



SYMPOSIUM

SeaBase: A Multispecies Transcriptomic Resource and Platform for Gene Network Inference

Antje H. L. Fischer,^{*,†} Dmitry Mozzherin,^{*} A. Murat Eren,^{*} Kristen D. Lans,^{*} Nathan Wilson,^{*} Carlo Cosentino[‡] and Joel Smith^{1,*}

^{*}Marine Biological Laboratory, Woods Hole, MA 02543, USA; [†]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA; [‡]Systems & Control Engineering, University of Magna Graecia, 88100 Catanzaro, Italy

From the symposium “A New Organismal Systems Biology: How Animals Walk the Tight Rope between Stability and Change” presented at the annual meeting of the Society for Integrative and Comparative Biology, January 3–7, 2014 at Austin, Texas.

¹E-mail: jcsmith@mbl.edu

Synopsis Marine and aquatic animals are extraordinarily useful as models for identifying mechanisms of development and evolution, regeneration, resistance to cancer, longevity and symbiosis, among many other areas of research. This is due to the great diversity of these organisms and their wide-ranging capabilities. Genomics tools are essential for taking advantage of these “free lessons” of nature. However, genomics and transcriptomics are challenging in emerging model systems. Here, we present SeaBase, a tool for helping to meet these needs. Specifically, SeaBase provides a platform for sharing and searching transcriptome data. More importantly, SeaBase will support a growing number of tools for inferring gene network mechanisms. The first dataset available on SeaBase is a developmental transcriptomic profile of the sea anemone *Nematostella vectensis* (Anthozoa, Cnidaria). Additional datasets are currently being prepared and we are aiming to expand SeaBase to include user-supplied data for any number of marine and aquatic organisms, thereby supporting many potentially new models for gene network studies. SeaBase can be accessed online at: <http://seabase.core.cli.mbl.edu>.

Introduction

Newly developed sequencing techniques, such as next-generation sequencing, which allow sampling of the entire transcriptome, are currently reducing the “barriers to entry” for complex systems in any number of new model organisms (Ekblom and Galindo 2011). These techniques allow researchers for the first time to acquire data on an appropriate scale. However, the handling and analyses of these datasets remains challenging.

Why marine models?

The sea provides an abundance of diverse organisms suitable as models for addressing many diverse questions in life science. One of the best-studied marine organisms is the sea urchin, which has been used as a model for developmental biology for over 100 years (McClay 2011). Sea urchins have come to represent

one of the pioneering models for the application of functional genomics to determine the gene regulatory network (GRN) for embryogenesis (Davidson et al. 2002, 2010). The features that have made the sea urchin an excellent model—e.g., easy to control fertilization, synchronized cultures in simple conditions that are easy to follow, and simple gene delivery by microinjection. (McClay 2011)—are shared by many aquatic animals.

Aquatic organisms hold promise as new models for several biomedically relevant fields, such as regeneration, cancer research, and aging. Marine organisms like starfish and sea cucumbers will greatly complement our understanding of regenerative processes (Lawrence 1991; García Arrarás et al. 1998, 1999; Hernroth et al. 2010; Mashanov et al. 2012; Sun et al. 2013). Sea cucumbers extrude their viscera as a defense mechanism, followed by major

regenerative organogenesis within a few weeks (García Arrarás et al. 1998, 1999). Starfish can regenerate lost arms, and some are even capable of regenerating the entire body from just one arm, using this as a defense mechanism or even to reproduce asexually via fission (Mladenov et al. 1986; Carneveli 2006; Hernroth et al. 2010). Other marine organisms such as annelids and the lamprey also exhibit impressive regenerative capabilities (Bely 2006). Lampreys are the only vertebrate known to achieve full functional recovery following injury to the spinal cord (Selzer 1978; Lurie and Selzer 1991; Bloom and Morgan 2011; Lau et al. 2013; Zhang et al. 2013); although, interestingly, not all neurons survive the injury (Barreiro-Iglesias and Shifman 2012; Busch and Morgan 2012). The study of these organisms can greatly complement biomedical understanding of regenerative processes.

Amphibians like axolotl and newt are excellent models for controlled studies of the resistance to cancer; see overview by Oviedo and Beane (2009). In lentectomized eyes of newt only the dorsal iris regenerates the new lens, while simultaneously the ventral iris, which is incapable of regeneration, forms tumors when carcinogenic chemicals are applied during regeneration (Okamoto 1997). Thus, amphibian models illustrate perfectly the well-known nexus between regeneration, development, and cancer biology (Waddington 1935), with cancer-resistance in amphibians being linked directly to regenerative processes and potential (Brookes 1998). Marine organisms might also offer new avenues in aging research. Classical aging models that have been used to map out important genes and pathways are *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster* (Guarente and Kenyon 2000; Kennedy 2008; Austad 2009). Although these examples have a short lifespan, other animals (including a large number of marine species) offer excellent models of longevity (Bodnar 2009). Several species of non-colonial marine invertebrates can live for more than 100 years, and some of these grow and reproduce during their entire life (Bodnar 2009). Studying these animals may reveal new and efficient strategies to understand and slow down aging (Bodnar 2009). In some cases, for example, as in asexually reproducing *Hydra vulgaris*, there seems to be a lack of aging mechanisms entirely; thus, longevity is extreme (Martínez 1998). The oldest recorded non-colonial animal is a marine organism: *Arctica islandica*, the ocean quahog clam, which was 405 years old when it was collected (Wanamaker et al. 2008). In fact, the lifespan varies widely among bivalves and even closely related species

differ markedly. While the scallop *Argopecten irradians* has a lifespan of only 18–22 months, the closely related *Argopecten purpuratus* lives for up to 7 years (Estabrooks 2007). Future studies on length of the telomere, stem-cell function, and the response to reactive oxygen species or pathways regulating specific factors such as sirtuins, offer points of comparison with mammalian studies. There also are discovery-efforts designed to identify novel factors and pathways that regulate longevity.

The use of marine models in evolutionary developmental biology (evo–devo)

While biomedical research still focuses on a small number of model organisms and is only now beginning to slowly include new model systems, “evo–devo” researchers have begun to increase the number of model species dramatically—and in fact many of those new models are marine organisms. “Evo–devo” research began to rise in the 1980s, when the *Hox* gene family was discovered and researchers realized that the same regulatory genes were shared across all major metazoan taxa and that developmental data could contain a strong and marked evolutionary signal (Raff 2000). At the center of “evo–devo” is the attempt to understand the molecular basis for differences in the organization of animal body plans and the origin of novel structures (Müller 2007; Martindale and Hejnol 2009). Non-bilaterian animals such as sponges, ctenophores and cnidarians are of particular interest (Miller 2009) and all of them are aquatic, predominantly marine, organisms. A selection of emerging non-bilaterian “evo–devo” model systems is listed in Table 1.

“Evo–devo” research also began to expand the focus from ecdysozoans (fly and *C. elegans*) and deuterostomes (mice, frogs, and fish) to the third major but yet understudied bilaterian taxon—the Lophotrochozoa—to study the origin of segmentation, sense organs and the centralized nervous system. Lophotrochozoans include several taxa with spiral cleavage such as mollusks, annelids, nemerteans, and planarians and other taxa such as bryozoans, brachiopods, and phoronids (Tessmar-Raible and Arendt 2003; Dunn et al. 2008; Helmkamp et al. 2008). Examples of emerging lophotrochozoan models are summarized in Table 1. New marine deuterostomes and ecdysozoan models begin to complement our views on animal evolution and Table 1 provides a small selection of new model species.

While “evo–devo” is paving the way for a broad selection of new model organisms, far more examples motivating research into new model systems,

Table 1. A selection of marine evo–devo model systems

Taxon	Species name	Citation
Non-bilaterian models		
Porifera (sponge)		
Homosclero-morpha	<i>Oscarella lobularis</i>	(Ereskovsky et al. 2009)
Demosponge	<i>Amphimedon queenslandica</i>	(Gauthier and Degnan 2008; Adamska et al. 2010, 2011)
Ctenophora (comb jelly)	<i>Mnemiopsis leydii</i>	(Pang and Martindale 2008; Fischer et al. 2014)
	<i>Beroe ovata</i>	(Pang and Martindale 2008)
	<i>Pleurobrachia pileus</i>	(Pang and Martindale 2008)
Cnidaria		
Anthozoa	<i>Nematostella vectensis</i>	(Darling et al. 2005; Genikhovich and Technau 2009b; Fischer et al. 2013; Stefanik et al. 2013; Wolenski et al. 2013)
	<i>Edwardsiella lineata</i>	(Reitzel et al. 2009; Stefanik et al. 2014)
	<i>Acropora millepora</i>	(Miller and Ball 2000)
Hydrozoa	<i>Hydra vulgaris</i>	(Galliot and Schmid 2002; Steele 2002; Bosch 2009)
	<i>Hydra magnipapillata</i>	(Galliot and Schmid 2002; Steele 2002; Bosch 2009)
	<i>Hydractinia echinata</i>	(Frank et al. 2001; Galliot and Schmid 2002)
	<i>Podocoryne carnea</i>	(Masuda-Nakagawa et al. 2000; Galliot and Schmid 2002)
	<i>Clytia hemisphaerica</i>	(Houliston et al. 2010)
Cubozoa	<i>Tripedalia cystophora</i>	(Piatigorsky and Kozmik 2004)
Scyphozoa	<i>Aurelia</i> sp.	(Yuan et al. 2008)
Bilaterian models		
Lophotrochozoa		
Annelida	<i>Platynereis dumerilii</i>	(Arendt et al. 2002; Raible et al. 2005; Schneider and Bowerman 2007; Dray et al. 2010; Fischer and Henrich 2010)
	<i>Capitella teleta</i>	(Tsutsumi and Kikuchi 1984; Dill et al. 2007; Meyer and Seaver 2009, 2010)
	<i>Hydroidis elegans</i>	(McDougall et al. 2008)
Gastropoda	<i>Crepidula fornicata</i>	(Hejnal et al. 2007; Henry and Perry 2008; Henry et al. 2010; Lyons et al. 2012)
	<i>Lottia gigantea</i>	(Grande and Patel 2009)
	<i>Ilyanassa obsoleta</i>	(Sweet 1998; Lambert and Nagy 2001; Dickinson and Croll 2003; Goulding 2009; Lambert 2009)
	<i>Patella vulgata</i>	(Damen 1996; Damen and Dictus 1996; Dictus and Damen 1997; Lartillot et al. 2002)
Cephalopoda	<i>Sepia officinalis</i>	(Grimaldi et al. 2008; Bassaglia et al. 2013)
Bivalvia	<i>Mytilus</i>	(Bierne et al. 2002; Dyachuk and Odintsova 2009)
Scaphopoda	<i>Antalis entails</i>	(Wanninger and Haszprunar 2002; Gonzales et al. 2007)
Polyplaco-phora	<i>Mopalia muscosa</i>	(Friedrich et al. 2002; Haszprunar et al. 2002)
	<i>Chiton olivaceus</i>	(Haszprunar et al. 2002)
Nemertini	<i>Cerebratulus lacteus</i>	(Henry and Martindale 1998; Jondelius et al. 2004)
	<i>Nemertoderma westbladi</i>	(Henry and Martindale 1998; Jondelius et al. 2004)
Platyhelminths	<i>Maritigrella crozieri</i>	(Rawlinson 2010; Lapraz et al. 2013)
Bryozoa	<i>Membranipora membranacea</i>	(Gruhl 2010)
Brachiopoda	<i>Terebratalia transversa</i>	(Freeman 1993; Altenburger and Wanninger 2009)
	<i>Novocrania anomala</i> (<i>Crania anomala</i>)	(Nielsen 1991; Freeman 2000)
Phoronida	<i>Phoronis vancouverensis</i>	(Freeman and Martindale 2002; Santagata 2004)

(continued)

Table 1. Continued

Taxon	Species name	Citation
Acoela	<i>Neochildia fusca</i>	(Boyer et al. 1996; Henry et al. 2000)
	<i>Symsagittifera roscoffensis</i>	(Semmler et al. 2008)
	<i>Convolutriloba longifissura</i>	(Hejnal and Martindale 2008, 2009; Sikes and Bely 2008)
Deuterostomia		
Urochordata	<i>Ciona intestinalis</i>	(Satoh 2003; Kumano and Nishida 2007)
	<i>Oikopleura dioica</i>	(Stach et al. 2008)
Echinoderms		
Asteroida	<i>Patiria miniata</i>	(Hinman and Davidson 2007; Hinman et al. 2007; Yankura et al. 2010)
Echinoida	<i>Strongylocentrotus purpuratus</i>	(Davidson et al. 1998; Angerer and Angerer 2003; McClay 2011)
	<i>Heliocidaris erythrogramma</i>	(Wray and Raff 1990; Raff 1992)
Holuturia	<i>Apostichopus japonicus</i> (<i>Stichopus japonicus</i>)	(Nakano et al. 2006)
Hemichordata	<i>Saccoglossus kowalevskii</i>	(Lowe et al. 2003, 2006; Gerhart et al. 2005)
	<i>Ptychodera flava</i>	(Tagawa et al. 1998, 2001; Peterson et al. 1999)
Cephalo-chordata	<i>Branchiostoma floridae</i>	(Holland et al. 1992, 1994; Moret et al. 2004; Bertrand and Escriva 2011; Vopalensky et al. 2012)
	<i>Branchiostoma lanceolatum</i>	(Bertrand and Escriva 2011; Oulion et al. 2012)
Ecdysozoa		
Crustacea	<i>Artemia salina</i>	(Blanchard 1986; Harzsch and Glötzner 2002; Wildt and Harzsch 2002; Copf et al. 2003)
Nematoda	<i>Enoplus brevis</i>	(Voronov and Panchin 1998; Sommer 2000)
Priapulida	<i>Priapulidus caudatus</i>	(Wennberg et al. 2008)

This list focuses on non-bilaterian and lophotrochozoan systems, but also include some representatives of deuterostomes and ecdysozoans. Datasets of species in bold and underlined are already available on SeaBase. Datasets of species in bold are currently being processed to upload them to SeaBase.

especially marine animals, abound. One example is the development of the sea anemone *Aiptasia* as a model for symbiotic and commensal relationships (Sunagawa et al. 2009; Lehnert et al. 2012). Other examples are cuttlefish, squid, and Octopus as new models of neurophysiology and complex behaviors, such as learning and camouflage (Fiorito and Scotto 1992; Williamson and Chrachri 2004; Mäthger et al. 2009; Chiao et al. 2011).

The enabling technology of genomics

Functional genomics opens avenues of research into many organisms, promising valuable insight into any number of areas of biomedical and basic research. RNA-Seq and targeted genetic perturbations provide key tools for interrogating complex systems (Ozsolak and Milos 2011). This set of advanced technologies will provide new insights into the transcriptional regulatory network underlying processes like the regulation of development and growth, longevity, and regeneration. Importantly, these techniques can be applied with relatively few restrictions. Thus, the acquisition of data on the appropriate scale is in many cases not limiting. The bottleneck, however, is in

interpreting the vast amount of data generated (Zhang et al. 2012). This problem is especially limiting in non-model or emerging-model organisms.

Here, we introduce SeaBase, a web resource for transcriptomic data for aquatic animals. Our first goal is to enable more effective sharing of data to open new avenues for collaboration and for analysis of data. Our second goal is to support tools for analyzing transcriptomic data to determine testable models of the transcriptional regulatory network. Finally, we will briefly compare SeaBase with other databases. The first dataset on SeaBase is a developmental RNA-Seq time series of the sea anemone *Nematostella vectensis*, a non-bilaterian species. Similar datasets for other species currently are being processed and will be available in the near future.

Material and methods

Culture and spawning of *Nematostella*

Nematostella vectensis was cultured as previously described (Darling et al. 2005; Genikhovich and Technau 2009a; Stefanik et al. 2013). Animals were kept in the

dark at 17°C. Spawning was induced by exposure to 9 h of bright light in combination with a heat-shock (25°C). Male and female animals were kept separated to control for time of fertilization. All samples were collected in one synchronized spawning of several males and females and all eggs were fertilized synchronously. After fertilization the eggs were washed with 2% cystein solution (pH 7.4) for 7 min followed by three washes with sterile filtered 50% natural seawater. The eggs and embryos were kept at an ambient temperature of 17°C ± 0.5°C.

Collection of samples, extraction of mRNA, preparation, and sequencing of the libraries

For each sample, 300 embryos were counted and collected. Sampling began with unfertilized eggs (0 h), and was continued at 1 h intervals until 19 h. Eggs for the 0 h time-point were not fertilized before removing the gelatinous mass. For most samplings, two samples (A and B) were collected. Samples for the same time-point were always collected within 10 min of each other. Only one sample per time-point was available from 3 h to 10 h. The collected eggs and embryos, respectively, were centrifuged at 2000 rpm for 30 s, supernatant water removed, and 300 µL lysis buffer, supplied by the Dynabead kit (Invitrogen), added. The tissue was homogenized with a pestle. The samples were stored in lysis buffer at -80°C. The libraries were prepared as described previously (Tulin et al. 2013). The libraries underwent paired-end sequencing with 100-bp read length on an Illumina HighSeq 1000. The raw data are available at WHOAS (Woods Hole open access server), doi: 10.1575/1912/5981.

Analyses of bioinformatics

After trimming and removing reads of low quality, ribosomal RNA and other contaminating RNAs as described by Tulin et al. (2013) using Bowtie 2.0.0-beta6 (Langmead et al. 2009), the reads were mapped against a *Nematostella* reference transcriptome (Tulin et al. 2013). Fragments per kilobase of exon per million fragments mapped (FPKM) values were determined using RSEM (Li and Dewey 2011) and further quantified based on the known concentration of the ERCC controls. Finally, the number of molecules per embryo was determined and plotted over time as described by Tulin et al. (2013).

Database

SeaBase is a web-application. It is built using a variety of open source technologies. The code of SeaBase itself is licensed under the open source

MIT license and is freely available for study or reuse. Changes to the code are tracked in a code repository using the revision control system Git (<http://git-scm.com>) and the code is publicly available at <https://github.com/EOL/seabase>. Transcript sequences, results of the regression analysis, and information about orthology are stored using a MySQL relational database. The Graphical User Interface of the web application is created using the Sinatra framework written in Ruby. Several open source libraries are used for communication between the database and the web front end.

The integrated graphs of the data are generated by Google Charts (<https://developers.google.com/chart/>), a proprietary JavaScript graphing library with a freely available API (application programming interface), as well as the open source D3.js library (<http://d3js.org/>). Network graph analysis is performed using the Gephi API (<http://gephi.org>). Blast searches against stored data are performed with code from [wwwblast](http://www.ncbi.nlm.nih.gov/staff/tao/URLAPI/wwwblast/) (Altschul et al. 1990, 1997; Schäffer et al. 2001). SeaBase is served using the Nginx web server and Unicorn application server on an Ubuntu Linux machine.

Results

The sea anemone *N. vectensis* is the first organism hosted by SeaBase

SeaBase is available online at: <http://seabase.core.cli.mbl.edu> (Fig. 1). SeaBase can accommodate transcriptome data for any number of aquatic animals. For the pilot, we have used data from the starlet sea anemone, *N. vectensis*, an emerging model system for development, evolution, genomics, regeneration, and ecology (Darling et al. 2005; Genikhovich and Technau 2009b; Gilmore et al. 2013; Stefanik et al. 2013). The sampling methods and data analyses for this test-case dataset are described in the “Material and methods” section. Notably, these data were collected as part of a high-resolution time-series with 20 time-points, one sample per hour from fertilization to the onset of gastrulation. A primary function of SeaBase is visualizing time-series expression plots.

Sharing and searching of genes

One of the major goals of SeaBase is to provide a platform for sharing and searching transcriptome data. SeaBase can be searched by UniProt accession number, gene annotation, or protein name (Fig. 2). For example, “polymerase alpha”, “P17918” and “Pcna” are all valid searches. An autofill function helps users rapidly identify genes of interest (Fig. 2). SeaBase annotations are currently based on

SeaBase provides a platform for sharing, searching and analyzing quantitative gene expression data. Please feel free to look around, find your favorite gene (or genes) and plot their temporal expression, or browse RNA-seq data from your favorite marine/aquatic organism.

Have data to contribute? Please let us know. We will be happy to include it.

We currently support a plotting function, e.g., for differential expression analysis or for time series expression plots. We will be adding more functions soon. Next on the list is a clustering tool for identifying synexpression networks. Stay tuned: more tools - all geared for determining Gene Networks - are to follow.

Interested in the code? Checkout our [GitHub repository](#)

For updates and to give us suggestions, follow us on Twitter @MBLSeaBase.

Marine Biological Laboratory | MIT License | Version 0.0.2

Fig. 1 SeaBase is a new online database that is currently hosting a developmental RNA-Seq time-series for *N. vectensis*. It can be accessed at: <http://seabase.core.cli.mbl.edu>. SeaBase can be searched by gene name or via BLAST.

Search term can be a UniProt Accession, Gene annotation, or Protein name. For example 'polymerase alpha',

Found ortholog page displays dynamics of transcription for the first 20 hours of embryogenesis starting from t

Marine Biological Laboratory | MIT License | Version 0.0.2

A2APC3:TtI19:Probable tubulin polyglutamylase TTL9
A4Q9E4:TtI12:Probable tubulin polyglutamylase TTL2
A4Q9E5:TtI13:Tubulin monoglycylase TTL3
A4Q9E8:TtI6:Tubulin polyglutamylase TTL6
A4Q9F0:TtI7:Tubulin polyglutamylase TTL7
A4Q9F4:TtI11:Tubulin polyglutamylase TTL11
A4Q9F6:TtI13:Tubulin polyglutamylase TTL13
P05214:Tuba3a:Tubulin alpha-3 chain
P20357:Map2:Microtubule-associated protein 2
P38585:TtI:Tubulin--tyrosine ligase

Fig. 2 SeaBase can be searched via search terms such as protein names, UniProt Accession number or gene annotation. An autofill function helps to identify available candidates.

hits in UniProt databases. User-supplied datasets may also use UniProt annotations or annotations from genome projects or other user-supplied annotation files.

SeaBase can be searched by BLAST

SeaBase locally runs NCBI BLAST (Basic Local Alignment Search Tool) to search all comp_IDs stored on SeaBase (Fig. 3). Each comp_ID identifies a unique sequence that has been assembled by Trinity (Grabherr et al. 2011), representing transcripts or splice variants of transcripts and transcript fragments. The BLAST function allows searching of the entire dataset and the ability to identify orthologs, paralogs, and splice variants of particular genes of interest.

SeaBase searches, whether via BLAST or a specific search term, return the developmental transcription level profile (number of molecules per embryo over developmental time) for any query, as well as all assembled sequences associated with that gene in FASTA format (Fig. 4).

Plotting functions

The second set of functions supported by SeaBase consists of basic plotting tools. As the current datasets represent the collection of RNA-Seq samples as a developmental time-series, the main purpose in the current release is to generate temporal expression plots. Thus, the expression levels of a gene of interest may be plotted over time (Fig. 4).



Fig. 3 SeaBase can be searched by BLAST. BLAST may be used to identify different orthologs, paralogs, or splice variants of any given gene. Each BLAST hit (the list of comp_IDs shown in blue on the bottom left) can be activated and will return the temporal expression profile of that comp_ID.

Cluster analysis

SeaBase aims to facilitate the analysis of gene network. Identifying genes with similar temporal profiles of expression can provide insight into the transcriptional regulatory network. For example, such profiles could indicate waves of zygotic gene activation.

SeaBase offers a clustering method based on a cosine similarity metric to determine sets or “clusters” of genes with similar expression profiles. The cosine similarity metric works as follows: each gene within a time-series dataset is represented by a vector in N -dimensional space. N is the number of dimensions depending on the number of time-intervals in the time-series. For the *N. vectensis* time-series, 20

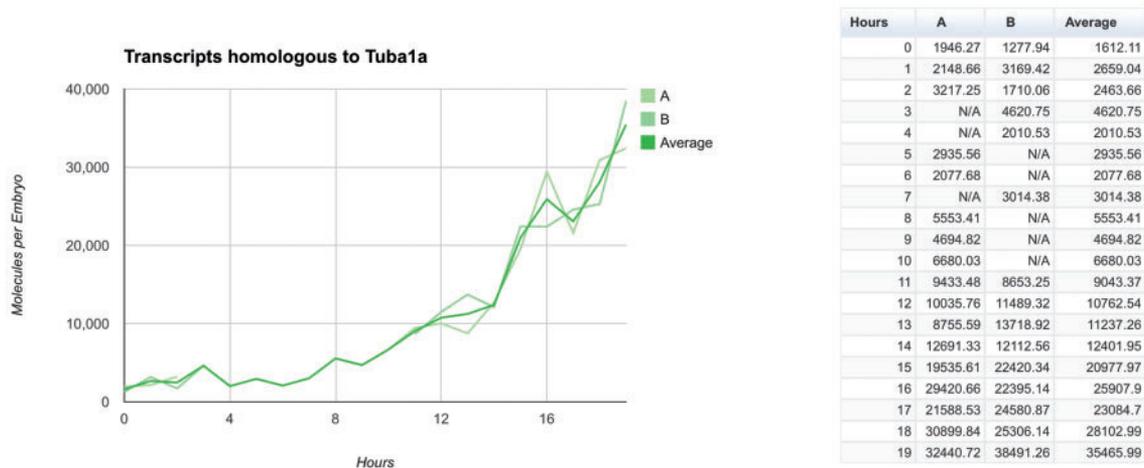
time-points were collected; thus, a 19-dimensional space is used to analyze the data. For each gene, the change in expression over each sampling-interval is represented by the corresponding coordinate of the associated vector in multi-dimensional space. We collected *Nematostella* samples every hour. Therefore, our sampling interval is 1 h and changes in level of gene expression over each 1-h interval is represented by their position within the 19-dimensional space. The cosine is determined for the angle between two vectors of the multi-dimensional space for all possible pairs of gene. Genes then are grouped, based on similar changes over time. Note that the cosine measure does not take the magnitude of the vector into account, but only the orientation. This implicitly introduces a

SeaBase
Home Search Blast

Search for *Nematostella vectensis* transcripts by Mouse orthologs Name: Search

***Nematostella vectensis* transcripts homologous to the mouse ortholog Tuba1a**

Mouse Ortholog: Tuba1a
UniProt Outlink: P68369: Tubulin alpha-1A chain



Transcripts

comp877_c0_seq1

```
>comp877_c0_seq1
TCAAGACGCTTCAATCCAAATATTTAGTGTGACGCTTGAATTAACCAATTAGAGCGTCGTAATAGTGGACAAAGAACAATTAGCATAACGAATCATATTTACACGCACCCAGTCAC
GAGGTGAGTCATTTACGCTTGCACCAATACTGTAGAGAGATCGTAACCTTACTACAGTTCGGAGCACTTCCATTACAGTTATATCCAAATCTCCTTCGAAATCAACCAACT
AATACAAAATCGGTGAGTGTATCTCTATCCACGTTGGACAAGCCGGTGTGCAGATTGGCAACGCCTGTTGGGAGTTGACTGCCTCGAACACGGAAATTACGCCCGATGGCCAGAT
GCCGAGCGACAAGACGATCGGTGGCGGCGATGACTCGTTCAACACTTCTTTAGCGAGACCCGCGCTGAAAGCACGTCGCCAGAGCTGTCTTTGTTGATTTAGAGCCAAACAGTA
GTCGATGAGGTTTCGCACCTGGAACCTACCCGACGCTTTTCCACCTGAGCAACTTATCACTGGCAAGAGGATGCTGCTAACAACTACGCCCGTGGACACTACACTGTCGAAAAG
AGTTGATCGATCTGGTGTCTGACAGAATCCGTAACCTGGCCGATCAATGACTGGTCTTCAAGGTTTCTTGTATCTCCATTCTTTCCGTTGGTGGCACGGGATCTGGATTTCCCTC
ACTGTTGATGGAACGCTGTCTGTTGACTACGGCAAGAAATCCAAGCTGGAGTTTGGCCATCACCCAGCTCCTCAGATCTCCACCGCTGTGGTTGAGCCCTACAACCTCCCTG
ACCACCCACACAACCTGGAGCACTCTGACTGTGCCTCATGGTAGACAACGAGGCCATCTATGACATCTGCCGTCGCAACCTGGACATGAAAGACCCCACTACACCAACCTGA
ACCGCTGATTGGCCAGATCGTGTCCCTCCATCACTGCCTCCCTGCGCTTTGATGGTGCCTGAATGTGGATCTGACTGAGTTCCAGACCAACTTGGTGCCTACCCCTCGTATCCA
CTTCCCTTAGCCACCTATGCCCAGTCACTCTGCTGAGAAGGCTACCACGAGCAACTGAGTGTGCTGAGATCACCAATGCCTGCTTTGAGCCCGCAACAGATGGTGA
TGTGACCCACGTCATGGCAAGTACATGGCCGCTGCTGCTGCTTTACCGTGGTGTGTTGTCGCCAAGGATGCAATGTGCCATTGGCCACATCAAGCAACGAGAGAACCATCCAGT
TCGTTGATGGTGGCCCACTGGATTCAGGTCGGCATCAACTACCGCCACCTACTGTTGCTCCCTGGCCGTCACCTGGCCAAGGTGCAGCCTGCCGTGGATTTGAGTAACAC
CACAGCCATCGCTGAGGCTGGGCTCGTCTGGATCACAAGTTGATGATGACGCAAGCGTGCCTTTGCTCCACTGGTATGTGGGTGAGGGTATGGAGGAAGGAGATTTCTCT
GAGGCTCGTGAGGATTCGGCAGCTCTTGAGAAGGATTACGAGGAGTCCGGTGTGACAGTGTGAGG
```

comp14036_c0_seq1

```
>comp14036_c0_seq1
GTAGAGCAGACAGCAGGCCATGTGCTTTCCGTGACGTGGGTACACTTAACCATCTGGTTAGCGGGTTCGAAGCAGGCATTGGTGATCTCGGCAACAGTGAGCTGCTCATGGTAA
GCCTTCTCAGCTGAGATGACAGGGGCATAGGTAACCAATGGGAAATGGATACGTTGGTATGGCCCAAGTTGGTCTGGAAATTCAGTGAGATCAACGTTTAAAGCCCATCAAAAAC
GAAGGCTGGCGGTGATGGAGAAACGATTTGACCAATGAGTCTGTTCAAGTTGGTGAAGTTGGGCGCTCTATGTCCAAAGTTACGGCGACAATAATCATAGATGGCCTCGTTGTC
GACCATAAAAGCCGAGTCAGAGTCTCCAAAGTGGTATGAGTTGTCAGATGGAATGTACGGCTCAACACAGGCAGTAGAGACCTGTGGTGTGGGTAGACGGCAAAATTCGAGC
```

Fig. 4 SeaBase can show the expression profile of a transcript or a particular comp_ID in a graph over time. The image shows the expression of transcripts homologous to tubulin a 1a from 0h post-fertilization (hpf) to 19hpf with the number of transcripts per embryo (on the Y-axis) over time (X-axis). The data points are also represented in a table (on the right). All comp_IDs that contribute to the total number of transcripts per embryo are listed with name and sequence underneath the graph.

normalization step, which is certainly desirable when dealing with a transcriptomic dataset, because the expressions of two transcripts can easily differ by orders of magnitude. A beta-version of this tool is currently online and a clustering dataset with 0h offset and 99.5% similarity is currently available under “Export”. This dataset can be exported and viewed using Gephi.

There are a number of advantages to clustering by this algorithm. First, it allows the generation of a

pair-wise correlation matrix for a window of expression. In addition, it allows for computing a pair-wise correlation matrix that incorporates delays. A 1h difference in time between windows may be informative for network inference. These values may be used to compute Granger causality (Granger 1988), to identify those pairs of gene that are correlated with a delay in time. A clustering dataset with 1h offset and 99% similarity can be downloaded from SeaBase.

Network inference

As mentioned previously, SeaBase uses cosine similarity to evaluate pair-wise correlations between every pair of genes, including different time-windows and with an option for delay. These data can be represented in graphical form, wherein the nodes of the network are the genes, and the edge connecting any two nodes represent the strength of the correlation between those two genes. Thus, these data can be used to infer a correlation network. SeaBase provides for the export of these data in an XML file for use with network graphing tools, such as Gephi (Bastian et al. 2009) or Cytoscape (Shannon et al. 2003; Smoot et al. 2011). As mentioned above, two examples of datasets are currently available for download.

SeaBase can host additional datasets

In addition to the *Nematostella* expression profile, SeaBase is open to host data from other animals and we are currently preparing to upload additional datasets. Users wishing to upload their own data should follow directions on SeaBase to contact the maintainers of the site.

Discussion

SeaBase is an online database (<http://seabase.core.cli.mbl.edu>) that hosts transcriptomic datasets for aquatic animals. SeaBase offers various tools to share, search, and analyze transcriptomic data and to support studies of gene regulatory network in non-model and emerging-model organisms.

The first dataset on SeaBase is a high-density developmental RNA-Seq time course of the sea anemone *N. vectensis*, covering developmental stages in 1-h intervals starting from unfertilized eggs and extending to the onset of gastrulation. To our knowledge, this is the first time transcriptional profiling has been done at such high density for an emerging model organism. We are currently preparing to add four more transcriptional profiling datasets for other marine invertebrates into SeaBase.

The transcriptome of *Edwardsiella lineata*, a parasitic anemone and the sister species of *N. vectensis* has recently been sequenced and is available on EdwardsiellaBase (Stefanik et al. 2014). EdwardsiellaBase offers a number of helpful search and browsing functions as well as the possibility for users to submit entries (Stefanik et al. 2014). While EdwardsiellaBase focuses on only one species, SeaBase is designed to host several species and will offer a comprehensive toolset to analyze transcriptomic datasets and infer gene interactions.

Next-generation sequencing has revolutionized transcriptomics since this technique was released in 2008 (Mortazavi et al. 2008; Sultan et al. 2008; Wilhelm et al. 2008). *De novo* transcriptome approaches became a particularly viable option for organisms without a reference genome or that have a poorly assembled genome (McGettigan 2013). Among animals without a sequenced reference genome is a large number of emerging marine model systems, many of which are listed in Table 1. Next-generation sequencing is therefore especially valuable for broadening biological sampling and extending the list of systems from the “big five” model species (*D. melanogaster*, *C. elegans*, *Xenopus*, *Danio rerio*, and *Mus musculus*) to multiple species, each particularly suitable for a specific biological question.

However, the amount of data next-generation sequencing creates has also caused problems for the management and storage of data (McGettigan 2013). SeaBase addresses this problem and offers a place to host, manage, and analyze RNA-Seq datasets for several species.

Online resources and databases have become incredibly valuable over the past three decades. Huge genomic databases like GenBank, launched in 1986 (Bilofsky et al. 1986; Bilofsky and Christian 1988) and Ensembl, launched in 2000 (Butler 2000; Hubbard and Birney 2000) as well as databases focused on one or a few closely related species like Flybase, launched in 1994 (Ashburner and Drysdale 1994; Consortium 1994) and Wormbase, launched in 2001 (Stein et al. 2001), are part of every researcher’s daily work. These databases are essential to host, store, find, share, and manage the large amount of data produced by the research community. The number of databases has been growing dramatically, reflecting a lasting demand. In March 2014, the Nucleic Acid Database collection included 1552 databases, including 58 new databases released in 2013 (Fernández-Suárez et al. 2014).

The main role of online databases is to make large datasets accessible to the research community (Baxevanis 2003) and sequencing efforts, such as the developmental transcriptome profile of *Nematostella*, will only be successful if the resulting data are available to the research community. Most importantly, online databases have to be easily searched and entries must be retrievable in a suitable, meaningful format (Baxevanis 2003).

SeaBase addresses these issues and aims to be the first database that provides the possibility to store, browse, and analyze RNA-Seq datasets for aquatic non-model animal systems all within one platform.

Acknowledgments

We want to thank Dr Sarah Tulin for technical advice on the preparation of the mRNA library. We want to thank Molly Philips and Kasia Hammer for technical support and help with the animal husbandry. We also want to thank the anonymous reviewers.

Funding

This work was supported by the Jewett Foundation for the development of SeaBase. JS and AHLF were supported by institutional funds from the Marine Biological Laboratory.

References

- Adamska M, Degnan BM, Green K, Zwafink C. 2011. What sponges can tell us about the evolution of developmental processes. *Zoology* 114:1–10.
- Adamska M, Larroux C, Adamski M, Green K, Lovas E, Koop D, Richards GS, Zwafink C, Degnan BM. 2010. Structure and expression of conserved Wnt pathway components in the demosponge *Amphimedon queenslandica*. *Evol Dev* 12:494–518.
- Altenburger A, Wanninger A. 2009. Comparative larval myogenesis and adult myoanatomy of the rhynchonelliform (articulate) brachiopods *Argyrotheca cordata*, *A. cistellula*, and *Terebratalia transversa*. *Front Zool* 6:1–14.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–10.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–402.
- Angerer LM, Angerer RC. 2003. 4 Patterning the sea urchin embryo: Gene regulatory networks, signaling pathways, and cellular interactions. *Curr Top Dev Biol* 53:159–98.
- Arendt D, Tessmar K, De Campos-Baptista M-I, Dorresteijn A, Wittbrodt J. 2002. Development of pigment-cup eyes in the polychaete *Platynereis dumerilii* and evolutionary conservation of larval eyes. *Development* 129:1143–54.
- Ashburner M, Drysdale R. 1994. FlyBase—the *Drosophila* genetic database. *Development* 120:2077–9.
- Austad SN. 2009. Is there a role for new invertebrate models for aging research? *J Gerontol A Biol* 64:192–4.
- Barreiro-Iglesias A, Shifman MI. 2012. Use of fluorochrome-labeled inhibitors of caspases to detect neuronal apoptosis in the whole-mounted lamprey brain after spinal cord injury. *Enzyme Res* 2012:835731.
- Bassaglia Y, Buresi A, Franko D, Andouche A, Baratte S, Bonnaud L. 2013. *Sepia officinalis*: A new biological model for eco-evo-devo studies. *J Exp Marine Biol Ecol* 447:4–13.
- Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. In: International AAAI conference on weblogs and social media: San Jose, California 361–2.
- Baxevanis AD. 2003. Using genomic databases for sequence-based biological discovery. *Mol Med* 9:185.
- Bely AE. 2006. Distribution of segment regeneration ability in the Annelida. *Integr Comp Biol* 46: 508–18.
- Bertrand S, Escriva H. 2011. Evolutionary crossroads in developmental biology: amphioxus. *Development* 138:4819–30.
- Bierne N, David P, Boudry P, Bonhomme F. 2002. Assortative fertilization and selection at larval stage in the mussels *Mytilus edulis* and *M. galloprovincialis*. *Evolution* 56:292–8.
- Bilofsky HS, Burks C, Fickett JW, Goad WB, Lewitter FI, Rindone WP, Swindell CD, Tung C-S. 1986. The GenBank genetic sequence databank. *Nucleic Acids Res* 14:1–4.
- Bilofsky HS, Christian B. 1988. The GenBank® genetic sequence data bank. *Nucleic Acids Res* 16:1861–3.
- Blanchard CE. 1986. Early development of the thorax and the nervous system of the brine shrimp *Artemia*. Dissertation, University of Leicester. 288 pp.
- Bloom OE, Morgan JR. 2011. Membrane trafficking events underlying axon repair, growth, and regeneration. *Mol Cell Neurosci* 48:339–48.
- Bodnar A. 2009. Marine invertebrates as models for aging research. *Exp Gerontol* 44:477–84.
- Bosch TC. 2009. Hydra and the evolution of stem cells. *Bioessays* 31:478–86.
- Boyer B, Henry J, Martindale M. 1996. Modified spiral cleavage: the duet cleavage pattern and early blastomer fates in the Acoel Turbellarian *Neochildia fusca*. *Biol Bull* 191:285.
- Brockes JP. 1998. Regeneration and cancer. *BBA-Rev Cancer* 1377:M1–11.
- Busch DJ, Morgan JR. 2012. Synuclein accumulation is associated with cell-specific neuronal death after spinal cord injury. *J Comp Neurol* 520:1751–71.
- Butler D. 2000. Ensembl gets a Wellcome boost. *Nature* 406:333–3.
- Carneveli M. 2006. Regeneration in Echinoderms: repair, regrowth, cloning. *Invertebrate Survival J* 3:64–76.
- Chiao C-C, Wickiser JK, Allen JJ, Genter B, Hanlon RT. 2011. Hyperspectral imaging of cuttlefish camouflage indicates good color match in the eyes of fish predators. *Proc Natl Acad Sci* 108:9148–53.
- Consortium F. 1994. FlyBase—the *Drosophila* database. *Nucleic Acids Res* 22: 3456–8.
- Copf T, Rabet N, Celniker SE, Averof M. 2003. Posterior patterning genes and the identification of a unique body region in the brine shrimp *Artemia franciscana*. *Development* 130:5915–27.
- Damen P. 1996. Spatial and temporal coincidence of induction processes and gap-junctional communication in *Patella vulgata* (Mollusca, Gastropoda). *Dev Genes Evol* 205:401–9.
- Damen P, Dictus WJAG. 1996. Organiser role of the stem cell of the mesoderm in prototroch patterning in *Patella vulgata* (Mollusca, Gastropoda). *Mech Dev* 56:41–60.
- Darling J, Reitzel A, Burton P, Mazza M, Ryan J, Sullivan J, Finnerty J. 2005. Rising starlet: the starlet sea anemone, *Nematostella vectensis*. *Bioessays* 27:211–21.
- Davidson EH. 2010. The regulatory genome: gene regulatory networks in development and evolution. Burlington: Academic Press.

- Davidson EH, Cameron RA, Ransick A. 1998. Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. *Development* 125:3269–90.
- Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh C-H, Minokawa T, Amore G, Hinman V, Arenas-Mena C. 2002. A genomic regulatory network for development. *Science* 295:1669–78.
- Dickinson AJG, Croll RP. 2003. Development of the larval nervous system of the gastropod *Ilyanassa obsoleta*. *J Comp Neurol* 466:197–218.
- Dictus WJAG, Damen P. 1997. Cell-lineage and clonal-contribution map of the trochophore larva of *Patella vulgata* (Mollusca). *Mech Dev* 62:312–26.
- Dill KK, Thamm K, Seaver EC. 2007. Characterization of twist and snail gene expression during mesoderm and nervous system development in the polychaete annelid *Capitella* sp. I. *Dev Genes Evol* 217:435–47.
- Dray N, Tessmar-Raible K, Le Gouar M, Vibert L, Christodoulou F, Schipany K, Guillou A, Zantke J, Snyman H, Béhague J, et al. 2010. Hedgehog signaling regulates segment formation in the annelid *Platynereis*. *Science* 329:339–42.
- Dunn CW, Hejnal A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD, et al. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452:745–9.
- Dyachuk V, Odintsova N. 2009. Development of the larval muscle system in the mussel *Mytilus trossulus* (Mollusca, Bivalvia). *Dev Growth Differ* 51:69–79.
- Eklblom R, Galindo J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107:1–15.
- Ereskovsky AV, Borchiellini C, Gazave E, Ivanisevic J, Lapébie P, Perez T, Renard E, Vacelet J. 2009. The Homoscleromorph sponge *Oscarellalobularis*, a promising sponge model in evolutionary and developmental biology. *BioEssays* 31:89–97.
- Estabrooks SL. 2007. The possible role of telomeres in the short life span of the bay scallop, *Argopecten irradians irradians* (Lamarck 1819). *J Shellfish Res* 26:307–13.
- Fernández-Suárez XM, Rigden DJ, Galperin MY. 2014. The 2014 Nucleic Acids Research Database Issue and an updated NAR online Molecular Biology Database Collection. *Nucleic Acids Res* 42:D1–6.
- Fiorito G, Scotto P. 1992. Observational learning in *Octopus vulgaris*. *Science* 256:545–7.
- Fischer AH, Henrich T. 2010. The normal development of *Platynereis dumerilii* (Nereididae, Annelida). *Front Zool* 7:31.
- Fischer AH, Pang K, Henry JQ, Martindale MQ. 2014. A cleavage clock regulates features of lineage-specific differentiation in the development of a basal branching metazoan, the ctenophore *Mnemiopsis leidyi*. *EvoDevo* 5:4.
- Fischer AH, Tulin S, Fredman D, Smith J. 2013. Employing BAC-reporter constructs in the sea anemone *Nematostella vectensis*. *Integr Comp Biol* 53:832–46.
- Frank U, Leitz T, Müller WA. 2001. The hydroid *Hydractinia*: a versatile, informative cnidarian representative. *Bioessays* 23:963–71.
- Freeman G. 1993. Regional specification during embryogenesis in the articulate brachiopod *Terebratalia*. *Dev Biol* 160:196–213.
- Freeman G. 2000. Regional specification during embryogenesis in the craniiform brachiopod *Crania anomala*. *Dev Biol* 227:219–38.
- Freeman G, Martindale M. 2002. The origin of mesoderm in phoronids. *Dev Biol* 252:301–11.
- Friedrich S, Wanninger A, Brückner M, Haszprunar G. 2002. Neurogenesis in the mossy chiton, *Mopalia muscosa* (Gould) (Polyplacophora): evidence against molluscan metamerism. *J Morphol* 253:109–17.
- Galliot B, Schmid V. 2002. Cnidarians as a model system for understanding evolution and regeneration. *Int J Dev Biol* 46:39–48.
- García Arrarás JE, Díaz Miranda L, Torres II, File S, Jiménez LB, Rivera Bermudez K, Arroyo EJ, Cruz W. 1999. Regeneration of the enteric nervous system in the sea cucumber *Holothuria glaberrima*. *J Comp Neurol* 406:461–75.
- García Arrarás JE, Estrada Rodgers L, Santiago R, Torres II, Díaz Miranda L, Torres Avillán I. 1998. Cellular mechanisms of intestine regeneration in the sea cucumber, *Holothuria glaberrima* Selenka (Holothuroidea: Echinodermata). *J Exp Zool* 281:288–304.
- Gauthier M, Degnan BM. 2008. The transcription factor NF-kappaB in the demosponge *Amphimedon queenslandica*: insights on the evolutionary origin of the Rel homology domain. *Dev Genes Evol* 218:23–32.
- Genikhovich G, Technau U. 2009a. Induction of spawning in the starlet sea anemone *Nematostella vectensis*, in vitro fertilization of gametes, and dejellying of zygotes. *Cold Spring Harb Protoc* 2009:pdb.prot5281.
- Genikhovich G, Technau U. 2009b. The starlet sea anemone *Nematostella vectensis*: an anthozoan model organism for studies in comparative genomics and functional evolutionary developmental biology. *Cold Spring Harb Protoc* 2009:pdb.emo129.
- Gerhart J, Lowe C, Kirschner M. 2005. Hemichordates and the origin of chordates. *Curr Opin Genet Dev* 15:461–7.
- Gilmore TD, Tarrant AM, Finnerty JR. 2013. A report from the second *Nematostella vectensis* research conference. *Dev Genes Evol* 223:207–11.
- Gonzales EE, van der Zee M, Dictus WJ, van den Biggelaar J. 2007. Brefeldin A or monensin inhibits the 3D organizer in gastropod, polyplacophoran, and scaphopod molluscs. *Dev Genes Evol* 217:105–18.
- Goulding MQ. 2009. Cell lineage of the *Ilyanassa* embryo: evolutionary acceleration of regional differentiation during early development. *PLoS One* 4:e5506.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29:644–52.
- Grande C, Patel NH. 2009. Nodal signalling is involved in left–right asymmetry in snails. *Nature* 457:1007–11.
- Granger CW. 1988. Causality, cointegration, and control. *J Econ Dyn Control* 12:551–9.
- Grimaldi A, Tettamanti G, Acquati F, Bossi E, Guidali ML, Banfi S, Monti L, Valvassori R, De Eguileor M. 2008. A hedgehog homolog is involved in muscle formation and organization of *Sepia officinalis* (mollusca) mantle. *Dev Dyn* 237:659–71.

- Gruhl A. 2010. Ultrastructure of mesoderm formation and development in *Membranipora membranacea* (Bryozoa: Gymnolaemata). *Zoomorphology* 129:45–60.
- Guarente L, Kenyon C. 2000. Genetic pathways that regulate ageing in model organisms. *Nature* 408:255–62.
- Harzsch S, Glötzner J. 2002. An immunohistochemical study of structure and development of the nervous system in the brine shrimp *Artemia salina* Linnaeus, 1758 (Branchiopoda, Anostraca) with remarks on the evolution of the arthropod brain. *Arthropod Struct Dev* 30:251–70.
- Haszprunar G, Friedrich S, Wanninger A, Ruthensteiner B. 2002. Fine structure and immunocytochemistry of a new chemosensory system in the Chiton larva (Mollusca: Polyplacophora). *J Morphol* 251:210–8.
- Hejnal A, Martindale MQ. 2008. Acoel development indicates the independent evolution of the bilaterian mouth and anus. *Nature* 456:382–6.
- Hejnal A, Martindale MQ. 2009. Coordinated spatial and temporal expression of Hox genes during embryogenesis in the acoel *Convolutriloba longifissura*. *BMC Biol* 7:65.
- Hejnal A, Martindale MQ, Henry JQ. 2007. High-resolution fate map of the snail *Crepidula fornicata*: The origins of ciliary bands, nervous system, and muscular elements. *Dev Biol* 305:63–76.
- Helmkamp M, Bruchhaus I, Hausdorf B. 2008. Phylogenomic analyses of lophophorates (brachiopods, phoronids and bryozoans) confirm the Lophotrochozoa concept. *P Roy Soc B Bio* 275:1927.
- Henry J, Martindale M. 1998. Conservation of the spiralian developmental program: cell lineage of the nemertean, *Cerebratulus lacteus*. *Dev Biol* 201:253–69.
- Henry J, Martindale M, Boyer B. 2000. The unique developmental program of the acoel flatworm, *Neochildia fusca*. *Dev Biol* 220:285–95.
- Henry JJ, Collin R, Perry KJ. 2010. The slipper snail, *Crepidula*: an emerging lophotrochozoan model system. *Biol Bull* 218:211–29.
- Henry JJ, Perry KJ. 2008. MAPK activation and the specification of the D quadrant in the gastropod mollusc, *Crepidula fornicata*. *Dev Biol* 313:181–95.
- Hernroth B, Farahani F, Brunborg G, Dupont S, Dejmek A, Nilsson Sköld H. 2010. Possibility of mixed progenitor cells in sea star arm regeneration. *J Exp Zool B* 314:457–68.
- Hinman VF, Davidson EH. 2007. Evolutionary plasticity of developmental gene regulatory network architecture. *Proc Natl Acad Sci* 104:19404–9.
- Hinman VF, Nguyen A, Davidson EH. 2007. Caught in the evolutionary act: precise cis-regulatory basis of difference in the organization of gene networks of sea stars and sea urchins. *Dev Biol* 312:584.
- Holland P, Holland LZ, Williams NA, Holland ND. 1992. An amphioxus homeobox gene: sequence conservation, spatial expression during development and insights into vertebrate evolution. *Development* 116:653–61.
- Holland PW, Garcia-Fernández J, Williams NA, Sidow A. 1994. Gene duplications and the origins of vertebrate development. *Development* 1994:125–33.
- Houlston E, Momose T, Manuel M. 2010. *Clytia hemisphaerica*: a jellyfish cousin joins the laboratory. *Trends Genet* 26:159–67.
- Hubbard T, Birney E. 2000. Open annotation offers a democratic solution to genome sequencing. *Nature* 403:825.
- Jondelius U, Larsson K, Raikova O. 2004. Cleavage in *Nemertoderma westbladi* (Nemertodermatida) and its phylogenetic significance. *Zoomorphology* 123:221–5.
- Kennedy B. 2008. The genetics of ageing: insight from genome-wide approaches in invertebrate model organisms. *J Int Med* 263:142–52.
- Kumano G, Nishida H. 2007. Ascidian embryonic development: an emerging model system for the study of cell fate specification in chordates. *Dev Dyn* 236:1732–47.
- Lambert D. 2009. Patterning the Spiralian Embryo: Insights from *Ilyanassa*. *Curr Top Dev Biol* 86:107–33.
- Lambert JD, Nagy LM. 2001. MAPK signaling by the D quadrant embryonic organizer of the mollusc *Ilyanassa obsoleta*. *Development* 128:45–56.
- 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.
- Lapraz F, Rawlinson KA, Girstmair J, Tomiczek B, Berger J, Jékely G, Telford MJ, Egger B. 2013. Put a tiger in your tank: the polyclad flatworm *Maritigrella crozieri* as a proposed model for evo-devo. *EvoDevo* 4:1–16.
- Lartillot N, Lespinet O, Vervoort M, Adoutte A. 2002. Expression pattern of Brachyury in the mollusc *Patella vulgata* suggests a conserved role in the establishment of the AP axis. *Development* 129:1411–21.
- Lau BY, Fogerson SM, Walsh RB, Morgan JR. 2013. Cyclic AMP promotes axon regeneration, lesion repair and neuronal survival in lampreys after spinal cord injury. *Exp Neurol* 250:31–42.
- Lawrence JM. 1991. Arm loss and regeneration in Asterozoa (Echinodermata). *Echinoderm Res* 1992:39–52.
- Lehnert EM, Burriesci MS, Pringle JR. 2012. Developing the anemone *Aiptasia* as a tractable model for cnidarian-dinoflagellate symbiosis: the transcriptome of aposymbiotic *A. pallida*. *BMC Genomics* 13:271.
- Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.
- Lowe CJ, Terasaki M, Wu M, Freeman RM Jr, Runft L, Kwan K, Haigo S, Aronowicz J, Lander E, Gruber C. 2006. Dorsoroventral patterning in hemichordates: insights into early chordate evolution. *PLoS Biol* 4:e291.
- Lowe CJ, Wu M, Salic A, Evans L, Lander E, Stange-Thomann N, Gruber CE, Gerhart J, Kirschner M. 2003. Anteroposterior patterning in hemichordates and the origins of the chordate nervous system. *Cell* 113:853–65.
- Lurie DI, Selzer ME. 1991. Axonal regeneration in the adult lamprey spinal cord. *J Comp Neurol* 306:409–16.
- Lyons DC, Perry KJ, Lesoway MP, Henry JQ. 2012. Cleavage pattern and fate map of the mesentoblast, 4d, in the gastropod *Crepidula*: a hallmark of spiralian development. *EvoDevo* 3:21.
- Martindale MQ, Hejnal A. 2009. A developmental perspective: changes in the position of the blastopore during bilaterian evolution. *Dev Cell* 17:162–74.
- Martínez DE. 1998. Mortality patterns suggest lack of senescence in hydra. *Exp Gerontol* 33:217–25.

- Mashanov VS, Zueva OR, Garcia-Ararras JE. 2012. Expression of *Wnt9*, *TCTP*, and *Bmp1/Tll* in sea cucumber visceral regeneration. *Gene Express Pattern* 12:24–35.
- Masuda-Nakagawa LM, Gröer H, Aerne BL, Schmid V. 2000. The HOX-like gene *Cnox2-Pc* is expressed at the anterior region in all life cycle stages of the jellyfish *Podocoryne carnea*. *Dev Genes Evol* 210:151–6.
- Mäthger LM, Denton EJ, Marshall NJ, Hanlon RT. 2009. Mechanisms and behavioural functions of structural coloration in cephalopods. *J Roy Soc Interface* 6: S149–63.
- McClay DR. 2011. Evolutionary crossroads in developmental biology: sea urchins. *Development* 138:2639–48.
- McDougall C, Hui JHL, Monteiro A, Takahashi T, Ferrier DEK. 2008. Annelids in evolutionary developmental biology and comparative genomics. *Parasite* 15:321–8.
- McGettigan PA. 2013. Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol* 17:4–11.
- Meyer NP, Seaver EC. 2009. Neurogenesis in an annelid: characterization of brain neural precursors in the polychaete *Capitella* sp. I. *Dev Biol* 335:237–52.
- Meyer N, Seaver E. 2010. Cell Lineage and Fate Map of the primary somatoblast of the polychaete annelid *Capitella teleta*. *Integr Comp Biol* 50:756–67.
- Miller DJ, Ball EE. 2000. The coral Acropora: what it can contribute to our knowledge of metazoan evolution and the evolution of developmental processes. *Bioessays* 22:291–6.
- Miller G. 2009. On the origin of the nervous system. *Science* 325:24–6.
- Mladenov PV, Carson SF, Walker CW. 1986. Reproductive ecology of an obligately fissiparous population of the sea star *Stephanasterias albula* Stimpson. *J Exp Marine Biol Ecol* 96:155–75.
- Moret F, Guillard JC, Coudouel S, Rochette L, Vernier P. 2004. Distribution of tyrosine hydroxylase, dopamine, and serotonin in the central nervous system of amphioxus (*Branchiostoma lanceolatum*): implications for the evolution of catecholamine systems in vertebrates. *J Comp Neurol* 468:135–50.
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5:621–8.
- Müller GB. 2007. Evo–devo: extending the evolutionary synthesis. *Nat Rev Genet* 8:943–9.
- Nakano H, Murabe N, Amemiya S, Nakajima Y. 2006. Nervous system development of the sea cucumber *Stichopus japonicus*. *Dev Biol* 292:205–12.
- Nielsen C. 1991. The development of the brachiopod *Crania (Neocrania) anomala* (OF Müller) and its phylogenetic significance. *Acta Zoologica* 72:7–28.
- Okamoto M. 1997. Simultaneous demonstration of lens regeneration from dorsal iris and tumour production from ventral iris in the same newt eye after carcinogen administration. *Differentiation* 61:285–92.
- Oulion S, Bertrand S, Belgacem MR, Le Petillon Y, Escriva H. 2012. Sequencing and analysis of the Mediterranean amphioxus (*Branchiostoma lanceolatum*) transcriptome. *PLoS One* 7:e36554.
- Oviedo NJ, Beane WS. 2009. Regeneration: The origin of cancer or a possible cure? *Semin Cell Dev Biol* 20:557–64.
- Ozsolak F, Milos PM. 2011. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet* 12:87–98.
- Pang K, Martindale MQ. 2008. Comb jellies (Ctenophora): a model for basal metazoan evolution and development. *Cold Spring Harb Protoc* 2008:pdb.emo106.
- Peterson KJ, Cameron RA, Tagawa K, Satoh N, Davidson EH. 1999. A comparative molecular approach to mesodermal patterning in basal deuterostomes: the expression pattern of Brachyury in the enteropneust hemichordate *Ptychodera flava*. *Development* 126:85.
- Piatigorsky J, Kozmik Z. 2004. Cubozoan jellyfish: an Evo/Devo model for eyes and other sensory systems. *Int J Dev Biol* 48:719–29.
- Raff RA. 1992. Direct developing sea urchins and the evolutionary reorganization of early development. *BioEssays* 14:211–8.
- Raff RA. 2000. Evo-devo: the evolution of a new discipline. *Nat Rev Genet* 1:74–9.
- Raible F, Tessmar-Raible K, Osoegawa K, Wincker P, Jubin C, Balavoine G, Ferrier D, Benes V, De Jong P, Weissenbach J, et al. 2005. Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*. *Science* 310:1325–6.
- Rawlinson KA. 2010. Embryonic and post-embryonic development of the polyclad flatworm *Maritigrella crozieri*; implications for the evolution of spiralian life history traits. *Front Zool* 7:12.
- Reitzel AM, Daly M, Sullivan JC, Finnerty JR. 2009. Comparative anatomy and histology of developmental and parasitic stages in the life cycle of the lined sea anemone *Edwardsiella lineata*. *J Parasitol* 95:100–12.
- Santagata S. 2004. Larval development of *Phoronis pallida* (Phoronida): implications for morphological convergence and divergence among larval body plans. *J Morphol* 259:347–58.
- Satoh N. 2003. The ascidian tadpole larva: comparative molecular development and genomics. *Nat Rev Genet* 4:285–95.
- Schäffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV, Altschul SF. 2001. Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res* 29:2994–3005.
- Schneider SQ, Bowerman B. 2007. b-Catenin asymmetries after all animal/vegetal-oriented cell divisions in *Platynereis dumerilii* embryos mediate binary cell-fate specification. *Dev Cell* 13:73–86.
- Selzer ME. 1978. Mechanisms of functional recovery and regeneration after spinal cord transection in larval sea lamprey. *J Physiol* 277:395–408.
- Semmler H, Bailly X, Wanninger A. 2008. Myogenesis in the basal bilaterian *Symsagittifera roscoffensis* (Acoela). *Front Zool* 5:14.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–504.
- Sikes JM, Bely AE. 2008. Radical modification of the A–P axis and the evolution of asexual reproduction in *Convolutriloba acoels*. *Evol Dev* 10:619–31.

- Smoot ME, Ono K, Ruschinski J, Wang P-L, Ideker T. 2011. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27:431–2.
- Sommer RJ. 2000. Evolution of nematode development. *Curr Opin Genet Dev* 10:443–8.
- Stach T, Winter J, Bouquet J-M, Chourrout D, Schnabel R. 2008. Embryology of a planktonic tunicate reveals traces of sessility. *Proc Natl Acad Sci USA* 105:7229–34.
- Steele R. 2002. Developmental signaling in Hydra: what does it take to build a “simple” animal? *Dev Biol* 248:199.
- Stefanik DJ, Friedman LE, Finnerty JR. 2013. Collecting, rearing, spawning and inducing regeneration of the starlet sea anemone, *Nematostella vectensis*. *Nat Protoc* 8:916–23.
- Stefanik DJ, Lubinski TJ, Granger BR, Byrd AL, Reitzel AM, DeFilippo L, Lorenc A, Finnerty JR. 2014. Production of a reference transcriptome and transcriptomic database (EdwardsiellaBase) for the lined sea anemone, *Edwardsiella lineata*, a parasitic cnidarian. *BMC Genomics* 15:71.
- Stein L, Sternberg P, Durbin R, Thierry-Mieg J, Spieth J. 2001. WormBase: network access to the genome and biology of *Caenorhabditis elegans*. *Nucleic Acids Res* 29:82–86.
- Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, Scherf M, Seifert M, Borodina T, Soldatov A, Parkhomchuk D. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 321:956–60.
- Sun L, Yang H, Chen M, Ma D, Lin C. 2013. RNA-Seq reveals dynamic changes of gene expression in key stages of intestine regeneration in the sea cucumber *Apostichopus japonicus*. *PLoS One* 8:e69441.
- Sunagawa S, Wilson EC, Thaler M, Smith ML, Caruso C, Pringle JR, Weis VM, Medina M, Schwarz JA. 2009. Generation and analysis of transcriptomic resources for a model system on the rise: the sea anemone *Aiptasia pallida* and its dinoflagellate endosymbiont. *BMC Genomics* 10:258.
- Sweet H. 1998. Specification of first quartet micromeres in *Ilyanassa* involves inherited factors and position with respect to the inducing D macromere. *Development* 125:4033.
- Tagawa K, Nishino A, Humphreys T, Satoh N. 1998. The spawning and early development of the Hawaiian acorn worm (hemichordate), *Ptychodera flava*. *Zool Sci* 15:85–91.
- Tagawa K, Satoh N, Humphreys T. 2001. Molecular studies of hemichordate development: a key to understanding the evolution of bilateral animals and chordates. *Evol Dev* 3:443–54.
- Tessmar-Raible K, Arendt D. 2003. Emerging systems: between vertebrates and arthropods, the Lophotrochozoa. *Curr Opin Genet Dev* 13:331–40.
- Tsutsumi H, Kikuchi T. 1984. Study of the life history of *Capitella capitata* (Polychaeta: Capitellidae) in Amakusa, South Japan including a comparison with other geographical regions. *Marine Biol* 80:315–21.
- Tulin S, Aguiar D, Istrail S, Smith J. 2013. A quantitative reference transcriptome for *Nematostella vectensis* early embryonic development: a pipeline for de novo assembly in emerging model systems. *EvoDevo* 4:16–16.
- Vopalensky P, Pergner J, Liegertova M, Benito-Gutierrez E, Arendt D, Kozmik Z. 2012. Molecular analysis of the amphioxus frontal eye unravels the evolutionary origin of the retina and pigment cells of the vertebrate eye. *Proc Natl Acad Sci* 109:15383–8.
- Voronov DA, Panchin YV. 1998. Cell lineage in marine nematode *Enoplus brevis*. *Development* 125:143–50.
- Waddington CH. 1935. Cancer and the theory of organisers. *Nature* 135:606–8.
- Wanamaker ADJ, Heinemeier J, Scourse JD, Richardson CA, Butler PG, Eiriksson J, Knudsen KL. 2008. Very long-lived mollusks confirm 17th century AD tephra-based radiocarbon reservoir ages for North Icelandic shelf waters. *Radiocarbon* 50:399–412.
- Wanninger A, Haszprunar G. 2002. Muscle development in *Antalis entalis* (Mollusca, Scaphopoda) and its significance for Scaphopod relationships. *J Morphol* 254:53–64.
- Wennberg SA, Janssen R, Budd GE. 2008. Early embryonic development of the priapulid worm *Priapulid caudatus*. *Evol Dev* 10:326–38.
- Wildt M, Harzsch S. 2002. A new look at an old visual system: structure and development of the compound eyes and optic ganglia of the brine shrimp *Artemia salina* Linnaeus, 1758 (Branchiopoda, anostraca). *J Neurobiol* 52:117–32.
- Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bähler J. 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453:1239–43.
- Williamson R, Chrachri A. 2004. Cephalopod neural networks. *Neurosignals* 13:87–98.
- Wolenski FS, Layden MJ, Martindale MQ, Gilmore TD, Finnerty JR. 2013. Characterizing the spatiotemporal expression of RNAs and proteins in the starlet sea anemone, *Nematostella vectensis*. *Nat Protoc* 8:900–15.
- Wray GA, Raff RA. 1990. Novel origins of lineage founder cells in the direct-developing sea urchin *Heliocidaris erythrogramma*. *Dev Biol* 141:41–54.
- Yankura KA, Martik ML, Jennings CK, Hinman VF. 2010. Uncoupling of complex regulatory patterning during evolution of larval development in echinoderms. *BMC Biol* 8:143.
- Yuan D, Nakanishi N, Jacobs DK, Hartenstein V. 2008. Embryonic development and metamorphosis of the scyphozoan *Aurelia*. *Dev Genes Evol* 218:525–39.
- Zhang C, Hanspers K, Kuchinsky A, Salomonis N, Xu D, Pico AR. 2012. Mosaic: making biological sense of complex networks. *Bioinformatics* 28:1943–44.
- Zhang G, Pizarro IV, Swain GP, Kang SH, Selzer ME. 2013. Neurogenesis in the lamprey CNS following spinal cord transection. *J Comp Neurol* 522:1316–32.