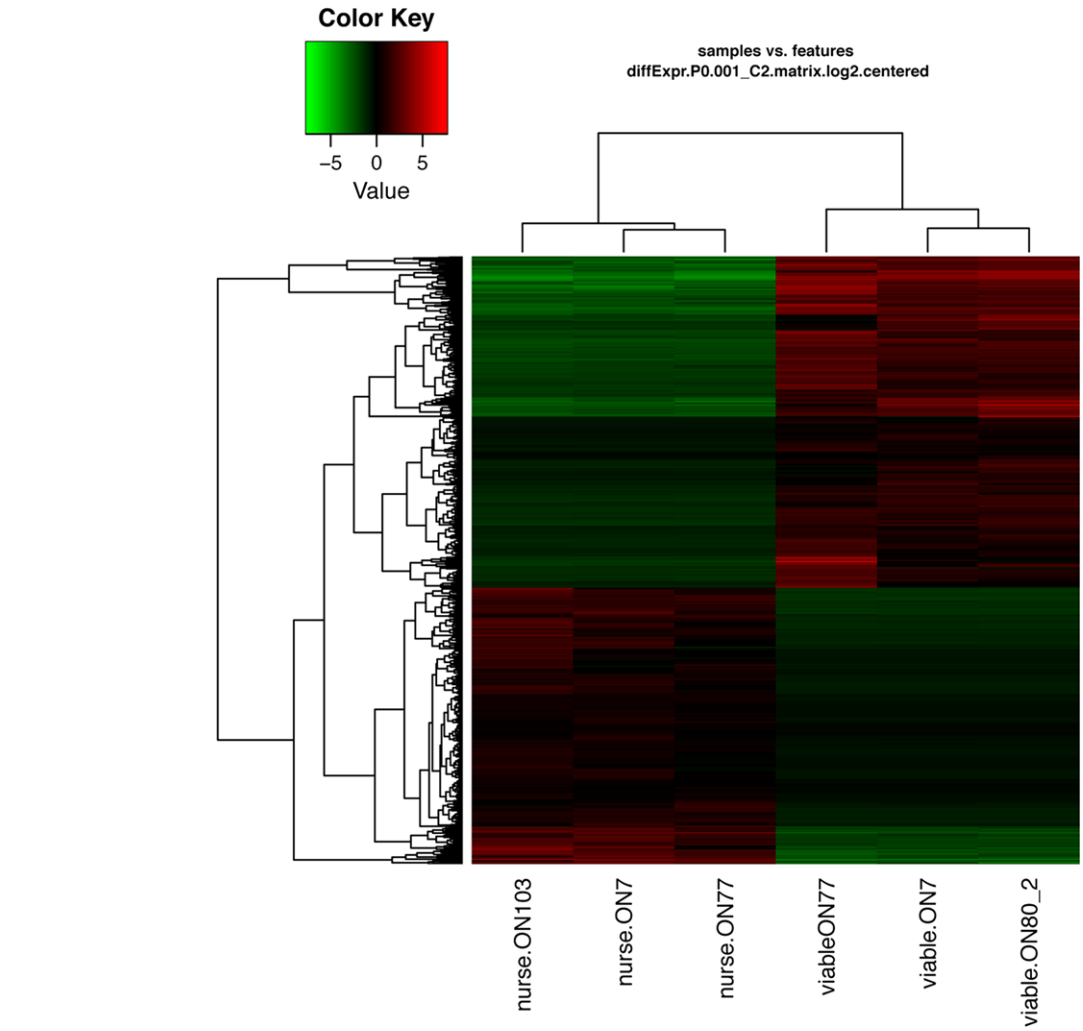

Figure 4. Differentially expressed transcripts of viable and gastrula-like nutritive embryos of *Crepidula navicella*. Log2 centered heat map showing relative expression levels of differentially expressed transcripts ($P < 0.001$, fold change = 4).

guidance molecule A. Repulsive guidance molecule A is associated with axon guidance and formation in vertebrate models (Monnier et al., 2002), and a homologue of neurolignin was found to function in postsynaptic plasticity associated with memory in *A. californica* (Choi et al., 2011). Transcripts implicated in shell secretion are also highly expressed in viable embryos including collagen alpha, peritrophin, and sushi van Willebrand factors (Zhang et al., 2012). Ferretin containing proteins have also been shown to be associated with early shell deposition in gastropod larvae, and are present here (Zhang et al., 2003). Yolk ferretin was also highly expressed in viable embryos, and is likely related to gut development and embryonic nutrition as transcripts have been shown to localize to the developing gut of the freshwater snail *Lymnaea stagnalis* (Bottke et al., '88). Another transcript that is likely involved in gut differentiation is cathepsin B. Although it is generally associated with immune response, Wang et al. (2008, 2011) showed that both RNA and protein localize to the developing gut of the bivalve *Meretrix meretrix*, and that treatment with a cathepsin B inhibitor reduced larval shell growth, potentially by interfering with vitellin degeneration. Known developmental pathways are also represented in viable embryos, notably Frizzled-5, a Wnt protein receptor. Other transcripts of interest include Nicolin-1, which has previously been characterized as a mammalian-only gene the product of which localizes to the nucleus (Backofen et al., 2002). It is present during development and may play a role in transcriptional regulation via Alu-containing elements (Chen et al., 2008).

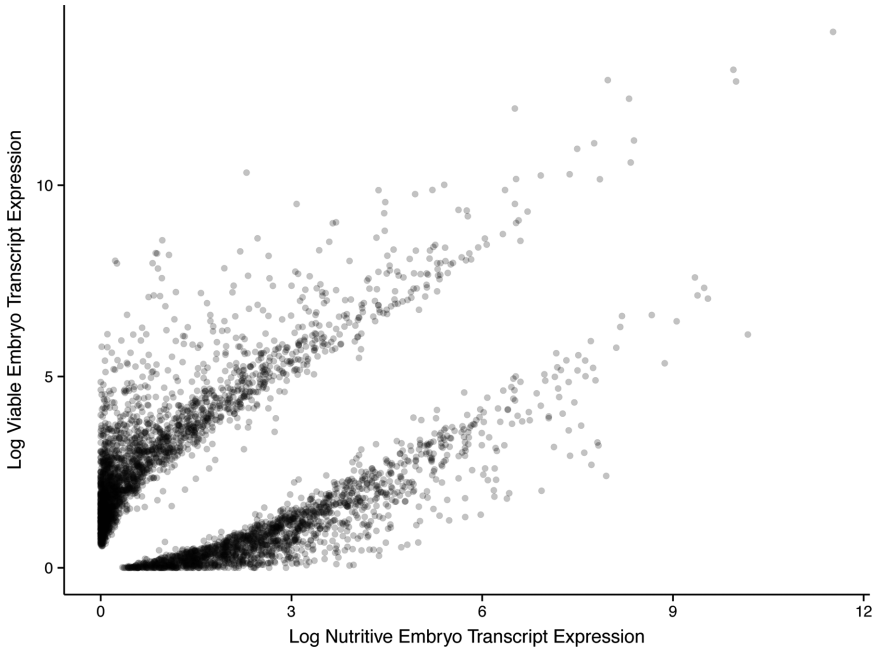

Figure 5. Relative expression of differentially expressed transcripts of viable and nutritive embryos of *Crepidula navicella*. Overlapping points appear darker.

Highly Expressed Transcripts in Nutritive Embryos. Highly expressed transcripts in nutritive embryos had a greater proportion of transcripts with a conserved identity (Supplementary Additional File 4, Table S3). The majority of these were associated with such processes as transcriptional regulation, translational activation, and ubiquitination. Some of the most differentially expressed transcripts (fold-change $> 64\times$, $P < 0.0001$) can be linked to ribosomal processes, such as nucleolin (Tajrishi et al., 2011). A number of transcripts were identified as members of conserved signal transduction pathways, including neurobeachin, a protein kinase A, which in *Drosophila* interacts with both notch and egfr (Shamloula et al., 2002); Wnt signalling associated transcript, Tankyrase-1 (Huang et al., 2009); interleukin-1 receptor-associated kinase 3, which is part of the Toll receptor pathway and is involved in signal transduction (Wesche et al., '99); Rho GTPase-activating protein 29 (Saras et al., '97); and regulatory-associated protein of mTOR, which is involved in controlling cell growth and mediating cellular response to nutritional levels (Kim et al., 2002), and has also been linked to ciliogenesis (Cardenas-Rodriguez et al., 2013). Downstream translational activators were also highly expressed in nutritive embryos, including transcripts of La-related protein, which is also part of the mTOR pathway and has been associated with cell division and apoptosis (Burrows et al., 2010). Other transcriptional regulators include elongation factor 2, and translational activator GCN1, which has been linked to apoptotic activity in *C. elegans* (Hirose and Horvitz, 2014). Apoptosis-related transcripts are well represented, and include NUAK family SNF1-like kinase 1, which has been directly linked to control of proliferation via p53 (Hou et al., 2011). Transcripts associated with ubiquitiniation also occurred frequently in nutritive embryos, with 35 transcripts identified as playing some role in the process, including numerous E3 ubiquitin protein ligases, which are linked to cell cycle control and protein handling in the cell (Teixeira and Reed, 2013). Argounaute-2, a member of the Piwi RNA induced silencing pathway, was also highly expressed in nutritive embryos (Hutvagner and Simard, 2008). A few transcripts have potential functions in tissue differentiation including netrin-1, which has conserved axon guidance and neural migration function in *C. elegans* and vertebrate models (Serafini et al., '94); dyslexia-associated protein KIAA0319-like protein, also thought to function in axon guidance (Poon et al., 2011); disks large associated protein 1, associated with postsynaptic transmission in vertebrates (Welch et al., 2004); and Pax-5, which is expressed in developing statocysts of larvae of *Haliotis asinina* (O'Brien and Degnan, 2003).

GOSeq Functional Enrichment. In order to link differentially expressed transcripts to their putative functions, we performed GO term enrichment analysis using GOSeq, an R Bioconductor package that determines enrichment of functional GO terms in samples of interest (i.e., differentially expressed transcripts)

Table 2. Top five GOSeq-enriched terms from biological process (BP), molecular function (MF), and cellular component (CC) categories highly expressed transcripts of viable embryos of *Crepidula navicella*

Category	*P*-value	Number DE in category	Number in category	GO term
GO:0051823	0.001154759	2	14	BP regulation of synapse structural plasticity
GO:0042832	0.001535492	2	15	BP defense response to protozoan
GO:0097067	0.003187186	2	20	BP cellular response to thyroid hormone stimulus
GO:0006665	0.003324676	3	101	BP sphingolipid metabolic process
GO:0031017	0.003366095	2	23	BP exocrine pancreas development
GO:0003774	3.14×10^{-09}	10	219	MF motor activity
GO:0005201	1.01×10^{-05}	6	128	MF extracellular matrix structural constituent
GO:0005509	2.26×10^{-05}	19	1,923	MF calcium ion binding
GO:0005198	0.000295405	6	247	MF structural molecule activity
GO:0004867	0.001055679	4	123	MF serine-type endopeptidase inhibitor activity
GO:0032982	4.93×10^{-15}	8	23	CC myosin filament
GO:0030016	6.72×10^{-13}	8	38	CC myofibril
GO:0016459	6.51×10^{-08}	8	164	CC myosin complex
GO:0005576	1.14×10^{-06}	20	1,679	CC extracellular region
GO:0005615	1.51×10^{-05}	12	809	CC extracellular space

Complete listings of GOSeq-enriched terms are available in Supplementary File 7, Table S4.

correcting for the increased likelihood of GO term enrichment correlating with longer transcripts. Longer transcripts are more likely to be sequenced due to chance, and therefore to appear to be more highly expressed in read counts (Young et al., 2010). Using the GO terms from the Trinotate annotation, GOSeq analysis found 172 GO terms to be enriched in transcripts that were highly expressed in viable embryos (Table 2; Supplementary File 7, Table S4), and 186 GO terms enriched in transcripts that were highly expressed in nutritive embryos (Table 3; Supplementary File 8, Table S5). Visual comparisons of these terms were made with treemaps produced with the online tool, REVIGO (http://revigo.irb.hr/) (Fig. 6). Treemaps present hierarchical data as nested rectangles, and provide an intuitive visualization of the data set. Here, enriched GO terms are clustered into shared categories, with the size of the rectangle indicating *P*-value (Supek et al., 2011). Highly expressed transcripts in the viable embryos were generally enriched in terms associated with developmental processes, clustering under the headings endocrine pancreas development, regulation of synapse structural plasticity, and sphingolipid metabolism. Several terms also clustered under defence response to protozoan. Highly expressed transcripts in nutritive embryos were enriched in terms associated with the negative regulation of apoptotic processes. Within this cluster (Fig. 6B), GO terms relating to both apoptotic processes and developmental processes are listed. For example, spleen development, spermatid development, thymus development and stem cell differentiation, all appear under the negative regulation of apoptosis supercluster.

Validation of Expression Analysis

Validation of the expression analysis was done with in situ hybridization to confirm the presence and localization of two highly expressed transcripts. As suggested by the in silico analysis, striated muscle myosin heavy chain was expressed only in viable embryos ($n = 10$ viable, 10 nutritive, all embryos from same brood) and is localized to developing retractor muscles (Fig. 7A and B). Expression of argonaute-2 is less distinct, but most apparent in nutritive embryos, appearing as discrete spots throughout the outer epithelium of some nutritive embryos ($n = 5/10$ nutritive embryos). No similar spots were seen in viable embryos ($n = 0/10$ viable embryos) (Fig. 7C and D).

DISCUSSION

Summary of Results

Using high-throughput sequencing, assembly, and analysis of differential expression of transcripts produced in two types of embryos of the direct developing *C. navicella*, gastrula-like nutritive embryos and viable embryos, we found that (1) just over 1,000 transcripts were differentially expressed between nutritive and viable embryos, (2) viable embryo transcripts were enriched in GO terms associated with development of various organs and organ systems, including shell, muscle, nervous, and digestive structures, (3) nutritive embryo transcripts were enriched in GO terms associated with negative regulation of apoptotic processes, as well as ubiquitination, transcriptional,

Table 3. Top five GOSeq-enriched terms from biological process (BP), molecular function (MF), and cellular component (CC) categories highly expressed transcripts of nutritive embryos of *Crepidula navicella*

Category	*P*-value	Number DE in category	Number in category	GO term
GO:0043066	7.12×10^{-05}	14	620	BP negative regulation of apoptotic process
GO:0045088	9.84×10^{-05}	4	37	BP regulation of innate immune response
GO:2001256	0.000927104	2	5	BP regulation of store-operated calcium entry
GO:0046628	0.001078326	2	8	BP positive regulation of insulin receptor signaling pathway
GO:0010609	0.001211662	2	4	BP mRNA localization resulting in posttranscriptional regulation of gene expression
GO:0015078	3.58×10^{-05}	6	88	MF hydrogen ion transmembrane transporter activity
GO:0008270	0.000139141	43	3,544	MF zinc ion binding
GO:0016874	0.000477693	8	268	MF ligase activity
GO:0050431	0.00115898	4	40	MF transforming growth factor beta binding
GO:0005520	0.001268673	2	11	MF insulin-like growth factor binding
GO:0009536	3.84×10^{-05}	5	31	CC plastid
GO:0000220	0.000673654	4	44	CC vacuolar proton-transporting V-type ATPase, V0 domain
GO:0000276	0.004882557	2	16	CC mitochondrial proton-transporting ATP synthase complex, coupling factor F(o)
GO:0031164	0.009943195	1	4	CC contractile vacuolar membrane
GO:0030133	0.012294674	2	44	CC transport vesicle

Complete listings of GOSeq-enriched terms are available in Supplementary File 8, Table S5.

and translational regulation, and (4) differential expression of transcripts of striated muscle-specific myosin heavy chain and argonaute-2 could be visualized in viable embryos and nutritive embryos, with the visualized expression matching the patterns of differential expression in the sequence data. Together with previous evidence, these results are consistent with the hypothesis that viable and nutritive embryos are alternative developmental phenotypes (Lesoway et al., 2014).

Viable Embryos Express Transcripts Associated with Developmental Processes

Highly abundant transcripts in viable embryos can be categorized into a number of general developmental processes. As expected from the phenotype of the viable embryos, transcripts associated with differentiation of muscles, nerves, shell, and the digestive system are highly abundant. For example, striated muscle myosin heavy chain was abundant in viable embryos and localized to developing musculature (Fig. 7A and B). Transcripts in viable embryos had a relatively low rate of identification, suggesting the presence of gastropod-specific transcripts. Although GO terms could be linked with several differentially expressed transcripts, there were challenges with the interpretation of enrichment data. We used the manually curated Swiss-Prot protein database to be more conservative in our functional analyses; however, it includes few well-annotated molluscan species. Several transcripts were associated with GO terms that did not reflect known functions in molluscs despite having BLAST hits to molluscan genes when verified against NCBI's nonredundant nr protein database (data not shown). For example, chitin binding is a functional term that appears frequently in our enrichment analyses (Fig. 6). Peritrophin, identified in our screen, has active sites that bind chitin and is associated with shell secretion in molluscs (Mann et al., 2012). However, in arthropods this motif is associated with food ingestion and possibly immune defence (Shen and Jacobs-Lorena, '98), suggesting that many of the GO terms associated with immune defence have shell deposition functions in molluscs. While this unbiased approach suggests candidates for further study, it demonstrates the challenges remaining for short read analysis in poorly represented systems (Amin et al., 2014), and the need for higher quality annotations of molluscan and lophotrochozoan animals in particular.

Apoptosis Plays a Role in Nutritive Embryo Differentiation

GO term enrichment analysis suggests that apoptosis is important in nutritive embryo differentiation, but contrary to

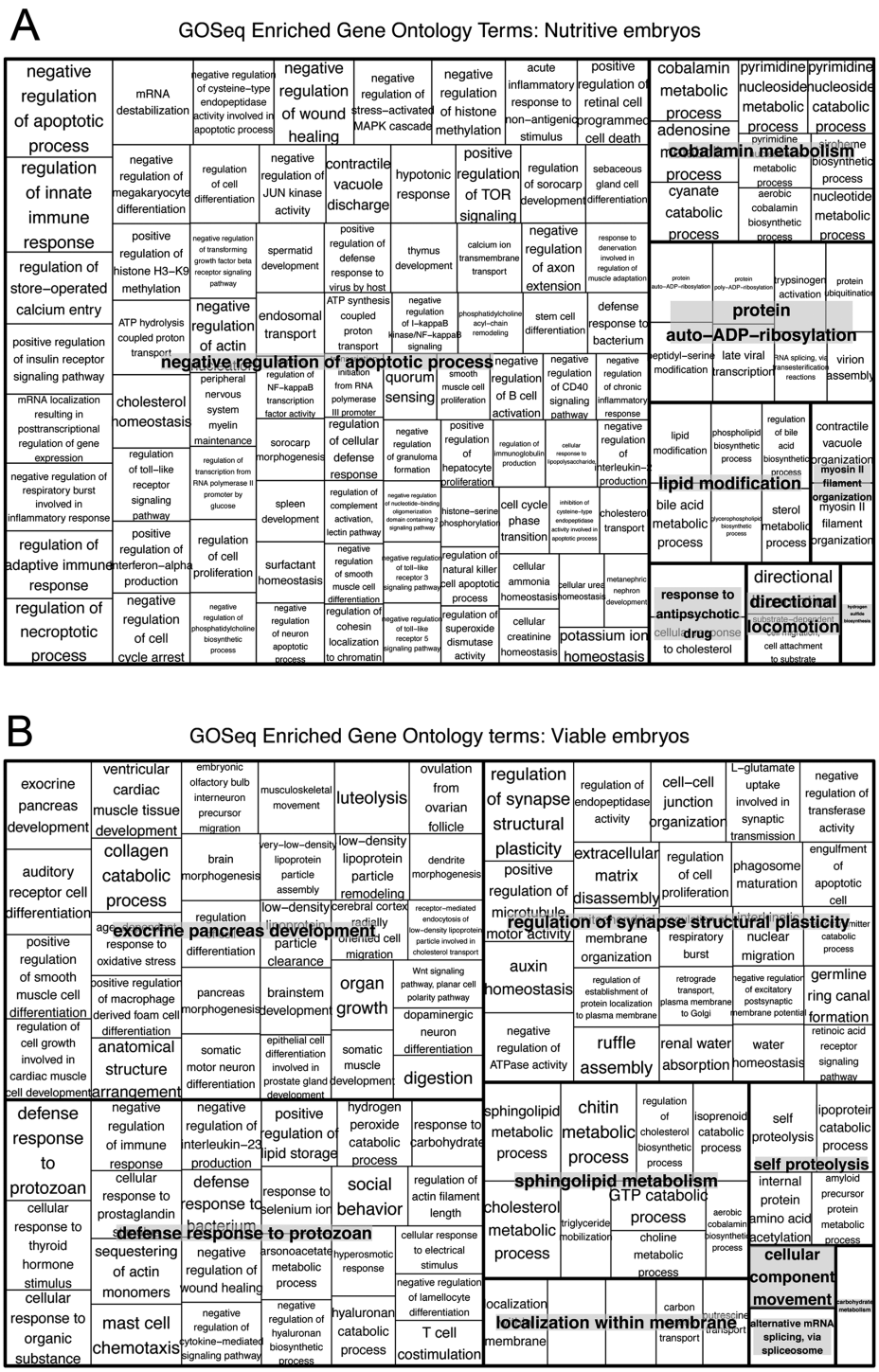

Figure 6. REVIGO treemaps showing enrichment of GO terms in viable and nutritive embryos of *Crepidula navicella*. (A) Biological processes GO terms enriched in viable embryos. (B) Biological processes GO terms enriched in nutritive embryos. Each small rectangle represents a cluster of GO terms, although fewer than 10 GO terms within any list clustered. The size of each rectangle reflects the enrichment *P*-value. Larger rectangles represent related "superclusters" (Supek et al. 2011).

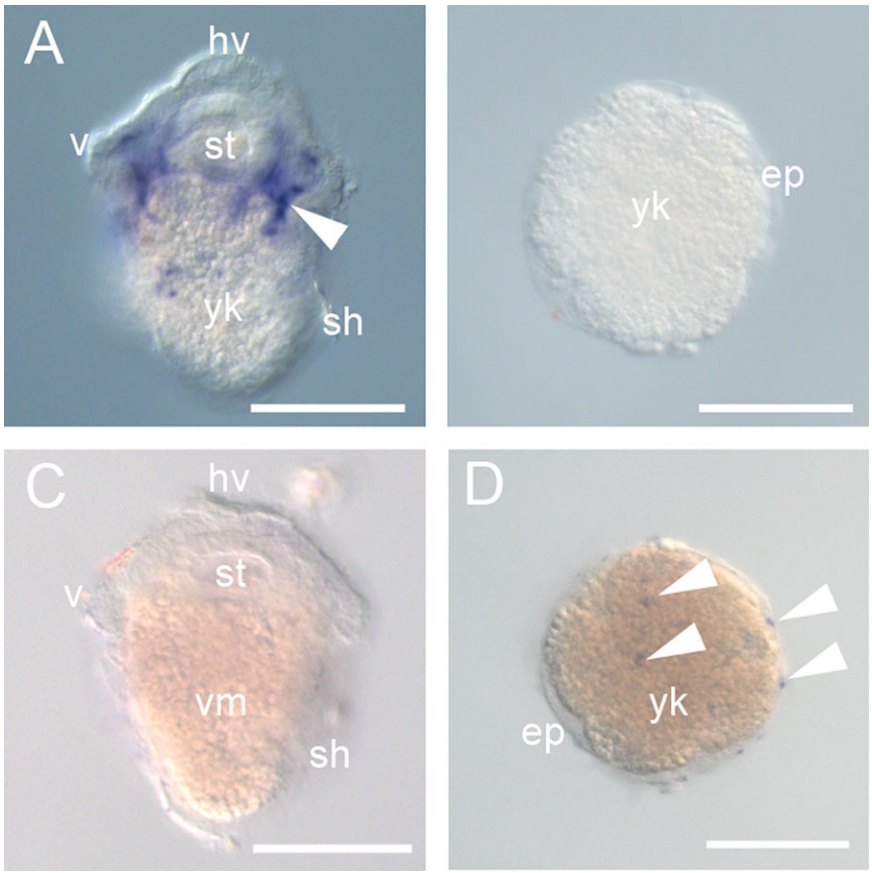

Figure 7. In situ hybridization validates differential expression as predicted by RNASeq analysis. Striated muscle myosin heavy chain mRNA is expressed in developing muscles of viable embryos (A), but not at all in nutritive embryos (B). Conversely, transcripts of Argonaute-2 are not expressed in viable embryos (C), but expressed as punctae in nutritive embryos (D). Arrowheads show probe expression. Scale = 100 μm. Viable embryos shown in ventral view, anterior up; nutritive embryos shown in animal view. ep, epithelium; hv, head vesicle; sh, shell; st, stomodeum; v, velar lobe; vm, visceral mass; yk, yolk.

expectations from nutritive embryo development in spionid polychaetes, apoptosis appears to be negatively regulated. Nutritive embryos of *B. proboscoidea* show DNA degredation and vesiculation (Smith and Gibson, '99). Similarly, nutritive embryos of *P. cornuta* show DNA degredation, changes in membrane characters associated with apoptosis, and nutritive embryo vesiculation as early as the eight-cell stage (Gibson et al., 2012). By contrast, nutritive embryos of *C. navicella* do not show outward signs of apoptosis until after gastrulation, when viable embryos are sufficiently developed to begin feeding, a slightly later stage than the current study (Lesoway et al., 2014). Gibson et al. (2012) suggested that apoptosis has been actively recruited in nutritive embryos of *P. cornuta* to produce degenerating embryos that can be ingested by viable embryos early in development. Similarly, the lack of apoptotic morphological indicators at this stage of development (Lesoway et al., 2014) suggests that negative regulation of apoptosis in nutritive embryos of *C. navicella* may have been actively recruited to maintain nonviable nutritive embryos within the capsule before feeding begins. Alternatively, this could be interpreted as a decline in abundance of the same transcripts and increased apoptotic activity in viable embryos. Cell death plays an important role in morphogenesis, including digit formation in vertebrates, metamorphosis in frogs, and tissue remodeling in *Drosphila* (Suzanne and Steller, 2013), and is known to play a role in degeneration of the larval apical organ in the gastropod *Ilyanassa* (Gifondorwa and Leise, 2006). Further investigation will be required to support either of these possibilities.

In addition to increased abundance of transcripts associated with negative regulation of apoptosis, nutritive embryos have high levels of transcripts associated with fundamental processes of transcription, translation, and protein modification. The high

level of functional conservation of these genes explains the relatively high proportion (nearly 60%) of highly abundant transcripts annotated. Many of these transcripts are associated with threshold processes. For example, ubiquitin-related transcripts are recognized as key players in protein localization and stabilization, and are important to cell cycle regulation; however, our knowledge of how control and downstream roles are regulated remains limited (Teixeira and Reed, 2013). In other words, apoptosis is playing an important role in nutritive embryos, but it is not clear if apoptosis is being negatively regulated as GOSeq analysis suggests. Correct timing and localization of expression of regulatory transcripts plays a key role in determining final developmental outcomes, and better knowledge of upstream events is needed to determine the function of apoptosis-related transcripts highlighted by our study.

Possible Evolutionary and Developmental Scenarios for Nutritive Embryos of *C. navicella*

This study, with previous developmental data (Lesoway et al., 2014), suggests a number of mechanisms that could lead to the nutritive embryo phenotype. The gastrula-like phenotype hints that development of gastrula-like nutritive embryos may be proceeding normally, but slowed relative to viable embryos. Indeed, we have previously shown that timing of early cleavage is highly variable in *C. navicella* and suggested that delayed cleavages could translate into downstream developmental delays (Lesoway et al., 2014). The nutritive embryo phenotype also suggests developmental arrest may occur during the transition from maternal to zygotic control of development. Although the timing of maternal to zygotic transition (MZT) is unknown in *C. navicella*, the onset of MZT is likely to be the stage shortly after gastrulation in embryos of another caenogastropod, *Ilyanassa obsoleta* (Collier, '61, Cooley, 2008). Blocking transcriptional activity during early development of *I. obsoleta* produces a phenotype that look strikingly similar to gastrula-like nutritive embryos of *C. navicella* (Cooley, 2008). Supporting this interpretation in *C. navicella* are high levels of transcriptional regulators identified in nutritive embryos, as well as the presence of transcripts such as Argonaute-2 (Fig. 7C and D), associated with transcript silencing and RNAi (Hutvagner and Simard, 2008). This and other transcriptional regulators could be functioning to silence maternal transcripts and produce zygotic transcripts. A similar hypothesis was put forward by West ('83) who suggested that nutritive embryos were due to "defects" in the oocyte, specifically an inability to transition to zygotically derived cytokinesis.

Upstream factors, including timing and localization of expression, likely play a major role in nutritive embryo determination. For example, mislocalization of maternal transcripts leads to nutritive embryo production in ants. In these social insects, both queens and workers are able to produce different types of nutritive (trophic) embryos; true trophic embryos that show no concentration or localization of the maternal determinant, Vasa, during oogenesis, and "failed" embryos, which show concentration of Vasa during oogenesis, but fail to localize this maternal determinant correctly at the anterior of the embryo (Khila and Abouheif, 2008, 2010). A similar mechanism could be at work in nutritive embryos of *C. navicella*, and detailed examination of localization of maternal transcripts in early embryos will be required to support this hypothesis and to understand the mechanism by which this is disrupted. The current study suggests numerous potential targets for future work.

Nutritive Embryos as an Alternative Phenotype: Limitations and Future Possibilities

Considering the adaptive significance of nutritive embryos raises a key question – are nutritive embryos an adaptive alternative developmental phenotype or a neutral degeneration? In the case of *C. navicella*, the combination of morphological data (Lesoway et al., 2014) and molecular evidence (current study) suggests that nutritive embryos are an adaptive developmental phenotype. However, the current study is limited to a single timepoint. Without knowledge of initial states of transcript abundance or the initial divergence between embryo types, we are unable to confirm the molecular basis for a developmental switch between the viable and nutritive developmental pathways. That being said, if nutritive embryos are the result of neutral degeneration, we would expect that they show no specific phenotype or a great deal of phenotypic variation. We would not expect a distinct molecular signature, and nutritive embryo breakdown would likely begin early. If nutritive embryos are an alternative phenotype, we would expect there to be a distinct and regular phenotype with few anomalies, occurring in large numbers, produced via a specific patterning mechanism, and available as nutrition at an appropriate developmental time.

Our previous work considering the timing of embryonic development in *C. navicella* showed that nutritive embryos form the majority (93% of embryos on average) of reproductive output, that there are two nutritive embryo phenotypes, with the gastrula-like nutritive embryos studied here showing a very regular phenotype, and that nutritive embryos do not begin to disintegrate until viable embryos are capable of ingesting them (Lesoway et al., 2014). Add to this the suggestion of the current study that apoptosis is negatively regulated in gastrula-like nutritive embryos, and the weight of evidence suggests that nutritive embryos are an alternative developmental phenotype.

CONCLUSIONS

Our analyses revealed a large number of differentially expressed transcripts between viable and nutritive embryos. We were able to identify many of the transcripts associated with the differentiation of nutritive and viable embryos. The current study and others like it are providing a wealth of data from which to continue to expand our understanding of the development of nonmodel systems, particularly molluscs and other

Table 4. Sample information for embryos used for RNASeq

Female ID	Embryo type	Number of capsules	Approximate number of embryos per capsule	Approximate number of individual embryos	Fixation date	Embryo age, dpl (days postlaying)
ON7	Viable	12	11	132	Sept. 20, 2011	7
ON7	Nutritive	12	44	532	Sept. 20, 2011	7
ON77	Viable	17	13	227	Nov. 1, 2011	10
ON77	Nutritive	9	99	890	Nov. 1, 2011	10
ON80.2	Viable	12	11	132	Oct. 31, 2011	8
ON103	Nutritive	10	Data not available	Data not available	Sept. 18, 2011	7

Approximate number of embryos is calculated based on an average value from embryo counts of the contents of subset of capsules (where available).

lophotrochozoans (Moroz et al., 2006; Joubert et al., 2010; Bai et al., 2013; Amin et al., 2014; Ho et al., 2014). This snapshot of the transcriptome of *C. navicella* shows that nutritive embryos produce a large number of transcripts associated with the negative regulation of apoptosis. This suggests that apoptosis may be actively delayed in nutritive embryos. The finding of distinct transcriptomic profiles distinguishing viable and nutritive embryos contributes further support to the view that nutritive embryos in *C. navicella* are a developmental polyphenism rather than a neutral degeneration (Lesoway et al., 2014), making this a useful system for further study of the development and evolution of an alternative phenotype.

METHODS

RNA Preparation and Sequencing

Females of *C. navicella* were collected from Playa Venado at Veracruz, Panama (8.886°N, 79.596°W), and maintained in the marine facilities at Naos Island Laboratories at the Smithsonian Tropical Research Institute in Panama, as described previously (Lesoway et al., 2014). Voucher specimens are available at the Redpath Museum, McGill University, Montreal, Canada. Having previously described development of nutritive embryos in *C. navicella,* we focused our efforts on the differentiation stage of development, as there are no obvious markers that distinguish nutritive embryos from viable embryos prior to that point (Lesoway et al., 2014) (Fig. 1). Females laid embryos in lab; capsules were removed from females at the earliest point that viable embryos could reliably be identified, 7–8 days after being laid (Fig. 1C–E). Capsules were rinsed several times in autoclaved, 0.22 μm filtered seawater. All embryos of a brood were removed from their capsules using fine forceps, categorized and separated into three categories: viable embryos, gastrula-like nutritive embryos (Fig. 1D), and postgastrula-like nutritive embryos (sensu Lesoway et al., 2014). Only viable and gastrula-like nutritive embryos were analyzed further. Embryos were transferred to Eppendorff tubes, excess seawater removed and immediately frozen at −80°C. Embryos from four different females were used (Table 4); matched samples of viable and nutritive embryos from two females (four samples), and unmatched samples of viable and nutritive embryos from different females (two samples). All embryos of each type from a brood were included in samples.

Total RNA was extracted from frozen embryos using Trizol (Invitrogen) following the manufacturer's protocol. Initial quantity and quality of total RNA was measured using a Nanodrop spectrophotometer, and three high-quality samples (260/280 ratios greater than 1.8) each of viable and gastrula-like nutritive embryos were sent to the Génome Québec Innovation Centre in Montréal, Canada. There, the quality and integrity of the RNA samples was verified using an Agilent Bioanalyser (Supplementary Additional File 9, Fig. S4). As is typical for many lophotrochozoans, our samples showed only a single RNA peak, likely due to a central hidden break in the 28S rRNA subunit (Ishikawa, '77). We were therefore unable to calculate a RNA Integrity Number (RIN), as reported in other gastropod species (Spade et al., 2010), but used high-quality, nondegraded RNA for downstream isolation of mRNA and library preparation (Illumina TruSeq Library Preparation kit) and sequencing. A total of six samples were used for 2 × 100 bp paired-end sequencing on two lanes of the Illumina HiSeq 2000 (Génome Québec).

Assembly

Demultiplexed, trimmed reads were scanned for overall quality using FASTQC (v0.10.1). Using adapter sequence obtained from Genome Québec, adapter and low-quality bases were removed from the raw reads using Trimmomatic-0.30. In addition to removing known adapter sequence, low-quality reads were removed using a sliding window four bases long, requiring an average Phred quality score of at least 20 within the sliding window. Low-quality reads (Phred score < 3) were removed from the head and tail of all reads, as well as the first 13 bases of each read, which were of lower quality than the rest of the read due

to a known priming bias in Illumina sequencing (Hansen et al., 2010). Reads shorter than 25 bases were also removed. Quality controlled reads from all samples were pooled and assembled using Trinity (Grabherr et al., 2011) (Release 2013-02-25). To reduce assembly time, digital normalization was implemented within Trinity. Normalization and assembly steps were done using the default parameters set by Trinity (maximum read depth 30×). High-performance computations were performed using the Guillimin supercomputer from McGill University managed by Calcul Québec and Compute Canada.

Differential Expression

In order to eliminate possible contamination in the assembled transcripts, we used the standalone version of Deconseq (v0.4.3), and compared assembled transcripts to RefSeq databases of bacteria, archaea, virus, protozoa, plastid, and human reference genome v38, using alignment identity cutoffs of 90% and coverage of 95%. Reads from each sample were then mapped to the cleaned de novo assembly using Bowtie (Langmead et al., 2009) as implemented in Trinity. A table of counts for each mapped read was produced with RSEM (Li and Dewey, 2011). Counts were normalized and compared using the Bioconductor package EdgeR (Robinson et al., 2010). All mapping and differential expression steps were implemented using perl scripts provided with Trinity, using the default parameters set by Trinity (Haas et al., 2013).

Annotation

A local BLAST installation was used to query the de novo assembly against the Uniprot Swiss-Prot manually curated protein database (provided by Trinotate, see details below), and annotated *Aplysia* protein database. Likely protein translations were determined using Transdecoder (release Jan. 16, 2014). An annotated version of the assembly was produced using Trinotate (release Jul. 8, 2014), which incorporated data from the results of nucleotide and protein BLAST searches against Swiss-Prot.

GO Term Enrichment

Differentially expressed genes that were identified via orthology to genes identified in the Uniprot database (Trinotate annotation) were analysed with GOSeq (Young et al., 2010) as implemented in Trinity (rel2014-07-17). The resulting lists of enriched GO terms and associated *P*-values were visualized with RESVIGO (Supek et al., 2011), which removes redundant GO terms and produces intuitive graphical outputs.

Validation: In Situ Hybridization

Two highly differentially expressed genes were selected from the assembly to verify their presence in viable and nutritive embryos of *C. navicella*; striated muscle myosin heavy chain that was highly expressed in viable embryos, and argonaute-2 that was highly expressed in nutritive embryos based on differential expression analysis. Methodological details of embryo fixation, gene cloning, in situ hybridization, and imaging are provided in Supplementary File 10, Additional Methods. Briefly, we isolated fragments of the above genes using primers designed based on our sequencing results (Supplementary File 11, Table S6) from cDNA synthesized from RNA extracted as above. Fragments were cloned and sequenced to confirm their identity. The resulting clones were used as templates to make into DIG-labelled RNA probes for use in in situ hybridization in both viable and nutritive embryos.

Availability of Supporting Data

The raw reads supporting the results of this article are available in the GenBank SRA depository under accession number SRP068776. The assembly has been deposited under the accession GELE00000000. The version described in this paper is the first version, GELE01000000. Data sets supporting the annotation and differential expression results of this article are included within the article and its additional files.

ACKNOWLEDGMENTS

Members of the Collin lab were instrumental in collection and maintenance of animals in Panama. Collection and export permits were kindly granted by the Autoridad de los Recursos Acuáticos de Panamá (ARAP). Computations were made on the supercomputer Guillimin from McGill University, managed by Calcul Québec and Compute Canada. The operation of this supercomputer is funded by the Canada Foundation for Innovation (CFI), NanoQuébec, Réseau de Médicine Génétique Appliquée (RMGA) and the Fonds de recherche du Québec – Nature et technologies (FRQNT). We thank Owen McMillan for support of this project, and Marie-Julie Favé for valuable comments on a previous version of the manuscript. Support for this work was provided by STRI, NSF grant IOS-1019727 to R. C., and a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant to E. A., and an FRQNT doctoral scholarship to M. P. L.

Authors' contributions

M.P.L. carried out the RNA extractions, transcriptome assembly and RNASeq analysis, and wrote the manuscript. E.A. and R.C. contributed to the conception and design of the study and edited the manuscript. All authors read and approved the final manuscript.

LITERATURE CITED

Abouheif E, Wray GA. 2002. Evolution of the gene network underlying wing polyphenism in ants. Science 297:249–252.

Amin S, Prentis PJ, Gilding EK, Pavasovic A. 2014. Assembly and annotation of a non-model gastropod (*Nerita melanotragus*) transcriptome: a comparison of *de novo* assemblers. BMC Res Notes 7:488.

Backofen B, Jacob R, Serth K, Gossler A, Naim HY, Leeb T. 2002. Cloning and characterization of the mammalian-specific nicolin 1 gene (NICN1) encoding a nuclear 24 kDa protein. Eur J Biochem 269:5240–5245.

Bai Z, Zheng H, Lin J, Wang G, Li J. 2013. Comparative analysis of the transcriptome in tissues secreting purple and white nacre in the Pearl Mussel *Hyriopsis cumingii*. PLoS One 8:e53617.

Bottke W, Burschyk M, Volmer J. 1988. On the origin of the yolk protein ferritin in snails. Roux's Arch Dev Biol 197:377–382.

Brakefield PM, Gates J, Keys D, Kesbeke F, Wijngaarden PJ, Montelro A, French V, Carroll SB. 1996. Development, plasticity and evolution of butterfly eyespot patterns. Nature 384:236–242.

Burrows C, Abd Latip N, Lam SJ, Carpenter L, Sawicka K, Tzolovsky G, Gabra H, Bushell M, Glover DM, Willis AE, Blagden SP. 2010. The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. Nucleic Acids Res 38:5542–5553.

Cardenas-Rodriguez M, Irigoin F, Osborn DP, Gascue C, Katsanis N, Beales PL, Badano JL. 2013. The Bardet-Biedl syndrome-related protein CCDC28B modulates mTORC2 function and interacts with SIN1 to control cilia length independently of the mTOR complex. Hum Mol Genet 22:4031–4042.

Chen L-L, DeCerbo JN, Carmichael GG. 2008. Alu element-mediated gene silencing. EMBO J 27:1694–1705.

Choi YB, Li HL, Kassabov SR, Jin I, Puthanveettil SV, Karl KA, Lu Y, Kim JH, Bailey CH, Kandel ER. 2011. Neurexin–neuroligin transsynaptic interaction mediates learning-related synaptic remodeling and long-term facilitation in aplysia. Neuron 70:468–481.

Collier JR. 1961. Nucleic acid and protein metabolism of *Ilyanassa* embryo. Exp Cell Res 24:320–326.

Collin R. 2003. Worldwide patterns in mode of development in calyptraeid gastropods. Mar Ecol Prog Ser 247:103–122.

Collin R. 2004. Phylogenetic effects, the loss of complex characters, and the evolution of development in calyptraeid gastropods. Evolution 58:1488–1502.

Collin R. 2012. Nontraditional life-history choices: what can "intermediates" tell us about evolutionary transitions between modes of invertebrate development? Integr Comp Biol 52:128–137.

Collin R, Spangler A. 2012. Impacts of adelphophagic development on variation in offspring size, duration of development, and temperature-mediated plasticity. Biol Bull 223:268–277.

Collin R, Chaparro OR, Winkler F, Veliz D. 2007. Molecular phylogenetic and embryological evidence that feeding larvae have been reacquired in a marine gastropod. Biol Bull 212:83–92.

Cooley J, 2008. Cell fate determination along the anterior-posterior axis in the Mud Snail *Ilyanassa obsoleta*. University of Arizona. Tucson, AZ. p. 103.

Crespi B. 1992. Cannibalism and trophic eggs in subsocial and eusocial insects. In: Elgar M, Crespi B, editors. Cannibalism: ecology and evolution among diverse taxa. New York: Oxford University Press. p 176–213.

Elgar M, Crespi B. 1992a. Cannibalism: ecology and evolution among diverse taxa. Oxford: Oxford University Press.

Elgar M, Crespi B. 1992b. Ecology and evolution of cannibalism. In: Elgar M, Crespi B, Eds. Cannibalism: Ecology and Evolution among Diverse Taxa. New York: Oxford University Press. 1–12.

Gallardo CS. 1977. Two modes of development in the morphospecies *Crepidula dilatata* (Gastropoda: Calyptraeidae) from Southern Chile. Mar Biol 39:241–251.

Gallardo C, Garrido O. 1987. Nutritive egg formation in the marine snails *Crepidula dilatata* and *Nucella crassilabrum*. Int J Invertebr Reprod 11:239–254.

Gibson GD. 1997. Variable development in the spionid *Boccardia proboscidea* (Polychaeta) is linked to nurse egg production and larval trophic mode. Invertebr Biol 116:213–226.

Gibson G, Hart C, Coulter C, Xu HX. 2012. Nurse eggs form through an active process of apoptosis in the spionid *Polydora cornuta* (Annelida). Integr Comp Biol 52:151–160.

Gifondorwa DJ, Leise EM. 2006. Programmed cell death in the apical ganglion during larval metamorphosis of the marine mollusc *Ilyanassa obsoleta*. Biol Bull 210:109–120.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494–1512.

Hansen KD, Brenner SE, Dudoit S. 2010. Biases in Illumina transcriptome sequencing caused by random hexamer priming. Nucleic Acids Res 38:e131.

Hirose T, Horvitz HR. 2014. The translational regulators GCN-1 and ABCF-3 act together to promote apoptosis in *C. elegans*. PLoS Genet 10:e1004512.

Ho KKY, Leung PTY, Ip JCH, Qiu JW, Leung KMY. 2014. *De novo* transcriptomic profile in the gonadal tissues of the intertidal whelk *Reishia clavigera*. Mar Pollut Bull 85:499–504.

Hou X, Liu JE, Liu W, Liu CY, Liu ZY, Sun ZY. 2011. A new role of NUAK1: directly phosphorylating p53 and regulating cell proliferation. Oncogene 30:2933–2942.

Huang SM, Mishina YM, Liu S, Cheung A, Stegmeier F, Michaud GA, Charlat O, Wiellette E, Zhang Y, Wiessner S, Hild M, Shi X, Wilson CJ, Mickanin C, Myer V, Fazal A, Tomlinson R, Serluca F, Shao W, Cheng H, Shultz M, Rau C, Schirle M, Schlegl J, Ghidelli S, Fawell S, Lu C, Curtis D, Kirschner MW, Lengauer C, Finan PM, Tallarico JA, Bouwmeester T, Porter JA, Bauer A, Cong F. 2009. Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature 461:614–620.

Hutvagner G, Simard MJ. 2008. Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol 9:22–32.

Ishikawa H. 1977. Evolution of ribosomal RNA. Comp Biochem Phys B: Comp Biochem 58:1–7.

Joubert C, Piquemal D, Marie B, Manchon L, Pierrat F, Zanella-Cleon I, Cochennec-Laureau N, Gueguen Y, Montagnani C. 2010. Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and shell: focus on biomineralization. BMC Genomics 11:613.

Khila A, Abouheif E. 2008. Reproductive constraint is a developmental mechanism that maintains social harmony in advanced ant societies. Proc Natl Acad Sci USA 105:17884–17889.

Khila A, Abouheif E. 2010. Evaluating the role of reproductive constraints in ant social evolution. Philos Trans R Soc Lond B Biol Sci 365:617–630.

Kim DH, Sarbassov DD, Ali SM, King JE, Latek RR, Erdjument-Bromage H, Tempst P, Sabatini DM. 2002. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. Cell 110:163–175.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.

Ledón-Rettig CC, Pfennig DW. 2011. Emerging model systems in eco-evo-devo: the environmentally responsive spadefoot toad. Evol Dev 13:391–400.

Lesoway MP, Abouheif E, Collin R. 2014. The development of viable and nutritive embryos in the direct developing gastropod *Crepidula navicella*. Int J Dev Biol 58:601–611.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323.

Mann K, Edsinger-Gonzales E, Mann M. 2012. In-depth proteomic analysis of a mollusc shell: acid-soluble and acid-insoluble matrix of the limpet *Lottia gigantea*. Proteome Sci 10:28.

Miyakawa H, Imai M, Sugimoto N, Ishikawa Y, Ishikawa A, Ishigaki H, Okada Y, Miyazaki S, Koshikawa S, Cornette R. 2010. Gene up-regulation in response to predator kairomones in the water flea, *Daphnia pulex*. BMC Dev Biol 10:45.

Moczek AP, Nijhout HF. 2002. Developmental mechanisms of threshold evolution in a polyphenic beetle. Evol Dev 4:252–264.

Monnier PP, Sierra A, Macchi P, Deitinghoff L, Andersen JS, Mann M, Flad M, Hornberger MR, Stahl B, Bonhoeffer F, Mueller BK. 2002. RGM is a repulsive guidance molecule for retinal axons. Nature 419:392–395.

Moroz LL, Edwards JR, Puthanveettil SV, Kohn AB, Ha T, Heyland A, Knudsen B, Sahni A, Yu F, Liu L. 2006. Neuronal Transcriptome of *Aplysia*: neuronal compartments and circuitry. Cell 127:1453–1467.

O'Brien EK, Degnan BM. 2003. Expression of Pax258 in the gastropod statocyst: insights into the antiquity of metazoan geosensory organs. Evol Dev 5:572–578.

Perry JC, Roitberg BD. 2006. Trophic egg laying: hypotheses and tests. Oikos 112:706–714.

Pfennig DW. 1992. Polyphenism in spadefoot toad tadpoles as a logically adjusted evolutionarily stable strategy. Evolution 46:1408–1420.

Poon MW, Tsang WH, Chan SO, Li HM, Ng HK, Waye MM. 2011. Dyslexia-associated KIAA0319-like protein interacts with axon guidance receptor Nogo Receptor 1. Cell Mol Neurobiol 31:27–35.

Rivest BR. 1983. Development and the influence of nurse egg allotment on hatching size in *Searlesia dira* (Reeve, 1846) (Prosobranchia, Buccinidae). J Exp Mar Biol Ecol 69:217–241.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

Saras J, Franzen P, Aspenstrom P, Hellman U, Gonez LJ, Heldin CH. 1997. A novel GTPase-activating protein for Rho interacts with a PDZ domain of the protein-tyrosine phosphatase PTPL1. J Biol Chem 272:24333–24338.

Schmieder R, Edwards R. 2011. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PLoS ONE 6:e17288.

Serafini T, Kennedy TE, Gaiko MJ, Mirzayan C, Jessell TM, Tessier-Lavigne M. 1994. The netrins define a family of axon outgrowth-promoting proteins homologous to *C.* elegans UNC-6. Cell 78:409–424.

Shamloula HK, Mbogho MP, Pimentel AC, Chrzanowska-Lightowlers ZMA, Hyatt V, Okano H, Venkatesh TR. 2002. Rugose (rg), a *Drosophila* A kinase anchor protein, is required for retinal pattern formation and interacts genetically with multiple signaling pathways. Genetics 161:693–710.

Shen Z, Jacobs-Lorena M. 1998. A type I peritrophic matrix protein from the malaria vector *Anopheles gambiae* binds to chitin: cloning, expression, and characterization. J Biol Chem 273:17665–17670.

Smith HL, Gibson GD. 1999. Nurse egg origin in the polychaete *Boccardia proboscidea* (Spionidae). Invertebr Reprod Dev 35:177–185.

Smith KE, Thatje S. 2013. Nurse egg consumption and intracapsular development in the common whelk *Buccinum undatum* (Linnaeus 1758). Helgol Mar Res 67:1–12.

Spade DJ, Griffitt RJ, Liu L, Brown-Peterson NJ, Kroll KJ, Feswick A, Glazer RA, Barber DS, Denslow ND. 2010. Queen Conch (*Strombus gigas*) testis regresses during the reproductive season at nearshore sites in the Florida keys. PLoS ONE 5:e12737.

Spight TM. 1976. Hatching size and the distribution of nurse eggs among prosobranch embryos. Biol Bull 150:491–499.

Stabell OB, Ogbebo F, Primicerio R. 2003. Inducible defences in *Daphnia* depend on latent alarm signals from conspecific prey activated in predators. Chem Senses 28:141–153.

Supek F, Bosnjak M, Skunca N, Smuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE 6:9.

Suzanne M, Steller H. 2013. Shaping organisms with apoptosis. Cell Death Differ 20:669–675.

Tajrishi MM, Tuteja R, Tuteja N. 2011. Nucleolin: the most abundant multifunctional phosphoprotein of nucleolus. Commun Integr Biol 4:267–275.

Teixeira LK, Reed SI. 2013. Ubiquitin ligases and cell cycle control. Annu Rev Biochem 82:387–414.

Véliz D, Winkler FM, Guisado C. 2003. Developmental and genetic evidence for the existence of three morphologically cryptic species of *Crepidula* in northern Chile. Mar Biol 143:131–142.

Véliz D, Winkler FM, Guisado C, Collin R. 2012. A new species of *Crepipatella* (Gastropoda: Calyptraeidae) from northern Chile. Molluscan Res 32:145–153.

Wang X, Liu B, Wang G, Tang B, Xiang J. 2008. Molecular cloning and functional analysis of Cathepsin B in nutrient metabolism during larval development in *Meretrix meretrix*. Aquaculture 282:41–46.

Wang X, Liu B, Tang B, Xiang J. 2011. Potential role of Cathepsin B in the embryonic and larval development of clam *Meretrix meretrix*. J Exp Zool B Mol Dev Evol 316B:306–312.

Welch JM, Wang D, Feng G. 2004. Differential mRNA expression and protein localization of the SAP90/PSD-95-associated proteins (SAPAPs) in the nervous system of the mouse. J Comp Neurol 472:24–39.

Wesche H, Gao X, Li X, Kirschning CJ, Stark GR, Cao Z. 1999. IRAK-M is a novel member of the Pelle/interleukin-1 receptor-associated kinase (IRAK) family. J Biol Chem 274:19403–19410.

West DL. 1983. Reproductive biology of *Colus stimpsoni* (Prosobranchia, Buccinidae): 5. Nutritive Egg Formation. Veliger 25:299–306.

West-Eberhard MJ. 1986. Alternative adaptations, speciation, and phylogeny (A Review). Proc Natl Acad Sci USA 83:1388–1392.

West-Eberhard MJ. 2003. Developmental plasticity and evolution. Oxford: Oxford University Press.

West-Eberhard MJ. 2005. Developmental plasticity and the origin of species differences. Proc Natl Acad Sci USA 102:6543–6549.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 11:R14.

Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, Xiong Z, Que H, Xie Y, Holland PWH, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C, Meng J, Yang L, Liu J, Wen B, Zhang N, Huang Z, Zhu Q, Feng Y, Mount A, Hedgecock D, Xu Z, Liu Y, Domazet-Loso T, Du Y, Sun X, Zhang S, Liu B, Cheng P, Jiang X, Li J, Fan D, Wang W, Fu W, Wang T, Wang B, Zhang J, Peng Z, Li Y, Li N, Wang J, Chen M, He Y, Tan F, Song X, Zheng Q, Huang R, Yang H, Du X, Chen L, Yang M, Gaffney PM, Wang S, Luo L, She Z, Ming Y, Huang W, Zhang S, Huang B, Zhang Y, Qu T, Ni P, Miao G, Wang J, Wang Q, Steinberg CEW, Wang H, Li N, Qian L, Zhang G, Li Y, Yang H, Liu X, Wang J, Yin Y, Wang J. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490: 49–54.

Zhang Y, Meng Q, Jiang T, Wang H, Xie L, Zhang R. 2003. A novel ferritin subunit involved in shell formation from the pearl oyster (*Pinctada fucata*). Comp Biochem Phys B: Biochem Mol Biol 135:43–54.