January 2012

Comparative Developmental Transcriptomics of Echinoderms

Roy Vaughn
University of South Florida, revaughn@mail.usf.edu

Comparative Developmental Transcriptomics of Echinoderms

by

Roy Vaughn

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Cell Biology, Microbiology and Molecular Biology
College of Arts and Sciences
University of South Florida

Major Professor, James R. Garey, Ph.D.
Brian T. Livingston, Ph.D.
Kathleen Scott, Ph.D.
Kristina Schmidt, Ph.D.

Date of Approval:
June 7, 2012

Keywords: Brittle Star, Sea Urchin, Gene Regulatory Networks,
Transcription Factors, Gastrula

Table of Contents

List of Tables

List of Figures

Abstract

The gastrula stage represents the point in development at which the three primary germ layers diverge. At this point the gene regulatory networks that specify the germ layers are established and the genes that define the differentiated states of the tissues have begun to be activated. These networks have been well characterized in sea urchins, but not in other echinoderms. Embryos of the brittle star *Ophiocoma wendtii* share a number of developmental features with sea urchin embryos, including the ingression of mesenchyme cells that give rise to an embryonic skeleton. Notable differences are that no micromeres are formed during cleavage divisions and no pigment cells are formed during development to the pluteus larva stage. More subtle changes in timing of developmental events also occur. To explore the molecular basis for the similarities and differences between these two echinoderms, the gastrula transcriptome of *Ophiocoma wendtii* was sequenced and characterized.

I identified brittle star transcripts that correspond to 3385 genes in existing databases, including 1863 genes shared with the sea urchin *Strongylocentrotus purpuratus* gastrula transcriptome. I have characterized the functional classes of genes present in the transcriptome and compared them to those found in sea urchin. I then examined which members of the germ-layer specific gene regulatory networks (GRNs) of *S. purpuratus* are expressed in the *O. wendtii* gastrula. The results indicate that there is a shared “genetic toolkit” central to the

echinoderm gastrula, a key stage in embryonic development, though there are also differences that reflect changes in developmental processes.

The brittle star expresses genes representing all functional classes at the gastrula stage. Brittle stars and sea urchins have comparable numbers of each class of genes, and share many of the genes expressed at gastrula. Examination of the brittle star genes whose sea urchin orthologs are utilized in germ layer specification reveals a relatively higher level of conservation of key regulatory components compared to the overall transcriptome. I also identify genes that were either lost or whose temporal expression has diverged from that of sea urchins. Overall, the data suggest that embryonic skeleton formation in sea urchins and brittle stars represents convergent evolution by independent cooptation of a shared pathway utilized in adult skeleton formation.

Transcription factors are of central importance to both development and evolution. Patterns of their expression and interactions form the gene regulatory networks which control the building of the embryonic body. Alterations in these patterns can result in the construction of altered bodies. To help increase understanding of this process, I compared the transcription factor mRNAs present in early gastrula-stage embryos of the brittle star *Ophiocoma wendtii* to those found in two species of sea urchins and a starfish. Brittle star homologs were found for one third of the transcription factors in the sea urchin genome and half of those that are expressed at equivalent developmental stages in sea urchins and starfish. Overall, the patterns of transcription factors found and not found in brittle star resemble those of other echinoderms, with the differences

largely consistent with morphological differences. This study provides further evidence for the existence of deeply conserved developmental genetic processes, with various elements shared among echinoderms, deuterostomes, and metazoans.

Chapter One:
Introduction

THE GENETICS OF DEVELOPMENT AND EVOLUTION

The process of turning a one-celled fertilized egg into a complete multicellular animal with an array of highly specialized tissues and organs requires a system for implementing specific genetic instructions at particular times and places. A developing organism must be able to build itself from and with the materials and tools at hand—its own genes and proteins—in a process of increasingly refined self-organization. This depends on the ability of the gene and protein players to talk to each other.

Transcription of genes to make proteins is controlled by other proteins—the transcription factors. Each particular transcription factor can bind to certain nucleotide sequences—promoters or *cis*-regulatory elements—upstream from the coding regions of target genes. This binding of specific transcription factors to DNA can either allow or inhibit assembly and activation of the full transcriptional apparatus, thus either promoting or repressing transcription of that gene. The ability of a given transcription factor's DNA-binding domain to attach to a given

ABBREVIATIONS
cDNA = complementary DNA (synthesized from mRNA templates)
EST = expressed sequence tag, **GRN** = gene regulatory network
mRNA = messenger RNA, **NSM** = non-skeletogenic mesenchyme
PCR = polymerase chain reaction, **PMC** = primary mesenchyme
RACE = rapid amplification of cDNA ends
SMC = secondary mesenchyme, **TF** = transcription factor

gene's promoter element depends on the exact amino acid and nucleotide sequences of these regions, and thus their potential for precise chemical interactions.

Following fertilization, transcription factors initially target genes mostly for other transcription factors and signaling pathways. As development proceeds, early differences in the locations and concentrations of these molecules establish gene regulatory networks (GRNs), regional patterns of gene expression which channel different parts of the embryo to mature into different structures.

Mutations which alter the binding specificity of transcription factors or the promoter elements of the genes they regulate can result in changes to the location and/or timing of expression of the targeted genes, and therefore to alterations in the structures and functions influenced by those genes. Such rewiring of developmental gene regulatory networks is a major mechanism linking genetic variation to natural selection, and is the focus of study for evolutionary developmental biology, or "evo-devo".

ECHINODERMS AS MODEL ORGANISMS IN DEVELOPMENTAL BIOLOGY

As basal deuterostomes, echinoderms occupy an important phylogenetic position between chordates and all other animal phyla. Sea urchins (Class Echinoidea) have been used as model organisms in developmental biology for more than a century. Many of the landmark discoveries of biology were made, at least in part, through the study of sea urchin embryos, including fertilization, egg cytoplasmic determinants, cell-signaling, differential gene expression, and the

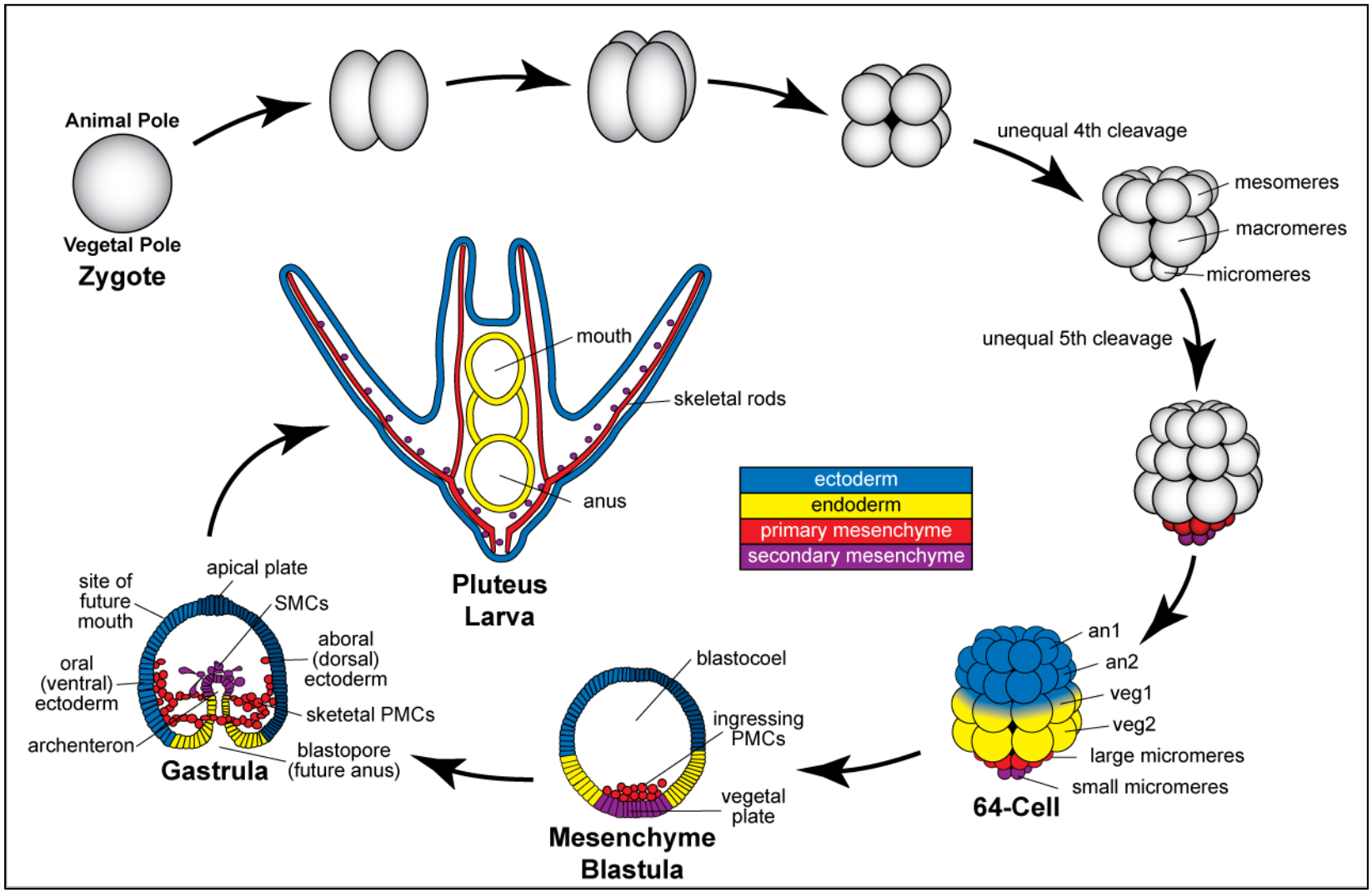

Figure 1.1: Development of Sea Urchin Embryos. (Modified from McClay 2011.)

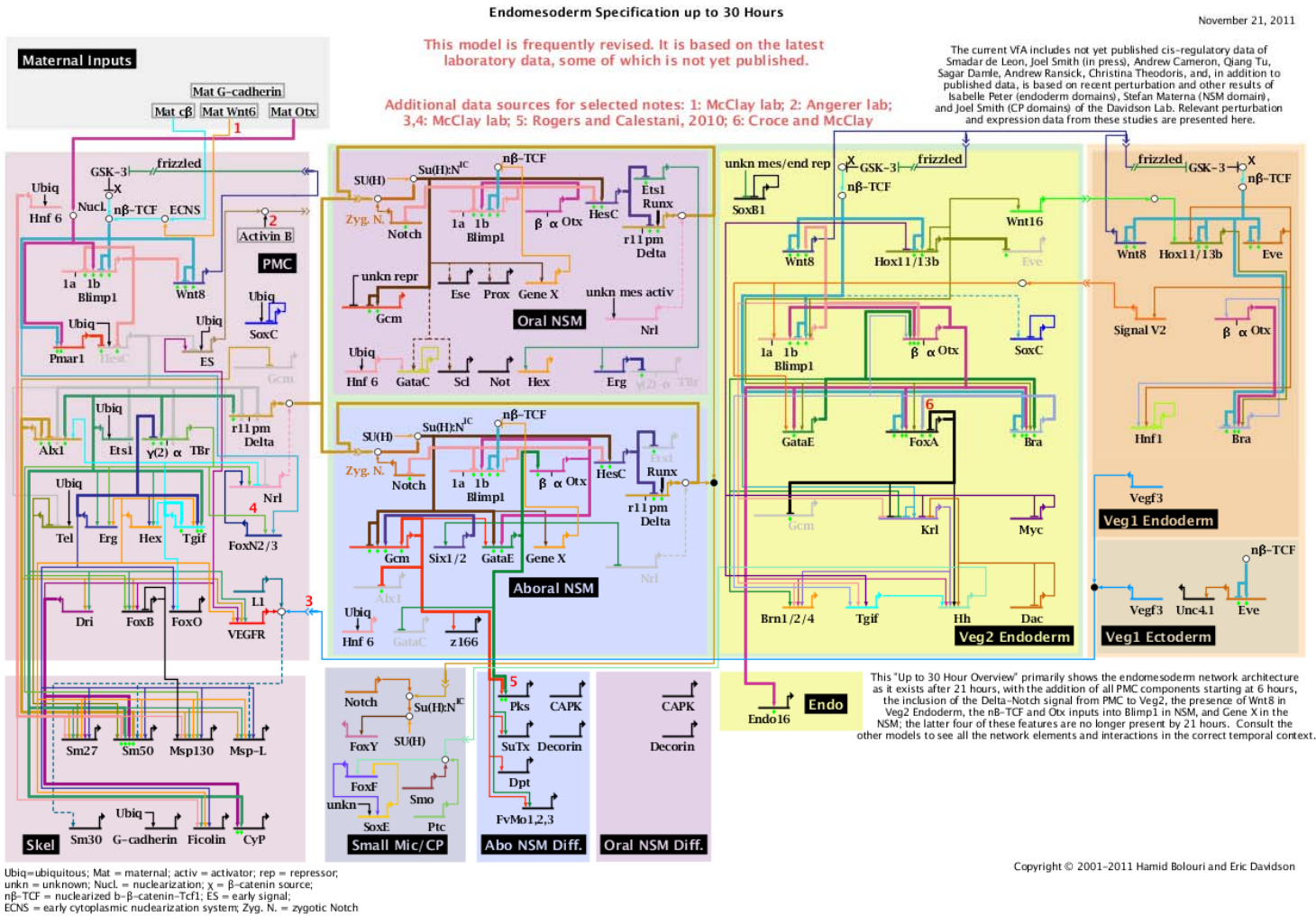

Figure 1.2: Endomesodermal Developmental Gene Regulatory Network for the Sea Urchin *Strongylocentrotus purpuratus.* The large blocks of color represent the different germ-layer regions of the early embryo. Early specification events are toward the top of the diagram, with progressively more refined specification in the middle, and the beginnings of tissue differentiation toward the bottom at around the time of gastrulation. The horizontal line above each gene name represents the promoter region for that gene. Arrow inputs to a promoter indicate transcription factors which promote transcription of the gene. Barred inputs signify transcriptional repression. Arrows leaving the end of a gene's promoter line symbolize transcription and expression. [Figure generated with BioTapestry software.]

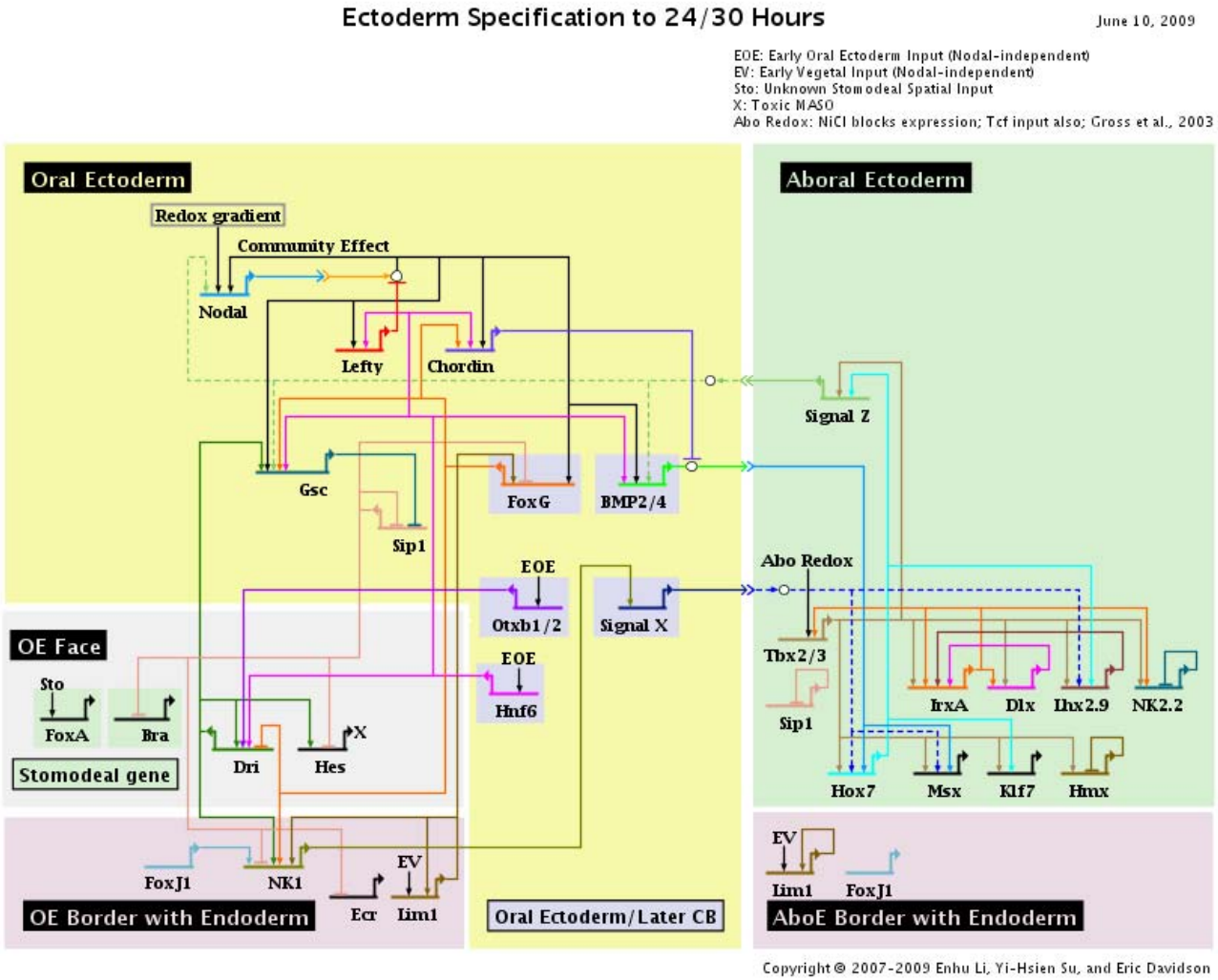

Figure 1.3: Ectodermal Developmental Gene Regulatory Network for the Sea Urchin *Strongylocentrotus purpuratus.* Symbolism is the same as for Figure 1.2 (page 4). [Figure generated with BioTapestry software – www.biotapestry.org]

mitotic spindle [reviewed in Pederson 2006, Briggs and Wessel 2006]. A number of urchin species have been widely used, including *Strongylocentrotus purpuratus, Lytechinus variegatus, Paracentrotus lividus, Heliocidaris erythrogramma,* and others. The key events of sea urchin embryonic development are illustrated in Figure 1.1.

The genome of *S. purpuratus* has been fully sequenced [Sea Urchin Genome Sequencing Consortium 2006, Cameron *et al* 2009], and its developmental gene regulatory network has been extensively studied in the stages from egg through gastrulation [Figures 1.2 & 1.3] [Davidson *et al* 2002a & 2002b, Oliveri and Davidson, 2004, Cameron *et al* 2009, Su 2009, Peter and Davidson, 2010]. Various techniques were used to uncover the structure of this network, including *in situ* hybridizations, quantitative PCR, and gain- and loss-of-function perturbation studies [Davidson *et al* 2002a & 2002b]. Similar studies have recently begun using the starfish *Patiria miniata* (formerly called *Asterina miniata*), and have produced many valuable insights into the ways that features of GRNs persist and/or change over the course of evolutionary time [Hinman *et al* 2003a, 2007a & 2007b; McCauley *et al* 2010]. The genetics of development in the remaining echinoderm classes have been studied very little, if at all. This project is the first to investigate these processes in the class Ophiuroidea, the brittle stars.

The precise phylogenetic position of brittle stars relative the other echinoderm classes is still unresolved due to conflicting molecular, morphological, and embryological evidence [Littlewood *et al* 1997, Harmon

2005]. Adult brittle stars somewhat resemble starfish (Class Asteroidea) in that they have 5 arms. However, the arms of brittle stars are long and very flexible—the name “ophiuroid” is from the Greek for “snake-like”—and are used for both locomotion and feeding. The arms of starfish are more like extensions of the central body, containing both the gonads and digestive caeca, and are generally held more-or-less stationary as the numerous tiny tube feet underneath are used for locomotion.

Brittle star development from egg to swimming larva is quite like that of sea urchins [Figure 1.1], with some notable differences. Brittle stars do not have unequal 4^{th} and 5^{th} cleavages, and thus do not produce micromeres. Nevertheless, mesenchyme cells apparently homologous to urchin micromeres ingress to form an embryonic skeleton very similar to that of urchins, giving the brittle star larva the same pluteus shape as the urchin larva, but lacking pigment cells. The larvae of the other echinoderm classes are very different in shape, and lack skeletons. Brittle stars are thus a promising group to explore just how similar or different their development gene regulatory network is to those of sea urchins and starfish. This project represents a first step toward investigating these questions,

RESEARCH DESIGN

Ophiocoma wendtii is a brittle star species living among the shallow water rocks and reefs of the Florida Keys. It was chosen for these studies because of

its easy availability and large size. The central body disk is often 2–2½ cm in diameter, with each arm 10–12 cm long.

A candidate gene PCR approach was initially tried. Degenerate PCR primers were designed based on the most conserved regions of transcription factor genes homologous between several echinoderms and other deuterostomes, including *S. purpuratus* and other urchin species, the starfish *Patiria miniata*, the hemichordates *Saccoglossus kowalevskii* and *Ptychodera flava,* the urochordate tunicate *Ciona intestinalis,* the cephalochodate *Branchiostoma floridae* ("Amphioxus"), and/or others as the available sequence data allowed. These primers were used to attempt to amplify the brittle star versions of these genes from *O. wendtii* cDNA from various embryonic stages. The resulting PCR products were then cloned, extended though RACE, and sequenced, with intentions for future experiments along the lines of those used to unravel the sea urchin GRN. This method led to effective cloning and sequencing of several brittle star genes. For several others genes, it was not successful, despite numerous attempts.

Newly available high-throughput pyrosequencing systems inspired a different approach. These technologies enable rapid, relatively inexpensive sequencing of a very large number of small DNA or cDNA fragments covering a large portion of an organism's genome or transcriptome. The results can then be analyzed for large-scale patterns and screened for genes of particular interest. The partial gene sequences can also greatly facilitate future full-length sequencing, expression and functional studies.

Messenger RNAs from early gastrula-stage embryos of the brittle star *Ophiocoma wendtii* were pyrosequenced using systems from 454 Life Sciences. The results have yielded important and intriguing additions to understanding the phylogenetic conservation and evolutionary remodeling of developmental gene regulatory networks.

Chapter Two:
Sequencing and Analysis of the Gastrula Transcriptome of the Brittle Star *Ophiocoma wendtii*

This chapter has been submitted to *EvoDevo* for publication under the same title, and is currently under revision.

BACKGROUND

Sea urchins (Class Echinoidea) have been used as model organisms in developmental biology for more than a century. Over the last two decades, intensive work has led to a fairly detailed understanding of the gene regulatory network (GRN) controlling the differentiation of the embryonic germ layers during development in the species *Strongylocentrotus purpuratus* [Davidson *et al* 2002a & 2002b, Oliveri and Davidson 2004, Su 2009, Peter and Davidson 2010, Cameron *et al* 2009]. An initial draft of the *S. purpuratus* genome was completed in 2006 [Sea Urchin Genome Sequencing Consortium 2006], and is now in its third revision [Cameron *et al* 2009]. Several expression databases for various embryonic stages have also been constructed, using expressed sequence tags (ESTs) [Poustka *et al* 1999, Lee *et al* 1999, Zhu *et al* 2001, Poustka *et al* 2003], microarrays [Wei *et al* 2006], and NanoString RNA counting [Materna *et al* 2010]. Here I begin to examine the conservation and divergence in the gene regulatory networks expressed at the gastrula stage in a member of a different echinoderm class, the Ophiuroidea. Our results indicate that there is a shared “genetic toolkit”

central to the echinoderm gastrula, a key stage in embryonic development, though there are differences that reflect changes in developmental processes.

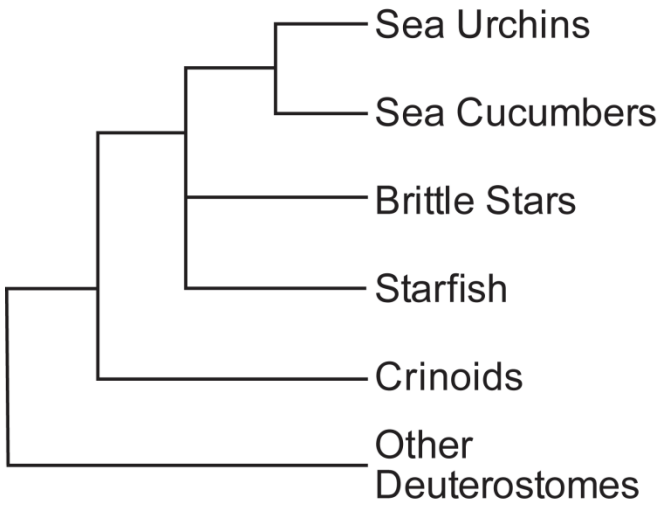

Figure 2.1: Phylogeny of Echinoderms
All evidence indicates that crinoids are the most basal. The other four groups all diverged within a very short geological timeframe around 500 million years ago. Urchins and sea cucumbers are generally considered to form a clade as the most derived. It remains unclear whether the brittle stars group more closely with this clade or with starfish, due to conflicts between molecular, morphological, and embryological evidence.

The echinoderms consist of five living classes: Asteroidea (starfish), Echinoidea (sea urchins and sand dollars), Ophiuroidea (brittle stars) Holothuroidea (sea cucumbers), and Crinoidea (sea lilies and feather stars). The crinoids appear first in the fossil record, and are clearly the most basal anatomically. The other four classes appear to have all diverged within a very short geological period around 500 million years ago [Paul and Smith 1984], and the exact phylogenetic relationship of the brittle stars to the other classes remains uncertain due to conflicts between molecular, morphological, and embryological evidence [Littlewood *et al* 1997, Harmon 2005] [Figure 2.1]. The embryos of all echinoderm classes share some features, including holoblastic

cleavage and similar cell movements during gastrulation. However, there are notable differences, such as the formation of micromeres in sea urchins but not brittle stars, the absence of pigment cells in brittle stars, and the formation of an embryonic skeleton in sea urchins and brittle star embryos, but not in the other groups. What is currently unclear is how these similarities and differences in development are reflected in the pattern of gene transcription. Davidson and Erwin [2006] have suggested that key gene regulatory subcircuits central to the formation of major morphological features ("kernels") are very highly conserved by stabilizing natural selection, both because they are critical to the formation of a complete viable body, and because their internal linkages and feedback loops make their component genes mutually dependent. A refinement of this idea is that some of the component transcription factors may be exchanged for others, as long the overall input/output logic and reliability of the circuit and its resulting function are maintained [Hinman and Davidson 2007]. This would suggest that many of the regulatory kernels shown to be important in sea urchin gastrulation would be conserved in the other echinoderm groups.

The set of genes that control skeleton formation in echinoderms may represent such a circuit under evolutionary constraints. All echinoderms form skeletons as adults, however, only sea urchins and brittle stars form extensive embryonic skeletal spicules. It has recently been shown that most of the same regulatory genes that underlie skeletogenesis in the sea urchin embryo are also expressed in the construction of the adult skeleton in both sea urchins and starfish [Gao and Davidson 2008]. The embryonic skeletons of sea urchins and

brittle stars are thus thought to be derived characters resulting from early activation of an adult gene regulatory network in the embryo.

The process of embryonic skeletogenesis has been extensively studied in sea urchins [Wilt and Ettensohn 2007]. Asymmetric fourth and fifth cleavages produce four small micromeres and four larger micromeres at the vegetal pole. Descendants of the larger micromeres ingress into the blastocoel just prior to gastrulation and become the primary mesenchyme (PMC), which soon produces the embryonic skeleton. Micromeres are a derived character unique to crown group sea urchins (euechinoids) [Ettensohn 2009]. Brittle stars and more basal sea urchin groups form very similar embryonic skeletons from apparently homologous mesenchymal cells without prior unequal cleavages [Wray and McClay 1988, this study].

I have sequenced and characterized the 40h gastrula transcriptome of the brittle star *Ophiocoma wendtii*. The gastrula stage was chosen because it represents the point in development at which the three primary germ layers diverge, with ingression of mesenchymal cells and invagination of the gut. At this point in sea urchins the gene regulatory networks that specify the germ layers are established and the genes that define the differentiated states of the tissues have begun to be activated. The early gastrula therefore expresses the greatest number and diversity of developmentally important genes. I report here that the brittle star gastrula expresses genes of all functional classes and appears to share many key developmental regulatory components with other echinoderms.

Some regulatory genes as well as genes expressed in differentiated tissues in the sea urchin gastrula were not found to be expressed in the brittle star gastrula.

METHODS

Animals and Embryos

Brittle stars (*Ophiocoma wendtii*) were collected from reefs and rubble piles in the shallow waters of Florida Bay near the Keys Marine Laboratory, Long Key, Florida, between April and October. Animals were sorted by sex, with gravid females identified by the presence of swollen purple gonads visible through the bursal slits. Sperm was obtained from 2-3 males by injection of 1-3mL of 0.5-1.0M KCl. Shedding of eggs from females was induced by a combination of heat and light shock. Animals were placed in containers in the dark with aeration at 30 to 32°C. Periodically the animals were exposed to bright light. Developing embryos were cultured at 25-27°C in filtered sea water. RNA was isolated using Trizol (Life Technologies, USA) following the manufacturer's protocol.

Characterization of Transcriptome Sequences

Sequencing and assembly of contiguous sequences was carried out as described by Meyer, *et. al.* [2009]. The comparisons of transcriptomes followed an all-by-all BLAST [Altschul *et al* 1997] approach where each comparison was databased. These results were then queried for the identification of orthologous genes using a reciprocal best BLAST (RBB) strategy, and for the identification of gene families following the method of Lerat, *et al.* [2005] as implemented

previously [Cooper *et al* 2010, Flynn *et al* 2010]. Gene families and singletons were then annotated using Homology Inspector (HomIn) software, a Java program that stores and queries a set of gene families using the database tool db4o for Java version 7.12. HomIn links gene families with annotation information including Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology Database (KO) categories [Kanehisa and Goto 2000], Clusters of Orthologous Groups of proteins (COGs) [Tatusov *et al* 1997 & 2003], Gene Ontology (GO) categories [Ashburner *et al* 2000], or any other available annotation.

Search for GRN Components

Glean3 predicted protein sequences for genes involved in the *S. purpuratus* developmental gene regulatory network were retrieved from SpBase [Cameron *et al* 2009] using the official gene name. These were used as queries to search the brittle star gastrula transcriptome sequences using TBLASTN at default settings. The best hit for each query was then used to search back against both sea urchin protein sequences and GenBank reference proteins using BLASTX. Sea urchin genes which had reciprocal best BLAST hits to brittle star with e-values of 1e-9 or better in both directions were designated as present in the brittle star gastrula transcriptome. These sequences can be found in GenBank using accession numbers JX60016 to JX60067.

Database and Analyses

Results from the automated BLAST searches were saved to a Microsoft Access database. This database and Microsoft Excel were used for the analyses involving presence/absence of expression, functional classes, and numbers of matches to other databases. Rarefaction curves were generated using EcoSim software [Gotelli and Entsminger 2011]. For the functional class analysis, KEGG ortholog clusters were used if they included genes from at least one animal taxon. When a KEGG cluster participates in more than one pathway within a functional class, it was counted only once within the larger functional class. (e.g. K00128 aldehyde dehydrogenase (NAD+) is part of five different pathways within the class of carbohydrate metabolism and two pathways in lipid metabolism, among many others, but was counted only once within each class in Figure 2.6B, and once in the total number of distinct KEGG animal clusters in Figure 2.6A.)

RESULTS AND DISCUSSION

Embryonic Development of *Ophiocoma wendtii*

The key stages of *Ophiocoma wendtii* development are shown in Figure 2.2. The egg is pigmented, and pigment granules are retained during cleavage stages, but disappear in the blastula. Cleavage is radial and holoblastic and is equal throughout cleavage, such that the micromeres characteristic of the sea urchin fourth cleavage division are not produced. A hollow blastula is formed, and cells ingress into the blastula to initiate gastrulation. The number of ingressing cells seems much larger than is typical in sea urchins, but we have not

quantitated the number or traced the lineage of individual cells. Archenteron formation occurs through invagination and convergent extension. A second group of mesenchyme cells forms at the tip of the archenteron and gives rise to the coelomic pouches, but no pigment cells appear. The skeletogenic mesenchyme cells gather in ventrolateral clusters as in sea urchins and begin to form the mineralized skeleton. The timing of development to hatching blastula is similar to sea urchins. However, following the invagination of endoderm, brittle star development proceeds at a slower rate relative to sea urchins. There is an initial invagination at 26-30h post-fertilization, but this persists for several hours before overt endomesoderm development proceeds. Also, unlike in sea urchins, the elongation of the skeletal rods is delayed relative to the extension of the archenteron, such that the archenteron has extended one third to halfway across the blastocoel before skeletal elements appear. When the gut is fully formed the skeleton is still composed of relatively small triradiate spicules. These then elongate such that the pluteus larva is very similar to that of sea urchins. The stage at which we isolated RNA for sequencing analysis is similar to Figure 2.2F. We chose that point when skeletal elements were first visible.

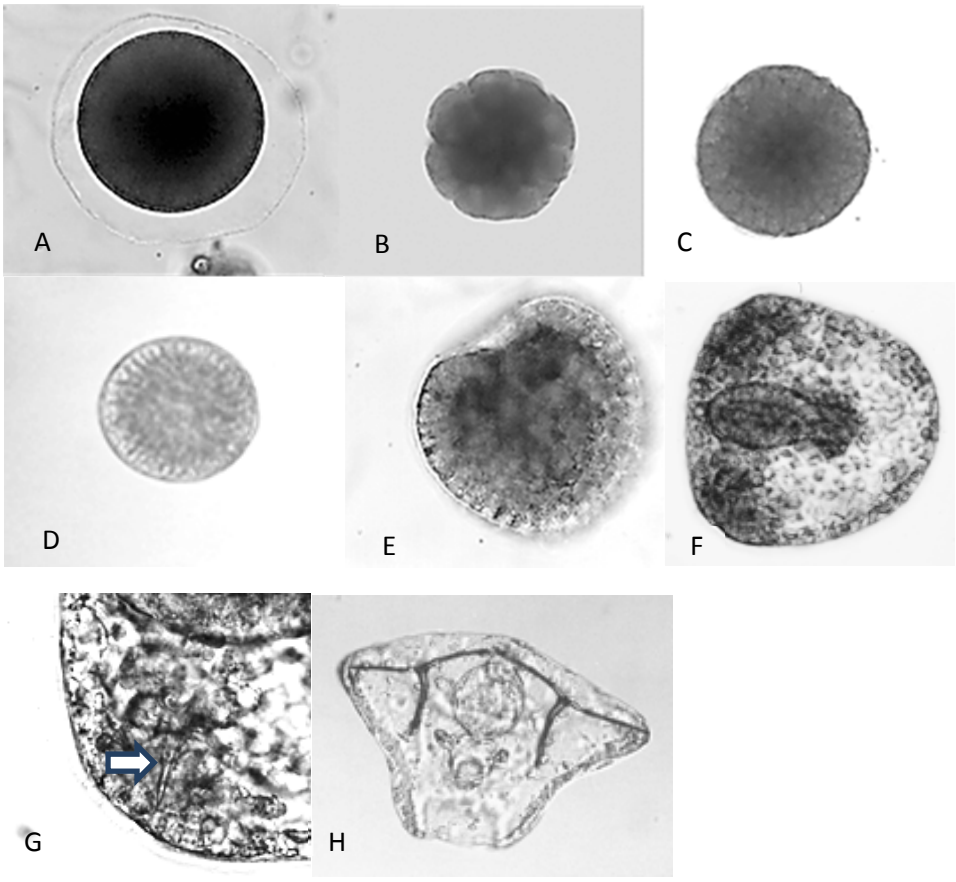

Figure 2.2: *Ophiocoma wendtii* Embryonic Development
Stages (**A**) egg, (**B**) 16 cell (5h), (**C**) hatched blastula (18h), (**D**) mesenchyme blastula (24h), (**E**) early gastrula (30h), (**F**) gastrula (40h), (**G**) ventrolateral cluster with skeletal spicule (arrow), (**H**) pluteus (80h).

Sequencing and Assembly

Pyrosequencing was performed on mRNA from gastrula stage brittle star embryos. After cleaning and trimming, there were 354,586 sequencing reads with a total of 75,031,136 basepairs [Figure 2.3A]. Lengths ranged from 16 to 439 basepairs, with approximately ¾ between 200 and 300. Less than 1% were longer than 300. Reads of 15 basepairs or shorter after trimming were not used for contig assembly. A total of 14,261 contigs were assembled, with a combined length of 5,488,581 basepairs [Figure 2.3B]. Median length increased by 23% over that of the unassembled reads (282 vs. 229), while average length

increased by 81% (384 vs. 212). Roughly ⅔ had lengths between 100 and 400 basepairs. The average number of reads per contig was 16.3, with a median of 5, a mode of 2, and a maximum of 8989. Coverage or depth ranged from 1x to 8549.4x, with an average of 7.1, median of 3.5, and standard deviation of 44.6 [Figure 2.3C].

Automated Annotation

Reciprocal best BLAST searches identified brittle star transcripts putatively corresponding to a total of 3385 orthologous genes in other databases [Figure 2.4]. The brittle star sequences were translated in all six reading frames and BLASTP was used to query the SpBase sea urchin Glean3 protein models. There were 3303 matches between brittle star and the sea urchin genome [Sea Urchin Genome Sequencing Consortium 2006]. Of these, 1863 also matched to the sea urchin combined UniGene transcriptome libraries [http://www.ncbi.nlm.nih.gov/UniGene/lbrowse2.cgi?TAXID=7668&CUTOFF=1000]. The KEGG Orthology database [http://www.genome.jp/kegg/ko.html] produced 1368 matches. More than two- thirds (2309 or 68%) of the identified brittle star genes had matches to more than one dataset. Almost a quarter (840 or 24.8%) matched to all three.

Note that the *O. wendtii* data were compared against each of the other datasets in Figure 2.4 separately. Therefore, brittle star sequences with hits to multiple datasets do not necessarily represent reciprocal best BLAST matches between every component in the annotation, but merely significant hits between

brittle star and more than one of the other datasets independently. Examination of the results reveals that the individual hits are mutually consistent in terms of genes identified.

The brittle star data have many times more sequences than the sea urchin gastrula UniGene set available on the NCBI UniGene database, and their average length is shorter. To assess whether we could make meaningful comparisons between these different data sets, we plotted the data as rarefaction curves. In ecology, rarefaction uses repeated random resampling of a large pool of samples to estimate the species richness as a function of the number of individuals sampled. Here we used it to estimate how thoroughly each data set represents the full transcriptome. In Figure 2.5, the curve for sea urchin has a much steeper initial slope, and therefore matches to a significant number of KEGG clusters even with many fewer sequences, probably because the sea urchin sequences are longer on average. The brittle star curve rises more gradually, but plateaus near the end, indicating that the sequencing captured most of the genes present in the transcriptome. If we assume that the two organisms express roughly the same number of genes at equivalent developmental stages, then the rarefaction curves indicate that this is indeed a meaningful comparison. The similar number of matches to the KEGG Orthology database for the two organisms also suggests this is the case.

Gene Functional Classes

O. wendtii sequences were compared to the KEGG Orthology database [http://www.genome.jp/kegg/ko.html] by reciprocal best BLAST [Figure 2.6A]. The KEGG Orthology database contains clusters of genes orthologous among a large number of organisms. Of the 3800 clusters relevant to animals, 1368 (36%) had significant matches to brittle star. Similarly, 1335 KEGG clusters (35%) had matches to sea urchin gastrula. These numbers include 840 KEGG clusters (22%) with matches to both organisms.

When sorted into functional classes [Figure 2.6B], an average of 43%, 39%, and 28% of the distinct KEGG clusters within each class had matches to brittle star, to sea urchin, and to both, respectively, with a range between 2% and 85%. Each KEGG functional class consists of a number of biochemical pathways. On average, 44%, 43%, and 30% of the KEGG clusters within each pathway had matches to brittle star, to sea urchin, and to both, respectively. Note that there is extensive overlap between the various KEGG functional classes and pathways, with many clusters falling into several different ones.

Overall, genes involved in metabolism and genetic information processing were the most highly conserved, as would be expected. The number of these “housekeeping” genes found in sea urchins and brittle stars are similar and the relationship between the number of genes observed in each group and the number shared between them is very consistent. There are fewer orthologs detected in the other KEGG orthology groups. Many pathways under “Organ Systems” and “Human Diseases” are vertebrate-specific and/or relate to

functions which do not operate extensively until later stages of development or after metamorphosis and would not be expected to be expressed at the gastrula stage. This is found to be true in both organisms. There is also more variation in the number of gene matches to sea urchins and brittle stars in these functional classes.

Genes involved with the cytoskeleton and cell junctions had considerably more matches to brittle star. Cell-adhesion genes are often large, with many exons, and with domains often repeated and shared between multiple genes [Whitaker *et al 2006*]. These characteristics, along with the short lengths of the brittle star sequences, have the potential to produce an artificially high number of BLAST hits. However, this pattern was the exception, not the rule across the other functional classes.

The sea urchin had a far greater number of matches to genes involved in endocytosis, lysosome and RNA degradation. Many of these genes again overlap with several other pathways, but there is no clear pattern to account for the disparity.

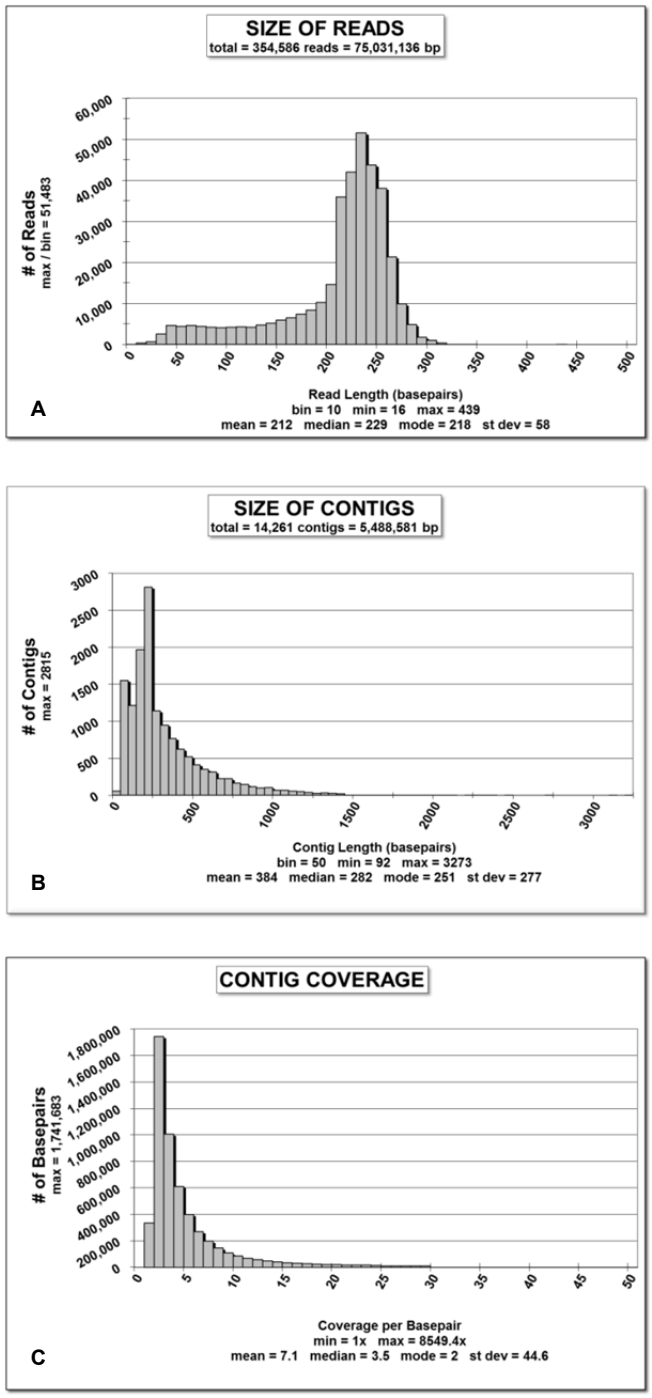

Figure 2.3: Pyrosequencing of Brittle Star Transcriptome
(**A**) After cleaning and trimming, 354,586 reads totaled 75,031,136 bp. Approximately ¾ had lengths between 200 and 300 bp. Less than 1% were longer than 300 bp. **(B)** A total of 14,261 contigs were assembled, with a combined length of 5,488,581 bp. Median length increased by 23% over that of the unassembled reads. Roughly ⅔ had lengths between 100 and 400 bp. Four percent were longer than 1000 bp, creating a long right-hand tail to the distribution. **(C)** The number of times a given nucleotide position is present in the reads used to assemble the contigs ranged from 1x to 8549.4x. Eighty-one percent were represented 1 to 5 times, while less than 1% had more than 100x coverage

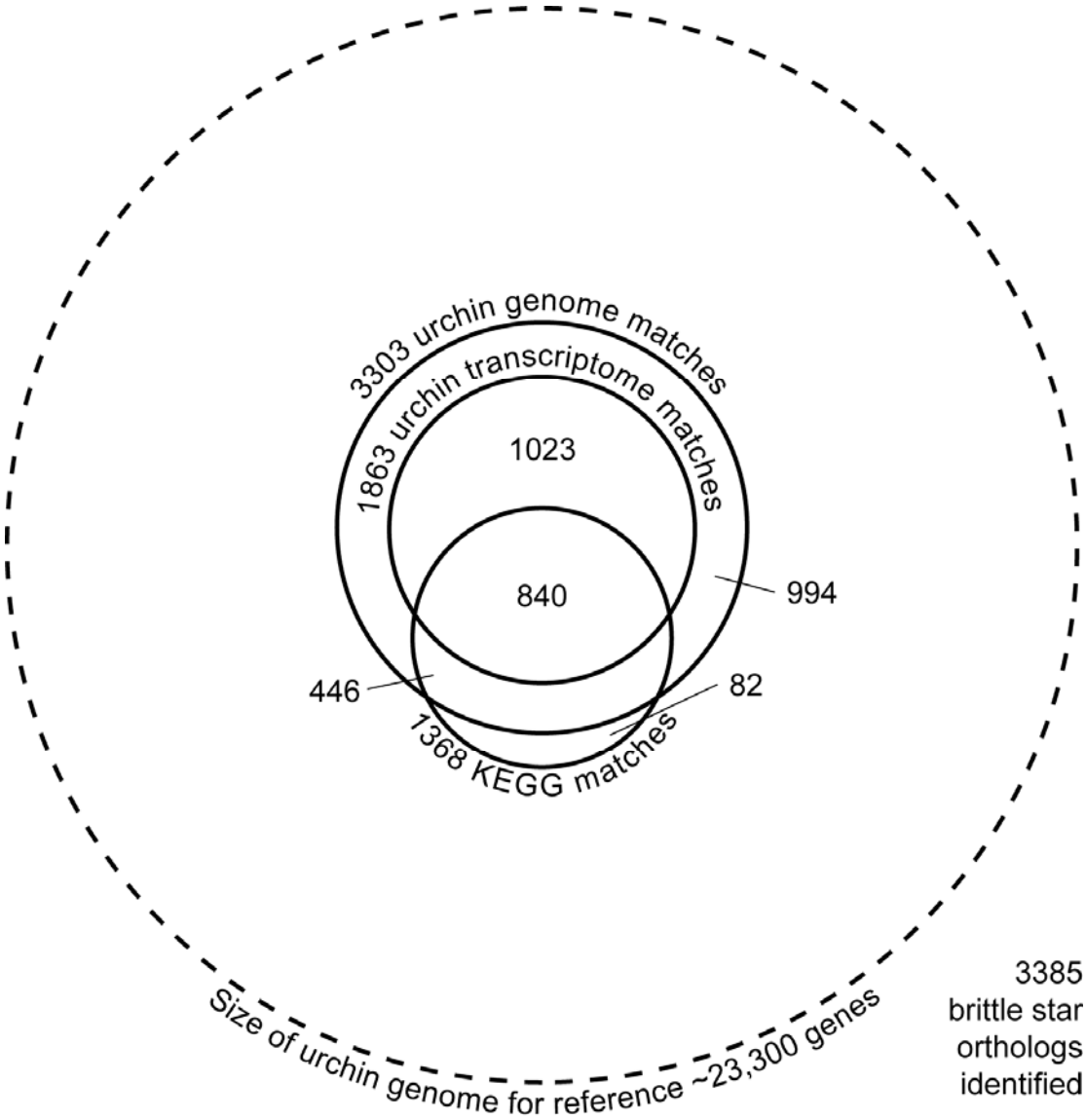

Figure 2.4: BLAST Identification of Brittle Star Genes
Automated BLAST was used to align *O. wendtii* cDNA sequences to both the genome and transcriptome of the sea urchin *S. purpuratus*, as well as to the KEGG Orthology database. The areas of the smaller circles represent the number of significant reciprocal best BLAST hits to the indicated datasets. Overlaps indicate matches of the same brittle star sequences to more than one dataset, and in nearly all such cases the matches from the different datasets are mutually consistent. For reference, the large dashed border represents the size of the *S. purpuratus* genome (~23,300 genes).

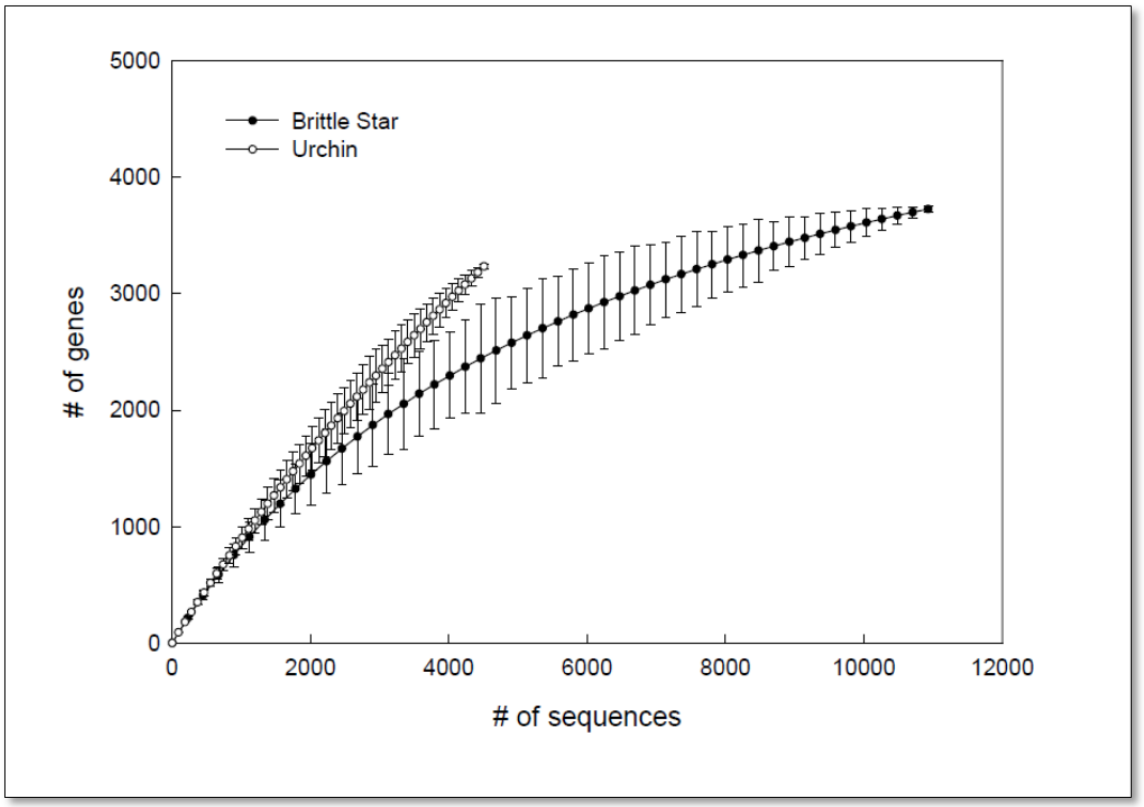

Figure 2.5: Rarefaction Curves for Sea Urchin and Brittle Star
The steeper initial slope for the sea urchin curve indicates matches to a significant number of KEGG clusters even with many fewer sequences. The brittle star curve rises more gradually, but becomes asymptotic at the right, indicating that the sequencing captured most of the genes present in the transcriptome. If the two organisms express roughly the same number of genes at equivalent developmental stages, then the rarefaction curves indicate that comparison of these two data sets is indeed meaningful.

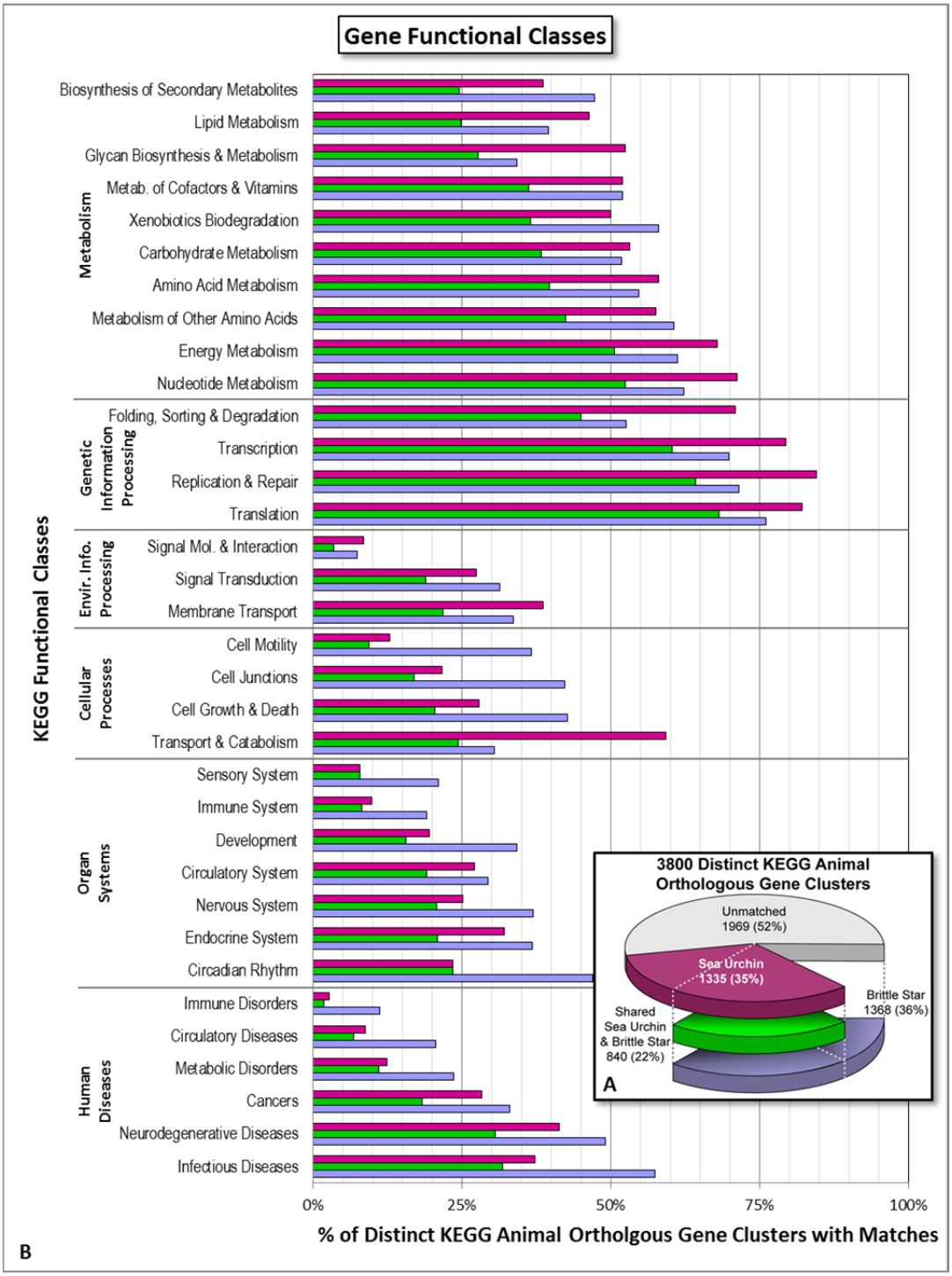

Figure 2.6: Gene Functional Classes Found in Brittle Star Transcriptome
(A) *O. wendtii* sequences were compared to the KEGG Orthology database by reciprocal best BLAST. Of 3800 distinct KEGG animal gene clusters, 36% had significant matches to brittle star (blue), and 35% had matches to sea urchin (purple). Green indicates the overlap between these two sets, i.e. KEGG clusters that match to both organisms (22%). **(B)** When sorted into functional classes, an average of 43%, 39%, and 28% of the KEGG clusters within each class had matches to brittle star, to sea urchin, or to both, respectively, with a majority of classes having similar representation in both organisms.

Comparison to Sea Urchin Developmental Gene Regulatory Network

The gene regulatory networks that underlie the differentiation of the basic tissue types in sea urchin embryos have been fairly well characterized. The temporal and spatial expression of these genes has been determined and many of the regulatory interactions between the various genes have been determined, either directly or inferred by interference with gene expression. The majority of these genes are expressed concurrently at the gastrula stage, which makes this stage an excellent point to identify a global set of genes important to the process of early cell differentiation. Here we use the sea urchin *Strongylocentrotus purpuratus* gastrula GRNs at 21-30h of development [Oliveri and Davidson 2004, Su 2009, Peter and Davidson 2010, Cameron *et al* 2009] as a reference to look for conservation of genes expressed in the brittle star gastrula at 40h of development, which is equivalent morphologically. At this point the skeletal spicules have just begun forming, the archenteron is one third to halfway across the blastocoel cavity and the equivalent of secondary mesenchyme has formed. The gut is not yet partitioned and no mouth has formed. The presence of the same genes expressed at the same stage in these two organisms would suggest a conservation of GRNs and a shared gastrula “toolkit” of proteins. The absence of genes expressed in either organism would indicate that there is either a temporal change in expression or that the gene is not expressed at all in the embryo of one group. Either is an indication of a change in a GRN. Reciprocal BLAST searches using the brittle star gastrula transcriptome data and the *S.*

purpuratus genome found homologs for a majority of genes involved in the sea urchin developmental gene regulatory network.

In sea urchins, a gradient of β-catenin initiated at the vegetal pole of the egg sets up and is soon reinforced by a circuit in the early embryo involving β--catenin/*lef1, wnt8, blimp1,* and *otx* in an intricate shifting relationship, creating a ring of gene expression which moves outward from the vegetal pole to specify endomesoderm [Smith *et al* 2007]. *Hox11/13b* is also soon involved in this circuit [Peter and Davidson 2010, Smith *et al* 2008]. Comparisons between sea urchins and starfish have revealed that just downstream from these early endomesoderm genes in the endoderm lies an extremely well-conserved kernel involving *blimp1/krox, otx, gatae, foxa,* and *brachyury* [Hinman *et al* 2003a]. In starfish, *tbr (t-brain)* is also part of this kernel, a role which is likely deeply ancestral, as it is also expressed in vegetal pole endoderm precursors in both sea cucumbers and hemichordates [Maruyama 2000, Tagawa *et al* 2000]. However, in sea urchins *tbr* has lost this role and has instead been co-opted into skeletogenesis [Hinman *et al* 2007b]. In sand dollars it appears to play both these roles [Minemura *et al* 2009].

Table 2.1 shows a comparison of some key endomesoderm and endoderm specific genes in the sea urchin with the transcripts present in the brittle star gastrula. Brittle stars express *β--catenin, lef1, otx, blimp1, wnt, hox/11/13b*, *and foxa* genes, suggesting that components of the endomesoderm and endoderm GRNs expressed early in development are conserved. *Gatae,* however, is not expressed.. Many animal phyla employ *gata* genes in gut

formation [Patient and McGhee 2002]. *Gatae* is a key component of the endoderm GRN in sea urchins and forms a feedback loop that maintains expression of these genes in the endoderm [Yuh *et al* 2004]. *Otx* and *blimp1* constitute another portion of that feedback loop [Peter and Davidson 2010], and this could be sufficient for endoderm differentiation in brittle stars. Two genes that are activated by *gatae* in *S. purpuratus*, *brachyury (bra)* and *krüppel (krl)*, are not expressed at gastrula stage in brittle stars. *Krüppel* expression in sea urchins is highest in the early blastula, and is mostly gone by the time of gastrulation in *S. purpuratus* [Howard *et al* 2001]. Its absence from the brittle star data may therefore reflect a small shift in timing and/or low transcript abundance at the onset of gastrulation. *T-brain* is not expressed in the brittle star gastrula. This would seem to indicate that *tbr* expression is not required for skeletogenesis in brittle star embryos as it is in sea urchins, or for endoderm formation as in starfish.

The endoderm in *S. purpuratus* is derived from two tiers of blastomeres formed during cleavage from the macromeres: Veg2, closest to the vegetal pole, and Veg1 above that. The Veg2 derived endoderm in *S. purpuratus* expresses *myc, brn1/2/4, tgif and dac* genes at the gastrula stage [Peter and Davidson 2011]. All of these are expressed in the brittle star gastrula [Table 2.1]. In contrast, the Veg1 genes *eve* and *hnf1* are not expressed in brittle stars. Together this suggests that a central early kernel of the endoderm GRN is conserved, although the expression of *gatae* and some genes it regulates are not. The expression of genes found in Veg 2 endoderm is also largely conserved.

The most likely explanation of our results is that the equivalent of Veg1 endoderm has not formed in the brittle star gastrula at the stage we examined. This suggests a heterochronic shift in the formation of the second tier of endoderm. This could also explain the absence of *brachyury*. It is a key player in gut formation in both protostomes and deuterostomes, though the details differ between taxa [Peterson *et al* 1999, Shoguchi *et al* 1999, Mitsunaga-Nakatsubo *et al* 2001, Gross and McClay 2003]. A shift in the timing of Veg1 endoderm formation could delay expression of *brachyury* in the brittle star. A less likely explanation is that a loss of this layer of endoderm has occurred in brittle stars, and that the gut is formed entirely by the equivalent of Veg 2 endoderm. *Endo16*, one of the major differentiation gene products in endoderm is not expressed in brittle star gastrula.

Following endomesoderm specification, Mesenchyme precursors all express *ets1/2, erg, and hex* in *S. purpuratus.* All three of these genes are expressed in the brittle star gastrula [Table 2.2]. The sea urchin skeletogenic primary mesenchyme derived from the micromeres, the homologous vegetal plate mesoderm in starfish, and the larval structures that produce the adult skeletons in both animals all express many of the same genes as sea urchin micromeres [Gao and Davidson 2008, Ettensohn *et al* 2007, McCauley *et al* 2010], and a majority of these genes were found in the brittle star gastrula transcriptome as well [Table 2.2]. In all cases of echinoderm skeleton formation studied, including brittle star embryos, *alx1* is expressed. This is consistent with reports that ectopic expression of *alx1* in sea urchin non-skeletogenic

mesenchyme (NSM) induces skeleton formation [Ettensohn *et al* 2007]. *Ets1* is expressed both maternally and zygotically, and is involved in all the above cases, activating a great number of downstream genes. *Ets1* and *alx1* were both found in the brittle star gastrula transcriptome. Just downstream from these in both sea urchin micromere and starfish vegetal plate mesoderm are a group of three genes, *erg, hex,* and *tgif,* which form a "lockdown" mechanism, stabilizing the specification state by feeding back to each other and to *tbr* and *ets1*, and feeding forward into tissue- specific differentiation genes [McCauley *et al* 2010]. All three were present in brittle star, as was *deadringer (dri)*, which appears to play a similar role in all the skeletogenic cases. *Tbr* has not been found in brittle stars. In starfish, *tbr* is seen in both endoderm (discussed above) and mesoderm [McCauley *et al* 2010]. It does not appear to be involved in adult skeletogenesis in either starfish or sea urchins. Its absence in brittle stars reinforces the idea that it was not part of the ancestral skeletal GRN and that its role in sea urchin embryonic skeleton formation is derived.

The downstream differentiation genes found in *S. purpuratus* skeletogenic cells at the gastrula stage are also found in the brittle star gastrula [Table 2.2]. The spicule matrix proteins of the sea urchin endoskeleton contain a single C-type lectin domain and repetitious stretches rich in proline and glycine [Livingston *et al* 2006, Mann *et al* 2008 & 2010]. The apparently loose constraints on primary structure in these proteins, and the resulting low sequence conservation make identification of brittle star homologs difficult. However, the brittle star gastrula transcriptome contains several transcripts encoding C-type lectin domains and

repetitive regions. Several other proteins, including Cyclophilin and Ficolin, are all expressed in sea urchin PMC cells and associated with the skeleton, though their exact functions remain unclear. The brittle star gastrula transcriptome contains matches for *cyclophilin* and *ficolin*, but not for *MSP130*, a major cell surface protein in sea urchin PMCs. Overall there is a remarkable conservation of the GRN leading to formation of mineralized tissue in the embryos of sea urchins and brittle stars.

In sea urchins, Delta-Notch signalling from the micromeres activates *gcm* in the adjacent NSM to form pigment cells [Ransick and Davidson 2006, Croce and McClay 2010]. Brittle star embryos do not form embryonic pigment cells. Neither do starfish, but they express *gcm* in ectoderm rather than mesoderm, and it does not depend on Delta signalling [Hinman *et al* 2007a]. Neither *notch* nor *delta* is expressed in the brittle star gastrula [Table 2.2]. *Gcm* is expressed in brittle star gastrula, but *gatac, gatae*. *six1/2,*and *scl* are not. This suggests that the GRN leading to pigment cells, not surprisingly, is not conserved in brittle stars. Likewise, most of the genes that are expressed in the *S. purpuratus* small micromeres (i.e *soxe, foxy*), which are not formed outside of euechinoids, are not expressed in the brittle star gastrula [Table 2.2].

In sea urchin ectoderm, Nodal patterns both the ventro-dorsal (oral-aboral) and left-right axes [Duboc and Lepage 2008a], but was not found in brittle star [Table 2.3]; nor was its antagonist Lefty, which soon limits Nodal to the ventral side [Duboc *et al* 2008b]. On the other hand, a number of genes downstream from Nodal and key to specification of different ectodermal regions [Saudemont

et al 2010] were found in brittle star [Table 2.3]. Most of the genes expressed in the *S.purpuratus* oral ectoderm are found in the brittle star gastrula transcriptome, including *chordin* and *BMP2/4*. Sea urchin BMP2/4 is expressed in the oral ectoderm, then diffuses to and specifies the aboral ectoderm by inhibiting Nodal [Lapraz *et al* 2009], while Chordin helps pattern neural tissue in the ciliary band at the oral/aboral border by excluding BMP2/4 activity from the oral side [Bradham *et al* 2009]. Genes that, in the sea urchin, are activated by Nodal-independent early oral ectoderm input are found to be expressed in brittle star gastrula. These include *otxb1/2* and *hnf6.* Of the sea urchin genes that are activated at the boundary of ectoderm and endoderm, *foxj* is expressed in brittle star gastrula, but *lim1* and *nk1* are not.

Genes that are expressed in the sea urchin aboral ectoderm are not as uniformly expressed in brittle star gastrula. Genes expressed by 12h of sea urchin development such as *sim* and *nk2.2* are expressed in brittle star gastrula, but not genes expressed later in sea urchin aboral ectoderm such as *hox7* and *msx*, or the differentiation genes *spec1* and *spec2a*. *Tbx2/3* is expressed in brittle star gastrula, but not *irxa* and *dlx*, which are activated by Tbx2/3 in sea urchins. Taken together this would suggest two heterochronic shifts in ectoderm determination between sea urchins and brittle stars. In sea urchins, all of the genes examined are expressed at the gastrula stage. It appears that in brittle stars patterning by Nodal and Lefty is complete by gastrula and these genes are no longer expressed. Oral ectoderm is determined and specification of the aboral

ectoderm is underway, but it appears that this process is not complete in the 40h brittle star gastrula.

Table 2.1: Comparison of *O. wendtii* gastrula transcripts to the *S. purpuratus* endodermal and endomesodermal gene regulatory networks

Gene	Found in *O.w.* gastrula	RBB to *S.p.* Genome	RBB to NCBI RefSeq Proteins	Role in *S.p.*
β-Catenin	Y	β-Catenin [SPU_004319]	*S.p.* β-Catenin [XP_786059.2]	Endoderm
Otx	Y	Otx [SPU_010424]	*S.p.* Otx [NP_999753.2]	Endoderm
Wnt	Y	Wnt5 [SPU_026277]	*S.k.* Wnt2 [NP_001158455.1]	Endoderm
Blimp1	Y	Blimp1/Krox [SPU_027235]	*B.f.* Zn-finger [XP_002587482.1]	Endoderm
Hox11/13b	Y	Hox11/13b [SPU_002631]	*S.p.* Hox11/13b [NP_999774.1]	Endoderm
Bra	N			Endoderm
Krl	N			Endoderm
Myc	Y	Myc [SPU_003166]	*S.p.* Myc [NP_999744.1]	Endoderm
SoxB1	Y	SoxB1 [SPU_022820]	*S.p.* SoxB1 [NP_999639.1]	Endoderm
Brn1-2-4	Y	Brn1-2-4 [SPU_016443]	*S.p.* Brn1-2-4 [XP_782909.2]	Endoderm
Tgif	Y	Tgif [SPU_018126]	*I.s.* Tgif [XP_002433653.1]	Endoderm
Hnf1	N			Endoderm
Eve	N			Endoderm
Hh	N			Endoderm
VEGF	Y	VEGF [SPU_030148]	*H.p.* VEGF BAI67115.1]	Endoderm
Dac	Y	Dac [SPU_028061]	*I.s.* Dachsund [XP_002407755.1]	Endoderm
Endo16	N			Endoderm
FoxA	Y	FoxA [SPU_006676]	*S.p.* FoxA [NP_001073010.1]	Endo+SMC
GataE	N			Endo+SMC
Kakapo	Y	Syne1 [SPU_013237]	*S.p.* Similar to CG33715-PD [XP_784190.2]	Endo+SMC
Apobec	Y	Hnrpr [SPU_019557]	*S.p.* Hnrpr [XP_793277.1]	Endo+SMC
Gelsolin	Y	Gelsolin [SPU_003985]	*S.p.* Gelsolin [XP_788777.1]	Endo+SMC

B.f. = Branchiostoma floridae, H.p. = Heliocentrotus pulcherrimus,
I.s. = Ixodes scapularis, O.w. = Ophiocoma wendtii,
S.k. = Saccoglossus kowalevskii, S.p. = Strongylocentrotus purpuratus

Table 2.2: Comparison of *O. wendtii* gastrula transcripts to the *S. purpuratus* mesenchymal gene regulatory network

Gene	Found in *O.w.* gastrula	RBB to *S.p.* Genome	RBB to NCBI RefSeq Proteins	Role in *S.p.*
HesC	Y	HesC [SPU_021608]	*S.p.* HesC [XP_796692.1]	Mesenchyme
Erg	Y	Erg [SPU_018483]	*S.p.* Erg [NP_999833.1]	Mesenchyme
Hex	Y	Hex [SPU_027215]	*S.p.* Hex [XP_001197103.1]	Mesenchyme
Ets1/2	Y	Ets1/2 [SPU_002874]	*S.p.* Ets1/2 [NP_999698.1]	Mesenchyme
Alx1	Y	Alx1 [SPU_025302]	*S.p.* Alx1 [NP_999809.1]	PMC
Tbr	N			PMC
Tgif	Y	Tgif [SPU_18126]	*I.s.* Tgif [XP_002433653.1]	PMC
FoxN2/3	N			PMC
Dri	Y	Dri [SPU_017106]	*S.p.* Dri [NP_999799.1]	PMC
FoxB	Y	FoxB [SPU_004551]	*S.p.* FoxB [NP_999797.1]	PMC
FoxO	Y	FoxO [SPU_009178]	*S.p.* FoxO [XP_001183650.1]	PMC
VEGFR	N			PMC
Delta	N			PMC
Spicule matrix genes	Possible	C-lectin [SPU_007882]	*S.p.* C-lectin [NP_999805.1]	Skeletal Differentiation
MSP130	N			Skeletal Differentiation
G-Cadherin	Y	G-Cadherin [SPU_015960]	*S.k.* G-Cadherin [XP_002741140.1]	Skeletal Differentiation
Ficolin	Y	Fic [SPU_023548]	*B.f.* Ficolin [XP_002594892.1]	Skeletal Differentiation
Cyclophilin	Y	CypL7 [SPU_008305]	*D.m.* Cyclophilin 1 [NP_523366.2]	Skeletal Differentiation
Gcm	Y	Gcm [SPU_006462]	*S.k.* Gcm [XP_002733441.1]	SMC
Notch	N			SMC
Six1/2	N			SMC
Hnf6	Y	Hnf6 [SPU_016449]	*S.p.* Hnf6 [NP_999824.1]	SMC
GataC	N			SMC
Scl	N			SMC
Pks	Y	Pks [SPU_028395]	*S.p.* Pks [NP_001239013.1]	SMC
FoxF	Y	FoxF [SPU_000975]	*S.p.* FoxF [XP_794135.1]	Small Micromeres
SoxE	N			Small Micromeres
FoxY	N			Small Micromeres

B.f. = Branchiostoma floridae, D.m. = Drosophila melanogaster, I.s. = Ixodes scapularis, O.w. = Ophiocoma wendtii, S.k. = Saccoglossus kowalevskii, S.p. = Strongylocentrotus purpuratus

Table 2.3: Comparison of *O. wendtii* gastrula transcripts to the *S. purpuratus* ectodermal gene regulatory network

Gene	Found in *O.w.* gastrula	RBB to *S.p.* Genome	RBB to NCBI RefSeq Proteins	Role in *S.p.*
Nodal	N			Oral Ectoderm
Lefty	N			Oral Ectoderm
Chordin	Y	Chordin [SPU_004983]	*S.k.* Chordin [NP_001158390.1]	Oral Ectoderm
Sip1	N			Oral Ectoderm
FoxG	N			Oral Ectoderm
BMP2/4	Y	BMP2/4 [SPU_000669]	*S.p.* BMP2/4 [NP_001116977.1]	Oral Ectoderm
FoxA	Y	FoxA [SPU_006676]	*S.p.* FoxA [NP_001073010.1]	Oral Ectoderm
Bra	N			Oral Ectoderm
Dri	Y	Dri [SPU_017106]	*S.p.* Dri [NP_999799.1]	Oral Ectoderm
Hes	Y	Hes [SPU_006814]	*S.k.* Hes1 [NP_001158466.1]	Oral Ectoderm
Hnf6	Y	Hnf6 [SPU_016449]	*S.p.* Hnf6 [NP_999824.1]	Oral Ectoderm
FoxJ1	Y	FoxJ1 [SPU_027969]	*S.p.* FoxJ1 [NP_001073013.1]	Ecto/Endo Border
Nk1	N			Ecto/Endo Border
Lim1	N			Ecto/Endo Border
Tbx2/3	Y	Tbx2/3 [SPU_023386]	*S.p.* Tbx2/3 [NP_001123280.1]	Aboral Ectoderm
Lhx2 (Lim2)	Y	Lhx2 [SPU_021313]	*M.m.* Lhx2 [NP_034840.1]	Aboral Ectoderm
Dlx	N			Aboral Ectoderm
Nk2.2	Y	Nk2.2 [SPU_000756]	*S.p.* Nk2.2 [NP_001123283.1]	Aboral Ectoderm
Hox7	N			Aboral Ectoderm
Msx	N			Aboral Ectoderm
Klf7	Y	Klf2/4 [SPU_020311]	*S.k.* Klf2 [NP_001161575.1]	Aboral Ectoderm
IrxA	N			Aboral Ectoderm
Hmx	N			Aboral Ectoderm

M.m. = Mus musculus, O.w. = Ophiocoma wendtii,
S.k. = Saccoglossus kowalevskii, S.p. = Strongylocentrotus purpuratus

Table 2.4: Conservation of genes between *S. purpuratus* and *O. wendtii*

		% conserved in *O. wendtii*
S. purpuratus Transcriptome		55
Early Gastrula GRN		65
Endoderm		53
	Veg2 Endoderm	70
	Veg1 Endoderm	20
Primary Mesenchyme		86
Non-Skeletogenic Mesenchyme		57
	Secondary Mesenchyme	64
	Small Micromeres	33
Oral Ectoderm		67
Aboral Ectoderm		58

CONCLUSIONS

The brittle star *Ophiocoma wendtii* exhibits radial holoblastic cleavages that are equal throughout, giving rise to uniform-sized blastomeres without the formation of the micromeres characteristic to sea urchins. Despite this, mesenchymal cells ingress and give rise to an embryonic skeleton, a developmental structure unique to echinoids and ophiuroids among the echinoderms. Mesenchymal cells also give rise to the coelomic pouches, but no pigment cells are formed in the embryo. Archenteron formation occurs much the same as in sea urchins, although there is a delay in gut elongation following invagination as well as in growth of the skeletal spicules initiated in the ventrolateral clusters. The resulting pluteus larva closely resembles that of sea urchins, albeit without pigment cells. The *O. wendtii* gastrula expresses genes from all functional classes at the gastrula stage. Brittle stars and sea urchins have comparable numbers of genes in most functional classes expressed at the gastrula stage.

A majority of the genes involved in the sea urchin gene regulatory network were also found in the brittle star gastrula transcriptome [Table 2.4]. The brittle star pyrosequencing data are completely consistent with our earlier results using a PCR- based candidate gene approach (not shown). For example, transcripts of *alx1, dri, gabp, ets1*, and *erg* were found by both methods, whereas *tbr, gatac,* and *gatae* were not. The percentage of genes involved in gene regulatory networks expressed in *S. purpuratus* gastrula that are also expressed in *O. wendtii* gastrula exceeds the percentage of transcripts conserved overall [Table 2.4]. However, this conservation is not uniform across the different tissue types found in echinoderm gastrulae. Some of these differences can be explained by heterochronic shifts in gene expression, although gene loss is also a possibility. Some of the endomesoderm genes that are expressed in sea urchin gastrula at declining levels could be undetectable by the brittle star gastrula stage. Examination of the aboral ectoderm genes expressed in *O. wendtii* relative to *S. purpuratus* indicates that specification of aboral ectoderm has begun but is delayed in the brittle star. The same could be true for the Veg1 endoderm. Other differences in gene expression correlate with differences in embryonic development. Brittle star embryos do not possess micromeres or pigment cells. The second lowest percentage of GRN genes conserved (33%) is seen in the genes expressed in *S. purpuratus* small micromeres and pigment cells [Table 2.4].

The highest percentage of GRN conservation is seen in the skeletogenic mesenchyme cells (PMCs in sea urchins). This is not surprising, since all adult

echinoderms form mineralized structures. The GRN and differentiation genes that lead to mineralized structures must be conserved in order for the adult skeleton to form. In sea urchins this GRN is activated in the embryo largely intact. The conservation of these genes in the *O. wendtii* gastrula suggests that is the case in brittle stars as well.

Hesc is a transcriptional repressor ubiquitously expressed in the sea urchin embryo, where its role is to keep the skeleton program off. In the sea urchin micromeres, *hesc* is itself repressed by Pmar1 in response to nuclearized β-catenin, thereby de-repressing the skeleton circuits [Revilla-i-Domingo *et al* 2007]. This double-negative *pmar1/hesc* gate appears unique to sea urchins as the mechanism that coupled the pre-existing programs of skeletogenesis and maternal β-catenin-mediated vegetal specification to produce the novelty of the embryonic skeleton, as it is not involved in adult sea urchin skeletogenesis [Gao and Davidson 2008, Ettensohn *et al* 2007]. Recent evidence suggests that other, as yet unknown mechanisms related to the unequal cleavage that produces the micromeres are also involved [Sharma and Ettensohn 2010]. Starfish, which do not build an embryonic skeleton, also express *hesc* throughout most of the embryo, but it appears to have no effect on mesodermal genes shared with sea urchin skeletogenesis, and *pmar1* has never been found in starfish [McCauley *et al* 2010].

Sea urchins express *pmar1* from fourth cleavage through mid-blastula, so it would not be expected to be seen in the *O. wendtii* gastrula transcriptome. Using PCR, our lab has searched for, but never found, *pmar1* transcripts from

any stage of brittle star development. We have, however, successfully amplified the *pmar1* homolog from brittle star genomic DNA, identified as such by the presence of a conserved intron [unpublished]. This suggests that activation of the adult skeletal GRN in embryos occurred differently in brittle stars than in sea urchins. Overall, the data suggest that embryonic skeleton formation in sea urchins and brittle stars represents convergent evolution by independent cooptation of a shared pathway utilized in adult skeleton formation.

Chapter Three:
A Survey of Transcription Factors in the Brittle Star *Ophiocoma wendtii* Gastrula Embryo

BACKGROUND

Transcription factors are a primary link between genotype and phenotype, and a major factor in evolution. The interactions between transcription factors, signalling molecules, and the genes which they target form gene regulatory networks (GRNs), the means by which developmental instructions scattered throughout the genome are organized and implemented at the proper places and times. Patterns of transcription in different regions of the egg and early embryo specify the primary germ layers and body axes. As development proceeds, the various GRN circuits active in each portion of the embryo turn on the genes for tissue-specific proteins and pathways, leading toward progressively more refined and specialized structures and physiology and mature functional organs.

Transcription factor proteins regulate transcription of target genes by binding to *cis*-regulatory DNA regions upstream of the coding sequence. Transcription factors are thus defined by their DNA-binding domains. Many also contain other functional domains such as phosphorylation sites or protein-protein interaction domains. Because of their critical roles in the construction of complete viable bodies, transcription factors tend to be very highly conserved at the amino acid sequence level within their important domains, as well as in the arrangements of the different domains within the protein. Regions outside these

domains have varying—and often very low—degrees of sequence homology. Mutations which alter the binding specificity of transcription factors or the promoter elements of the genes they regulate can result in changes to the location and/or timing of expression of those targeted genes, thereby altering aspects of the GRNs which build bodies, and resulting in bodies which look or function differently [Davidson and Erwin 2006, Ettensohn *et al* 2007, Hinman *et al* 2007a].

As basal deuterostomes, echinoderms occupy an important phylogenetic position between chordates and all other animal phyla. The GRN behind development of the purple sea urchin *Strongylocentrotus purpuratus* has been extensively studied in the stages from egg through gastrulation [Davidson *et al* 2002a & 2002b, Oliveri and Davidson 2004, Cameron *et al* 2009, Su 2009, Peter and Davidson 2010]. The equivalent GRNs in other echinoderm classes have so far been much less studied. As part of the *S. purpuratus* genome sequencing project [Sea Urchin Genome Sequencing Consortium 2006], systematic searches were conducted to find all the transcription factor genes in the urchin genome [Arnone *et al* 2006, Howard-Ashby *et al* 2006a & 2006b, Materna *et al* 2006a, Rizzo *et al* 2006, Tu *et al* 2006]. Here, I present the results of a search through the early gastrula-stage transcriptome of the brittle star *Ophiocoma wendtii* for homologs of all the *S. purpuratus* transcription factors. Roughly one third of the transcription factor genes in the *S. purpuratus* genome, and half of those expressed in gastrula-stage sea urchin embryos, were found in the brittle star

gastrula. Comparisons to the gastrula transcriptomes of the green sea urchin *Lytechinus variegatus* and the starfish *Patiria miniata* produced similar results.

METHODS

Protein sequences for the *S. purpuratus* transcription factor genes listed in Arnone *et al* 2006, Howard-Ashby *et al* 2006a & 2006b, Materna *et al* 2006a, Rizzo *et al* 2006, and Tu *et al* 2006 were retrieved from SpBase [http://www.spbase.org/SpBase/, Cameron *et al* 2009]. These were used as queries to search the brittle star gastrula transcriptome sequences using TBLASTN [Altschul *et al* 1997] at default settings. The best hit for each query was then used to search back against both sea urchin protein sequences and NCBI RefSeq proteins [http://www.ncbi.nlm.nih.gov/RefSeq/] using BLASTX. Urchin genes which had reciprocal best BLAST (RBB) hits to brittle star with e-values of 1e-9 or better in both directions were designated as high-confidence homologs and color-coded green in Tables 3.1–3.15 and Figures 3.1–3.2. Those with e-values of 1e-6 to 2e-9 were designated as possible homologs and coded yellow, while those with e-values worse than 1e-6 were considered poor matches and coded pink. Brittle star transcripts that matched best to different genes in different organisms were coded blue-purple. Genes with matches that fell into more than one of these quality categories were assigned an overall quality rating based on the preponderance of the evidence, or, if there was no clear preponderance, assigned to the lower quality category. Gastrula transcriptome sequences from *Lytechinus variegatus* [http://www.spbase.org/LV/index.php] and

Patiria miniata [http://www.spbase.org/PM/index.php] were then used to search against brittle star in the same way using TBLASTX. Genes were considered to be transcribed at a biologically significant level in *S. purpuratus* if they had a concentration of ≥150 transcripts/embryo at ~30 hours post-fertilization (early gastrula stage) according to SpBase [http://www.spbase.org/SpBase/, Cameron *et al* 2009] or Materna *et al* 2010 [http://vanbeneden.caltech.edu/~m/cgi-bin/hd-tc/plot.cgi], since this was the threshold used in previous studies of transcription factor expression in *S. purpuratus* [Howard-Ashby *et al* 2006b].

RESULTS AND DISCUSSION

Overview

The *S. purpuratus* genome contains 284 transcription factors, not including C_2H_2 zinc-finger genes [Figure 3.1a]. Ninety-nine (35%) of these had moderately good (1e-6 to 2e-9) or excellent (≤ 1e-9) matches to the brittle star gastrula transcripts, while 164 (58%) were not found. The remaining 21 (7%) had either ambiguous results or only poor matches.

At the 30h early gastrula stage, *S. purpuratus* expresses 166 (58%) of these transcription factor genes at biologically significant levels of at least 150 transcripts/embryo. Of the genes expressed, 82 (49%) had good matches to brittle star, while 44% were not found [Figure 3.1b]. The currently available transcriptome data for *Lytechinus variegatus* and *Patiria miniata* are less complete, containing 111 and 124 transcription factor genes respectively. The results, however, are quite similar to those for *S. purpuratus,* with 52% of *L.*

variegatus transcription factors and 57% of those from *P. miniata* having good or excellent brittle star matches [Figures 3.1c-d]. Forty-two percent of the transcription factor genes that were not found in the brittle star gastrula are also not expressed at this stage in any of the other three organisms.

Figure 3.2 depicts the number and quality of brittle star matches for each gene family. Within each family, the proportions of high-quality matches are rather consistent for all three transcriptomes. The bHLH, nuclear receptor, and homeobox families have many genes which are not expressed at the gastrula stage in urchins or starfish, and therefore would not be expected to match to brittle star gastrula. When unexpressed genes are excluded, the bHLH and homeobox families have results similar to those for the other families. With the exception of the nuclear receptors, which are more rapidly evolving, the three transcriptomes match to brittle star for 45-64% of the expressed genes in each family, with the exception of starfish in the bZip and Ets families, where it matches to brittle star at 69% and 80% respectively.

While the spatial expression of these genes, and the precise regulatory interactions between them are as yet unknown in brittle stars, the general patterns of what was found and not found in the sequencing results, in the context of broadly conserved developmental mechanisms and roles played by these genes in other organisms, can allow us to draw some tentative conclusions about the likely structure of the brittle star developmental gene regulatory network.

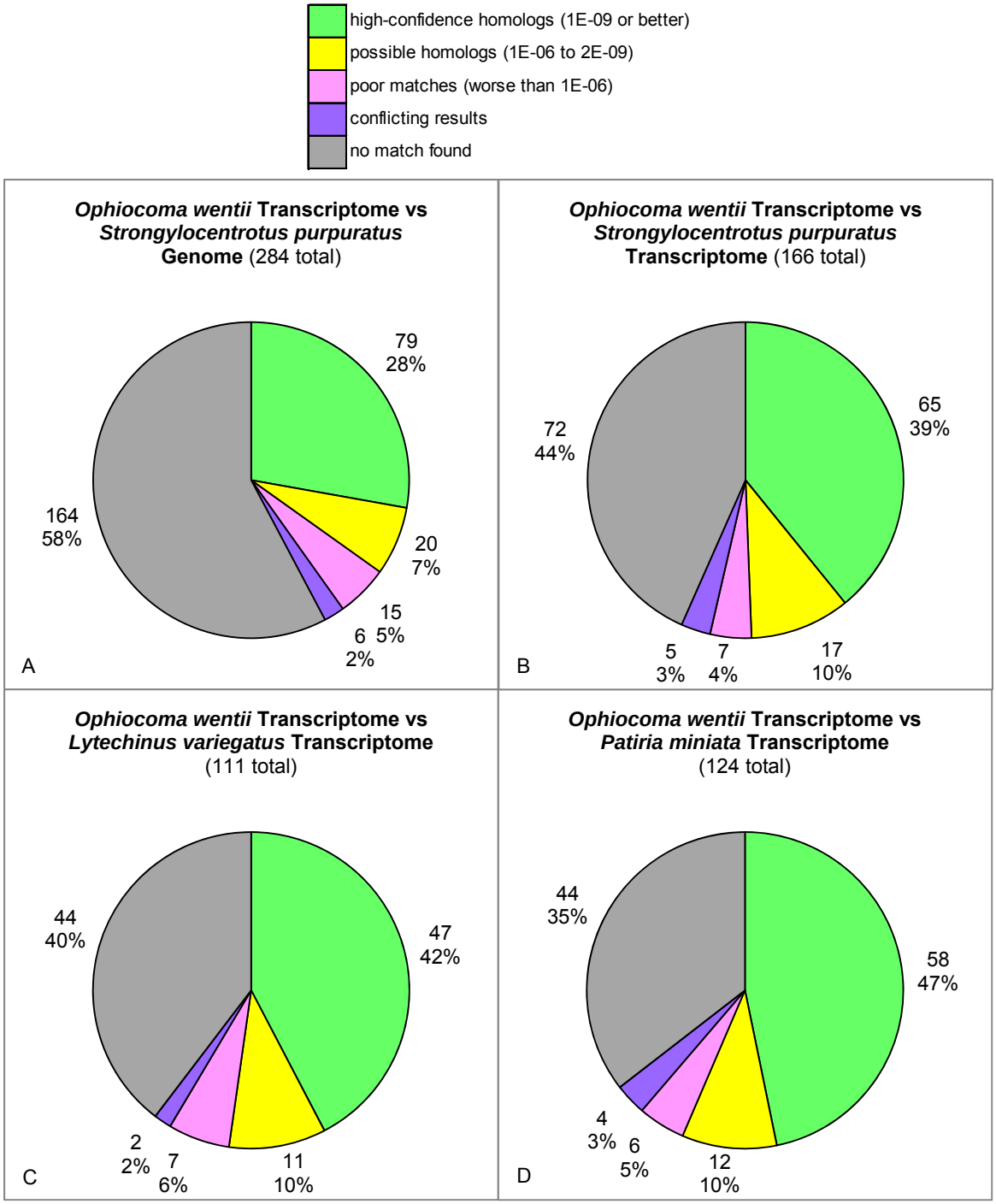

Figure 3.1: Transcription Factor Genes Found in Brittle Star Transcriptome.
Results of reciprocal best BLAST of transcription factor genes from other organisms vs. the brittle star gastrula transcriptome.

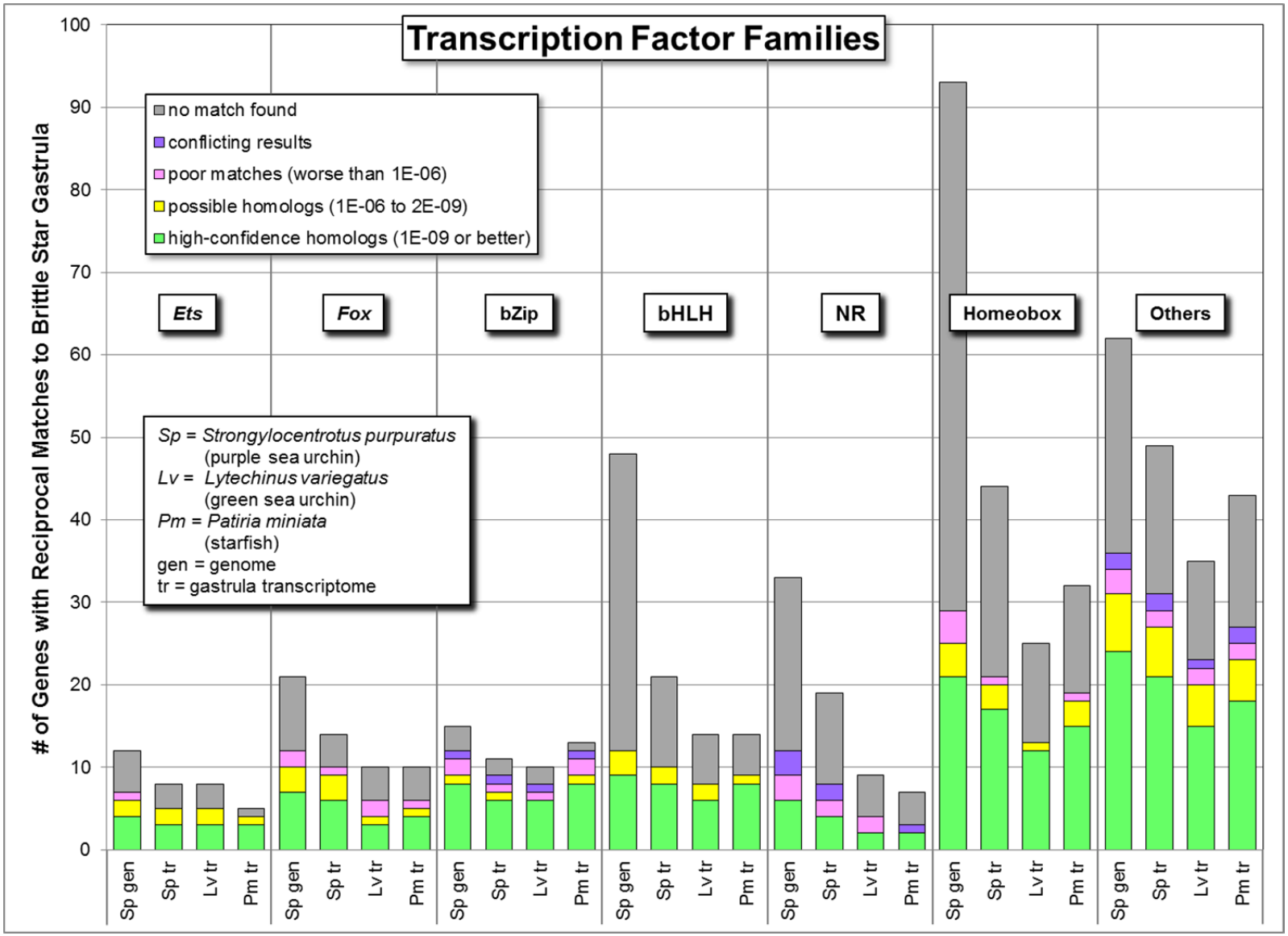

Figure 3.2: Transcription Factor Families
Results of reciprocal best BLAST of transcription factor genes from other organisms vs. the brittle star gastrula transcriptome.

Ets Gene Family

The Ets family of transcription factors is defined by the ETS winged helix-turn-helix DNA-binding domain. Most also contain a PNT (POINTED) domain which interacts with other proteins. Such interactions, as well as signalling pathways, especially MAPK, and phosphorylation state all contribute to the activity and sequence-binding specificity of particular Ets proteins [Sharrocks 2001].

Multiple *ets* genes have been found throughout the animal kingdom, from sponges to mammals [Degnan *et al* 1993], and duplications have greatly expanded the Ets family in vertebrates. The sea urchin genome includes eleven *ets* genes, one homolog corresponding to each vertebrate *ets* subfamily (except one which is mammal-specific) with 73-96% amino-acid sequence identity between urchin and human for eight of these [Rizzo *et al* 2006]. Good matches for six *ets* genes were found in brittle star [Table 3.1].

Sp-Ets1/2 is expressed at very high levels. Ubiquitous at the earliest stages, it soon localizes to the PMC cells, where it is central to embryonic skeleton formation. *Ets1/2* was found in brittle star, as were *Erg* and *Tel* which also participate in urchin PMCs, as well as SMCs in the case of *Erg*. Two other urchin *ets* genes, *Ese* and *Pea* are confined to the SMCs at gastrulation. *Ese* was not found in brittle star. *Pea* had only moderate e-values (~e-9) vs. *S. purpuratus*, though the match was to the PEA3 MAP kinase activation domain, and the matches between *Lytechinus* and brittle star were better (~e-12). *Ets4,* which helps define the nonvegetal portion of the early urchin embryo, and

activates the gene for hatching enzyme [Wei *et al* 1999], was also found in brittle star.

Of the *ets* genes not found in brittle star, *Elk* and *Erf* are expressed ubiquitously in urchin. *Sp-Pu1* is not expressed significantly until 72 hpf (pluteus stage). *Sp-Elf* is not expressed until late gastrula, and had only a poor match in brittle star.

Forkhead (Fox) Gene Family

The *forkhead* gene family, defined by the winged helix-turn-helix FORKHEAD [FKH or FOX] DNA-binding domain, is found in abundance in all animal phyla, with 15 *fox* genes in the cnidarian *Nematostella* [Magie *et al* 2005] and 5 in sponges [Adell and Muller, 2004]. The urchin genome includes 22 *fox* genes, 5 of which have 2 alleles each [Tu *et al* 2006]. The set of *fox* genes found in brittle star closely corresponds with those expressed in urchin gastrula [Table 3.2], and the differences are mostly consistent with morphological differences.

Sp-FoxA, B, and *P* are all involved with endoderm formation [Hinman *et al* 2003a, Oliveri *et al* 2006, Luke *et al* 1997]. *FoxB* is also important to PMC specification [Minokawa 2004], along with *FoxO. FoxC* is seen in the coelomic pouches [Ransick *et al* 2002]. *FoxJ1* is found in the apical plate, as is *FoxQ2,* which provides a vital link between the AV and OA axes [Yaguchi *et al* 2008]. All of these were also found in brittle star. Patterning of neuroectoderm through restriction of FoxQ2 orthologs to one pole of the early embryo by Wnt signalling from the opposite pole has been described in *Saccoglossus* [Darras *et al* 2011],

Branchiostoma [Yu *et al* 2007], and even the cnidarian *Clytia hemisphaerica* [Momose *et al* 2008].

Four other *fox* genes expressed in urchin gastrula [Tu *et al* 2006] were not found in brittle star. *Sp-FoxG* is seen in oral ectoderm, while *Sp-FoxM* is ubiquitous, and acts in the cell cycle in other organisms. The absence of *FoxN2/3* and *FoxY* is consistent with the reduced presence of SMCs in brittle star. *Sp-FoxN2/3* is very briefly (but importantly) expressed in PMCs at blastula stage, and is thereafter seen only in the small micromeres and veg2 SMCs [Rho and McClay 2011], while *Sp-FoxY* is apparently urchin-specific and confined exclusively to the small micromeres.

Two *fox* genes found in brittle star likely represent heterochronic shifts: at early gastrula, *Sp-FoxJ2* is weakly seen only in SMCs, but is present throughout the urchin embryo at earlier stages, while *Sp-FoxF* appears only later in gastrulation within the coelomic pouches [Tu *et al* 2006]. In brittle stars, the coelomic pouches form a bit earlier, and are complete before gastrulation is finished. The status of two other genes is questionable. *Sp-FoxK* is very highly expressed in both primary and secondary mesenchyme, as well as in aboral ectoderm [Tu *et al* 2006]. Its presence in brittle star is in doubt, as there are only poor-quality matches to the FHA domain (also found in many other unrelated proteins) and to a short segment just C-terminal to the FKH domain. *Sp-FoxQ1* is not significantly expressed at all in urchin embryos, and had only a very poor match in brittle star.

Basic Zipper (bZip) Gene Family

Basic Zipper transcription factors have hetero- or homo-dimerization domains ("zippers") in their C-terminal ends [Vinson *et al* 2002]. There are 15 urchin bZip genes, but the precise roles of most of these remain uncertain. Nine had good matches in brittle star [Table 3.3]. Notable among these are *Sp-Jun,* which is seen in the ingressing PMCs, and *Sp-Hlf,* found in the neurogenic ectoderm [Howard-Ashby *et al* 2006b, Burke 2006].

Basic Helix-Loop-Helix (bHLH) Gene Family

The urchin genome contains 48 basic helix-loop-helix genes [Howard-Ashby *et al* 2006b]. Many bHLH genes do not operate until the later differentiation stages of development. Thus, in urchin gastrula, only 21 (44%) of these are expressed, many at a rather low level. Ten of these were also found in brittle star, plus another 2 not present in urchin or starfish gastrulae [Table 3.4].

Arnt (Ahr Nuclear Translocator) forms heterodimers with several other bHLH proteins, including Ahr and Hif1a, which were not found in brittle star, as well as with Sim, which was found. In flies and mammals, the Sim-Arnt dimer is active in neurogenesis [Kinoshita *et al* 2004].

Hes (*Hairy Enhancer of Split*) has a complex expression pattern in several regions of the urchin embryo, but its precise functions are as yet unclear [Minokawa *et al* 2004]. The paralog *HesC* is expressed throughout most of the starfish embryo with an unknown role [McCauley *et al* 2010]. In urchins, *HesC* has been co-opted to function as part of the master switch mechanism for the

derived embryonic skeleton [Revilla-i-Domingo *et al* 2007]. *Mitf* is seen in the skeletal precursors and oral ectoderm [Howard-Ashby *et al* 2006b], while *Myc* is a cell-cycle control gene expressed around the urchin blastopore [Howard-Ashby *et al* 2006b, Peter and Davidson 2011]. *Usf* (*Upstream Stimulatory Factor*) is found in the SMCs and foregut, and *E12* is present throughout the embryo [Howard-Ashby *et al* 2006b]. All of these were found in the brittle star transcripts.

Notably absent from brittle star were *Ac/Sc* (*Achaete-Scute*) which is expressed in the urchin apical plate and is very widely employed in neurogenesis [Burke *et al* 2006, Wei *et al* 2009].

Nuclear Hormone Receptor (NR) Gene Family

Nuclear receptors are ligand-activated transcription factors. This family is faster evolving than other families [Howard-Ashby *et al* 2006b], making identification of homologs more difficult. Out of 33 urchin nuclear receptor genes, only 9 had good-quality brittle star matches. As with bHLH genes, many hormone receptors are not active until much later in urchin development, and very little is known about the ligands they bind or their specific functions. Nineteen NR genes are expressed in urchin gastrula, 6 of which had good matches in brittle star [Table 3.5]. The e-values for this gene family were also considerably larger on average than for other families. One gene for which more is known is *Fxr,* which is expressed in the oral ectoderm and functions in environmental sensing and immune response [Goldstone *et al* 2006].

The three *nr1h* urchin genes are recent duplications [Howard-Ashby *et al* 2006b]. All three match to the same brittle star cDNA, but with conflicting best e-values in the forward and reverse directions. It is therefore unclear whether the duplication events occurred before or after the last common ancestor. Similarly, the two urchin *ppar* genes are a recent duplication. Only *Ppar1* is expressed in urchin gastrula, but only *Ppar2* was found in brittle star, so this duplication is likely urchin-specific. The four *nr1m* genes are also a recent expansion. Urchins express two of these at gastrula, but none were found in brittle star.

Homeobox Gene Family

The homeobox or homeodomain family is large and diverse, with many subfamilies. The *S. purpuratus* genome contains 96 homeobox genes, 44 of which are expressed at gastrula stage [Tables 3.6–3.11].

The classical *hox* cluster [Table 3.7] contains 11 genes in sea urchins, only 2 of which participate in embryogenesis, while the rest are active only in adult development [Arenas-Mena *et al* 2000]. *Sp-Hox7* is expressed in the urchin archenteron, SMCs, and oral ectoderm [Dobias *et al* 1996] but was not found in brittle star. *Sp-Hox11/13b* helps specify the vegetal pole and gut [Smith *et al* 2008] and in fact, this was the only member of the extended *hox* class found in brittle star. The three urchin *parahox* genes, *Gsx, Lox,* and *Cdx* are not active until mid- to late-gastrula [Arnone *et al* 2006], and were not seen in brittle star.

In urchins and starfish, a single *Otx* (*Orthodenticle*) gene is subject to alternative promoters and alternative splicing to yield 2 or 3 distinct proteins

which play vital roles in oral ectoderm and the vegetal plate [Li 1997, Hinman *et al* 2003a & 2003b]. The splice-variants share the homeobox and C-terminal ends, but differ in their N-terminal regions. Brittle star had matches to all 3 of these regions, with the N-terminal sequence matching best to starfish *Otxβ-b,* the ortholog of urchin *Otxβ.* The β isoform is the ancestral version, and the one more widely employed, with the α variant unique to echinoderms [Hinman *et al* 2003b]. There were no matches to the α N-terminus, but if the urchin and starfish pattern is conserved, expression of the α form is declining at gastrula stage and greatly outnumbered by transcripts of *Otxβ.*

Several other homeobox genes with well-characterized roles in urchins were also found in brittle star. *Alx1* is near the top of the PMC skeletogenic program [Ettensohn *et al* 2003 & 2007]. *Hex* and *Tgif* function a bit later in the PMCs [Howard-Ashby 2006a], as well as in Veg2 endoderm for *Tgif* [Peter 2011]. Across the animal kingdom, *Pax6* plays a central early role in neural and especially retinal specification [Vopalensky and Kozmik 2009]. *Sp-Six3* is near the top of the circuit for the neurogenic animal pole signalling center, and, as in vertebrates, excludes Wnt signalling from this area [Wei *et al* 2009, Lagutin 2003]. *Nk2.2* is found in the ectoderm [Saudemont *et al* 2010], while *Nkx2.1* is seen only in non-neurogenic cells of the apical plate [Takacs *et al* 2004]. *Hnf6* helps pattern the oral ectoderm, and later the ciliary band, as well as contributing to activation of the skeletal differentiation genes, and to *GataC* in the NSM. [Otim *et al* 2004].

Sea urchin *Pmar1* is briefly expressed only in the cleavage-stage micromeres to initiate skeletogenesis [Oliveri *et al* 2003]. Our lab has cloned brittle star *Pmar1* from genomic DNA, but never found it expressed at any developmental stage. Nor was it found in the brittle star gastrula sequences. This lends support to the idea that brittle stars acquired their embryonic skeleton independently of sea urchins through early activation of the skeletal GRN by a different gene than *Pmar1.*

C_2H_2 Zinc-Finger Gene Family

The sea urchin genome contains 377 C_2H_2 zinc-finger genes, less than a third of which are expressed during embryogenesis, and the functions of the vast majority of these remain unknown [Materna *et al* 2006a]. The C_2H_2 zinc-finger domain is extremely common in a variety of proteins. Many of these are transcription factors, but many others are not, and there is no simple way to distinguish these through their sequences alone. Proteins often contain a large number of zinc-finger domains repeated in tandem [Laity *et al* 2001], and the amino acid sequences of these domains are generally very similar between different proteins, while regions outside the domains tend to be widely divergent even among known orthologs [Knight and Shimeld, 2001]. Because of these features, my reciprocal-best-BLAST approach proved unsuccessful with zinc-finger genes, and they were not included in this study.

Other Gene Families

The sea urchin genome contains another 62 transcription factor genes belonging to a number of smaller families [Tables 3.12–3.15], half of which were found in brittle star.

Brittle star gastrula contained *SoxB1* but not *SoxB2* [Table 3.12]. In urchins, these two factors play distinct but overlapping roles in patterning the ectodermal end of the animal-vegetal axis [Kenny *et al* 2003]. *Sp-SoxD* is seen in the ingressing end of the gut, while *Sp-SoxC* is seen in the same area, plus in several distinct patches of the ectoderm [Howard-Ashby *et al* 2006a], but was not found in brittle star. Neither was *SoxE,* which is seen in the urchin SMCs and later in the left coelomic pouch [Juliano *et al* 2006].

Smad proteins are the transcription factors at the ends of TGF-β, Activin, and BMP signalling pathways [Itoh *et al* 2000]. Sea urchins have 4 *smad* genes, one homolog for each of the vertebrate *smad* types, and brittle star had matches of 3 of these [Table 3.13]. Sp-Nodal signals through Sp-Smad2/3 in the oral ectoderm to limit neurogenesis to the apical plate [Yaguchi *et al* 2006]. *Smad2/3* had no reciprocal matches to brittle star, but *Smad1/5/8* had reciprocal matches to five separate regions covering both the MH1 and MH2 domains. It is possible that at least one of these sequences actually represents a transcript of *Smad2/3* in a region with greater sequence divergence.

Of the 6 urchin members of the *tbox* gene family, 3 are expressed at gastrula [Table 3.14]. One of these was found in brittle star: *Tbox2-3*, which contributes to morphogenic cell movements in the aboral regions of all three

urchin germ layers [Gross *et al* 2003]. *Tbr* (*Tbrain*) helps specify starfish endoderm, but in urchins has been co-opted into skeletogenesis [Hinman 2003a, 2007b]. Brittle star had only a very poor match to starfish *Tbr,* and none at all to urchin *Tbr.* Attempts at PCR amplification from various developmental stages were also unsuccessful, and I conclude that *Tbr* may have been lost in brittle stars, or at least not expressed during embryogenesis.

Similarly, no *gata* genes were found either in the brittle star sequences or through PCR [Table 3.15]. *Sp-GataC* is active in pigment cells, [Ransick and Davidson 2006] which brittle stars lack. *GataE* plays important conserved roles in urchin and starfish endoderm [Hinman and Davidson 2003c, Hinman *et al* 2003a, Kiyama and Klein 2007]. *Gata* factors are relatively variable even within the DNA-binding domain, and radically divergent elsewhere [Lowry and Atchley 200]. This, in conjunction with the short pyrosequencing reads, could potentially mask the presence of brittle star *gata* factors.

Several other transcription factors of note were found in brittle star [Table 3.15]. Cytoskeletal Actin 3a (*Sp-CyIIIa*) is activated in the aboral ectoderm by *Sp-Runt1* and is inhibited elsewhere by *Sp-Myb* [Coffman *et al* 1996 & 1997]. Sea urchin Gcm acts downstream of Delta-Notch signalling to specify pigment cells [Ransick *et al* 2002, Ransick and Davidson 2006]. Brittle stars lack pigment cells, so Gcm's role here is unknown. *Deadringer* (*Dri*) performs dual roles in the PMCs and the oral ectoderm [Amore *et al* 2003], Groucho is a ubiquitous cofactor which partners with Six3 in the animal hemisphere to compete with β-catenin for TCF

binding sites [Range *et al* 2005, Howard-Ashby *et al* 2006], thus excluding vegetalizing signals from the ectoderm.

CONCLUSIONS

The brittle star gastrula-stage embryo expresses genes from all of the transcription factor families. Roughly half of the transcription factor genes expressed at gastrula stage in the sea urchins *Strongylocentrotus purpuratus* and *Lytechinus variegatus* and in the starfish *Patiria miniata* were also found in brittle star gastrula.

These results are consistent with previous studies. BLAST comparisons of the entire *S. purpuratus* genome to the genomes of human, mouse, tunicate, fly, nematode and cnidarian found reciprocal-best-BLAST matches for between 15% and 25% of all sea urchin genes, while a comparison of the same sort between human and mouse genomes found homologs amounting of 58% of all mouse genes and 67% of all human genes [Materna *et al* 2006b]. Given that the degree of relatedness between echinoderm classes is intermediate between these two examples, 49%-57% reciprocal-best-BLAST matches among echinoderm gastrula transcription factors is in line with expectations.

Several other lines of evidence also suggest that the pyrosequencing captured the vast majority of the brittle star genes expressed at gastrula stage. The pyrosequencing results are entirely consistent with previous results from attempts at PCR amplification. *Alx1, Dri, Ets1*, and *Erg* were found at gastrula

stage by both methods, whereas *Pmar1,Tbr, GataC,* and *GataE* were not found in cDNA from any stage.

The brittle star pyrosequences also contained matches for all or very nearly all the proteins involved in several very highly conserved pathways, including mitosis, glycolysis, DNA replication, and ribosomal proteins. The protein sequences of transcription factors tend to be highly conserved in the important domains, but quite variable elsewhere. They are also sometimes redeployed into different temporal expression patterns in the course of evolution, unlike the critical housekeeping genes mentioned. On the other hand, housekeeping genes are by their nature very highly expressed, while transcription factors are typically found at only a few hundred transcripts per embryo. So it is entirely possible that some transcription factors were simply missed by the sequencing. However, the fact that so many of the most important transcription factors were found suggests that most were indeed captured by the sequencing.

Finally, the overall patterns of genes found and not found in brittle star are consistent in several aspects with the morphological differences between brittle stars and urchins, such as the presence of nearly all the skeleton genes, and the absence of many pigment cell genes. As described in Chapter 2, some of the visible events in the brittle star embryo are delayed or proceed at a slower pace than in urchins. So it is possible that some of the molecular events of specification in the Veg1 endoderm and aboral ectoderm are delayed as well.

While only very tentative conclusions about the brittle star developmental gene regulatory network can be drawn at this point, the evidence presented here

can be used to infer some of the larger features of this network. Many of the "kernels" [Davidson and Erwin 2006] central to echinoderm development appear to be intact. The majority of genes involved in early vegetal plate endomesoderm specification, apical pole neuroectoderm specification, and skeletogenesis in both sea urchins and starfish were found in brittle star. Given the relative degrees of conservation and innovation in these GRN circuits in these other echinoderms [Ettensohn *et al* 2007, Hinman and Davidson 2007a, Gao and Davidson 2008, McCauley *et al* 2010], it seems highly likely that brittle stars employ the same general input/output logic of these GRN kernels, as well as many of the specific interactions of the regulatory "wiring".

Key for Tables

Ow = *Ophiocoma wendtii* (brittle star) transcriptome
Sp = *Strongylocentrotus purpuratus* (purple sea urchin) proteome
Lv = *Lytechinus variegatus* (green sea urchin) transcriptome
Pm = *Patiria miniata* (starfish) transcriptome
GB = GenBank reference proteins

	high-confidence homologs (1E-09 or better)
	possible homologs (1E-06 to 2E-09)
	poor matches (worse than 1E-06)
*	conflicting results
—	no match found
—	not present in available gastrula transcriptome data
	present in genome but not in available gastrula transcriptome data
	no expression data available

Table 3.1: Ets Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
ElfA	—	—	—	—	—	—	—			
ElfB	FKK7YRX02MPLQR	2E-04	3E-04	—	—	—	—	Sp-ElfB	5E-03	2E-04
Elk	—	—	—	—	—	—	—			
Erf	—	—	—	—	—	—	—			
Erg	FKK7YRX02OL9IR	2E-16	4E-16	3E-14	1E-14	3E-24	*	Sp-Erg	1E-16	2E-16
Ese	—	—	—	—	—	—	—			
Ets1/2	FKK7YRX02MLP4M	8E-25	3E-25	5E-31	7E-32	—	—	Sp-Ets1/2	1E-27	8E-25
	FKK7YRX02NL359	4E-21	7E-21	3E-25	5E-26	—	—	Sp-Ets1/2	2E-22	4E-21
Ets4	contig13356	4E-48	4E-55	4E-59	9E-66	8E-46	5E-44	Sp-Ets4	1E-63	4E-48
Gabp	FKK7YRX02NWFJW	3E-36	4E-36	—	—	4E-16	*	*Saccoglossus kowalevskii* Gabp α subunit	2E-45	2E-31
Pea	contig06761	3E-09	5E-09	2E-12	3E-12	2E-06	6E-07	—	—	—
Pu1	—	—	—	—	—	—	—			
Tel	contig00179	1E-09	5E-09	—	—	—	—	*Drosophila melanogaster* Pointed isoform D	7E-59	2E-61

* FKK7YRX02OL9IR matches best to *Pm-Gabp*, and FKK7YRX02NWFJW matches best to *Pm-Erg*.

Table 3.2: Forkhead Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
FoxA	contig07901	7E-15	1E-15	4E-05	4E-06	2E-25	1E-25	Sp-FoxA	3E-13	7E-15
FoxABL	—	—	—	—	—	—	—			
FoxB	FKK7YRX02K5B9K	2E-12	1E-10	—	—	1E-21	8E-14	Sp-Fkh1 (Sp-FoxB)	2E-07	2E-12
FoxC	FKK7YRX02ON70S	8E-08	1E-07	—	—	—	—	Sp-FoxC	3E-05	8E-08
	FKK7YRX02ONMQ2	4E-07	2E-05	—	—	—	—	Sp-FoxC	9E-03	4E-07
FoxD	—	—	—	—	—	—	—			
FoxF	FKK7YRX02OCVUU	1E-38	4E-39	—	—	—	—	Sp-FoxF	6E-37	1E-38
FoxG	—	—	—	—	—	—	—			
FoxI	—	—	—	—	—	—	—			
FoxJ1	FKK7YRX02OIQ9X	4E-13	7E-13	6E-14	1E-14	—	—	Sp-FoxJ1	2E-10	4E-13
FoxJ2	FKK7YRX02O2DG3	5E-11	9E-11	2E-04	9E-06	—	—	Sp-FoxJ2	2E-08	5E-11
FoxK	FKK7YRX02K8XHX	4E-04	5E-04	1E-05	5E-05	8E-05	1E-05	*Branchiostoma floridae* hypoth. Fox protein	6E-02	1E-04
	FKK7YRX02MUOT4	3E-01	7E-05	1E-05	6E-06	4E-05	4E-07	Sp-FoxK	1E-03	3E-01
FoxL1	—	—	—	—	—	—	—			
FoxL2	—	—	—	—	—	—	—			
FoxM	—	—	—	—	—	—	—			
FoxN1/4	FKK7YRX02LTME4	1E-07	1E-06	1E-08	2E-09	8E-10	2E-09	Sp-FoxN1/4	3E-04	1E-07
FoxN2/3	—	—	—	—	—	—	—			
FoxO	contig01589	1E-43	4E-44	—	—	—	—	Sp-FoxO	6E-42	1E-43
FoxP	FKK7YRX02MXRWU	5E-32	7E-32	—	—	—	—	*Ailuropoda melanoleuca* FoxP1-like isoform 2	1E-29	3E-32
FoxQ1	FKK7YRX02NUCL3	1E-01	2E-01	—	—	—	—	—	—	—
FoxQ2	FKK7YRX02NDRLV	7E-32	1E-31	3E-38	7E-39	3E-40	4E-41	Sp-FoxQ2	3E-29	7E-32
FoxY	—	—	—	—	—	—	—			

Table 3.3: Basic Zipper Transcription Factor Family

Gene	Ow ID	Sp vs Ow	Ow vs Sp	Lv vs Ow	Ow vs Lv	Pm vs Ow	Ow vs Pm	Ow vs GB	Ow vs GB	GB vs Ow
Atf2	FKK7YRX02LHAEC	7E-06	4E-03	1E+00	—	3E-08	2E-07	*Sus scrofa* Atf2	3E-05	3E-07
Atf6b	contig14022	4E-07	4E-07	—	—	—	—	—	—	—
Creb	FKK7YRX02M1WF7	3E-17	5E-17	1E-10	3E-11	4E-31	5E-31	Sp-Notch2_7	7E-18	3E-17
Creb3L1	contig12373	—	—	3E-27	8E-31	1E-34	3E-37	*Saccoglossus kowalevskii* CREB-like	3E-08	1E-27
	FKK7YRX02OTWFS	6E-11	5E-08	4E-13	8E-19	1E-15	3E-16	Sp-Creb3L1	5E-12	6E-11
Creb3L3	contig01994 *	4E-07	3E-09	2E-06	*	8E-04	*	*Saccoglossus kowalevskii* Atf6β-like *	1.E-26	1E-28
Fos	FKK7YRX02N3R6W	5E-04	8E-04	—	—	1E-05	1E-05	—	—	—
Fra2	contig02088	2E-20	1E-21	2E-32	6E-32	1E-20	9E-23	Sp-Fra2	1E-40	2E-32
Giant	—	—	—	—	—	—	—			
Hlf	FKK7YRX02NVHQH	4E-20	2E-11	3E-24	2E-14	2E-26	1E-16	*Drosophila melanogaster* PAR-domain protein 1, isoform F	2E-37	5E-33
Jun	contig13321	1E-53	2E-47	4E-58	4E-53	2E-58	2E-50	Sp-Jun	4E-45	1E-53
Lztf1	FKK7YRX02MHLAX	1E-30	2E-30	—	—	2E-33	3E-33	Sp-Lztf1	3E-36	1E-30
Maf	contig00697	8E-18	3E-17	—	—	8E-19	2E-18	*Drosophila melanogaster* Maf-S	5E-91	7E-73
Nfe2	FKK7YRX02KPQGA	4E-23	6E-23	—	—	1E-20	3E-25	Sp-Nfe2	3E-27	4E-23
	FKK7YRX02K2KLS	1E-15	9E-11	—	—	1E-08	2E-09	*Saccoglossus kowalevskii* Nfe2-like	5E-11	5E-18
NflL3	—	—	—	—	—	—	—			
Xbp1	—	—	—	—	—	—	—			

* contig01994 matches to different genes in different organisms

Table 3.4: Basic Helix-Loop-Helix Transcription Factor Family

Gene	Ow ID	Sp vs Ow	Ow vs Sp	Lv vs Ow	Ow vs Lv	Pm vs Ow	Ow vs Pm	Ow vs GB	Ow vs GB	GB vs Ow
Ac/Sc	—	—	—	—	—	—	—			
Ac/Sc3	—	—	—	—	—	—	—			
Ahr_1	—	—	—	—	—	—	—			
Ap4	—	—	—	—	—	—	—			
Arnt	contig06184	5E-59	1E-57	—	—	—	—	*Drosophila melanogaster* Cycle	0.0	1E-157
Ato, Hath6	—	—	—	—	—	—	—			
AtoL1	—	—	—	—	—	—	—			
AtoL2	—	—	—	—	—	—	—			
Beta3b	—	—	—	—	—	—	—			
Bmal	—	—	—	—	—	—	—			
Clock	—	—	—	—	—	—	—			
E12	contig02328	2E-18	4E-06	7E-21	2E-13	—	—	Sp-E12	4E-11	2E-18
Ebf3	FKK7YRX02NS71F	3E-08	4E-08	—	—	—	—	*Saccoglossus kowalevskii* Ebf3	8E-08	9E-09
Hairy2/4, Hes	FKK7YRX02NKH5N	2E-07	2E-04	2E-08	2E-07	8E-07	2E-07	*Saccoglossus kowalevskii* Hes1	1E-04	3E-09
Hairy, HesB	—	—	—	—	—	—	—			
Hand	—	—	—	—	—	—	—			
HesC	FKK7YRX02NRDTB	6E-22	1E-21	1E-26	3E-27	1E-22	3E-23	Sp-HesC	2E-24	6E-22
Hey	—	—	—	—	—	—	—			
Hey4	—	—	—	—	—	—	—			

continued…

Table 3.4: Basic Helix-Loop-Helix Transcription Factor Family (continued)

Hif1a	—	—	—	—	—	—	—			
Id	—	—	—	—	—	—	—			
Mad	—	—	—	—	—	—	—			
Max	contig01703	4E-28	7E-28	—	—	5E-33	6E-32	*Drosophila melanogaster* Max	3E-88	5E-72
Mist	—	—	—	—	—	—	—			
Mitf	FKK7YRX02OWEWT	1E-07	2E-07	5E-09	2E-09	—	—	*Pediculus humanus corporis* Mitf	3E-08	5E-09
Mlx	—	—	—	—	—	—	—			
Mlx/lp	contig05473	8E-22	8E-22	—	—	—	—	Sp-Mlx/lp	2E-22	8E-22
Mnt	—	—	—	—	—	—	—			
Myc	FKK7YRX02NV3UF	2E-12	4E-11	4E-10	2E-10	8E-19	4E-19	Sp-Myc	2E-11	2E-12
MyoD	—	—	—	—	—	—	—			
MyoD2	—	—	—	—	—	—	—			
MyoD3	—	—	—	—	—	—	—			
Myor2	—	—	—	—	—	—	—			
Myor4	—	—	—	—	—	—	—			
Ncsl	—	—	—	—	—	—	—			
NeuroD1	—	—	—	—	—	—	—			
Ngn	—	—	—	—	—	—	—			
NtL, Nato3	—	—	—	—	—	—	—			
Nxf	—	—	—	—	—	—	—			
Olig3	—	—	—	—	—	—	—			
Par1	—	—	—	—	—	—	—			
Ptf1a	—	—	—	—	—	—	—			
Sage	—	—	—	—	—	—	—			
Scl	—	—	—	—	—	—	—			
Sim_1	FKK7YRX02N96LL	1E-14	2E-14	—	—	—	—	*Schistosoma mansoni* Single-minded	4E-15	5E-13
Srebp, bHLHB1	FKK7YRX02M90VR	5E-08	7E-08	7E-10	2E-10	2E-14	1E-14	*Danio rerio* Strebp1	6E-20	4E-19
Trh	—	—	—	—	—	—	—			
Usf	FKK7YRX02OW5VD	9E-32	1E-31	4E-37	9E-38	—	—	*Nomascus leucogenys* Usf2	4E-38	4E-32

Table 3.5: Nuclear Hormone Receptor Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Couptf1	—	—	—	—	—	—	—			
Dsf	—	—	—	—	—	—	—			
e78, e78a	FKK7YRX02NT6JT	1E-15	2E-15	—	—	—	—	Sp-e78	6E-16	1E-15
e78b, Rara	—	—	—	—	—	—	—			
Err	—	—	—	—	—	—	—			
Fax1	FKK7YRX02LQFET	3E-02	4E-02	—	—	—	—	—	—	—
Fxr	FKK7YRX02KYR42	5E-26	9E-34	—	—	6E-43	5E-43	Sp-Fxr	4E-38	5E-26
Gcnf1	FKK7YRX02KLFS8	3E-14	1E-14	—	—	5E-06	8E-10	Sp-Gcnf1	4E-15	3E-14
	FKK7YRX02MPSM0	1E-13	2E-14	—	—	9E-13	2E-16	Sp-Gcnf1	7E-11	1E-13
Grf	—	—	—	—	—	—	—			
Hnf4	FKK7YRX02L7ADE	1E-22	2E-22	—	—	—	—	*Drosophila melanogaster* Hnf4 isoform C	4E-29	3E-24
Nr1AB, Nr1x	FKK7YRX02NUJ5Y	2E-12	3E-12	—	—	—	—	Sp-Nr1x	1E-10	2E-12
Nr1h6	**FKK7YRX02K0S55**	1E-10	9E-16	—	—	—	—	Sp-Nr1h6	2E-16	1E-10
Nr1h6b	**FKK7YRX02K0S55**	5E-14	**9E-16**	—	—	—	—	**Sp-Nr1h6**	**2E-16**	**1E-10**
Nr1h6c	**FKK7YRX02K0S55**	**1E-13**	**9E-16**	—	—	—	—	**Sp-Nr1h6**	**2E-16**	**1E-10**
Nr1m1	—	—	—	—	—	—	—			
Nr1m2	—	—	—	—	—	—	—			
Nr1m3	—	—	—	—	—	—	—			
Nr1m4	—	—	—	—	—	—	—			
Nr2C	—	—	—	—	—	—	—			
Nr2e6	—	—	—	—	—	—	—			
Nr5a, FtzF	—	—	—	—	—	—	—			
Nurr1	—	—	—	—	—	—	—			
Pnr	—	—	—	—	—	—	—			
Ppar1	—	—	—	—	—	—	—			
Ppar2	FKK7YRX02MESJR	3E-16	5E-16	—	—	—	—	Sp-Ppar2	1E-16	3E-16
Rar	—	—	—	—	—	—	—			
Reverb	—	—	—	—	—	—	—			
Rora	—	—	—	—	—	—	—			
Rxr	—	—	—	—	—	—	—			
Shr2, Tr2.4	FKK7YRX02NRNVZ	2E-04	3E-04	—	—	—	—	***Ixodes scapularis* Retinoid X receptor ***	4E-05	**9E-10 ***
Thr	—	—	—	—	—	—	—			
Thrb	FKK7YRX02MWCT3	1E-02	2E-02	—	—	—	—	***Nematostella vectensis* predicted protein ***	4E-02	**6E-20 ***
Tll	—	—	—	—	—	—	—			

* GenBank sequences match back to FKK7YRX02L7ADE (Hnf4)

Table 3.6: Paired Homeobox Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Alx1	FKK7YRX02LUU3G	2E-18	5E-15	—	—	—	—	Sp-Alx1	2E-15	2E-18
Alx4	—	—	—	—	—	—	—			
Arx	—	—	—	—	—	—	—			
Arxl	—	—	—	—	—	—	—			
Arxl2	—	—	—	—	—	—	—			
Chx10	—	—	—	—	—	—	—			
Eyg	—	—	—	—	—	—	—			
Gsc	—	—	—	—	—	—	—			
Hbn	—	—	—	—	—	—	—			
Mbx1	—	—	—	—	—	—	—			
Otp	—	—	—							
Otx	contig6115	8E-06	2E-04	5E-07	9E-08	2E-09	6E-09	—	—	—
	FKK7YRX02LKG0R	2E-09	2E-07	2E-12	9E-14	2E-20	4E-22	*Cavia porcellus* OTX2-like isoform 1	4E-08	1E-08
	FKK7YRX02MPQH8	9E-07	3E-10	4E-09	6E-09	—	—	Sp-Otx	8E-08	8E-09
Pax1-9	FKK7YRX02L240Z	3E-14	2E-04	—	—	—	—	*Monodelphis domestica* cyclic nucleotide-gated cation channel alpha-3-like	7E-08	*
Pax2-5-8	FKK7YRX02NS6L1	5E-09	8E-09	—	—	—	—	*Saccoglossus kowalevskii* Pax2/5/8	7E-04	**
Pax4L	FKK7YRX02MV38W	6E-15	1E-14	—	—	—	—	*Saccoglossus kowalevskii* Pax6	2E-08	**
Pax6	contig01219	7E-29	1E-31	—	—	—	—	*Drosophila melanogaster* twin of eyeless	2E-153	1E-156
	FKK7YRX02NN4NX	2E-17	3E-23	—	—	4E-31	2E-30	*Saccoglossus kowalevskii* Pax6	8E-29	**
PaxA	—	—	—	—	—	—	—			
PaxB	contig04061	8E-63	8E-62	2E-73	1E-73	8E-54	4E-54	Sp-PaxB	1E-59	8E-63
PaxC	—	—	—	—	—	—	—			
Pbx_1	FKK7YRX02M3FZH	2E-32	4E-32	1E-39	7E-40	—	—	Sp-Pbx_1	9E-30	2E-32
Phb1	FKK7YRX02MVHDZ	7E-17	4E-18	2E-19	4E-22	4E-16	9E-18	Sp-Phb1	1E-15	7E-17
Phb2	contig06985	4E-17	2E-20	—	—	—	—	*Culex quinquefasciatus* consrv hyp protein	6E-31	3E-25
Phox2	FKK7YRX02K7TKF	2E-04	3E-04	—	—	—	—	—	—	—
Pitx1	—	—	—	—	—	—	—			
Pitx2	—	—	—	—	—	—	—			
Pitx3	—	—	—	—	—	—	—			
Pmar1	—	—	—	—	—	—	—			
Prx	—	—	—	—	—	—	—			
Rx	—	—	—	—	—	—	—			
Shox	—	—	—	—	—	—	—			
Unc4.1	—	—	—	—	—	—	—			

* GenBank matches back to contig03546

** GenBank matches back to contig04061 (PaxB) much better

Table 3.7: Extended Hox Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Barh1	—	—	—	—	—	—	—			
Barx	—	—	—	—	—	—	—			
Cdx	—	—	—	—	—	—	—			
Emx	—	—	—	—	—	—	—			
Eve	—	—	—	—	—	—	—			
Gbx	—	—	—	—	—	—	—			
Gsx	—	—	—	—	—	—	—			
Hb9	—	—	—	—	—	—	—			
Hox1	—	—	—	—	—	—	—			
Hox2	—	—	—	—	—	—	—			
Hox3	—	—	—	—	—	—	—			
Hox5	—	—	—	—	—	—	—			
Hox6	—	—	—	—	—	—	—			
Hox7	—	—	—	—	—	—	—			
Hox8	—	—	—	—	—	—	—			
Hox9/10	—	—	—	—	—	—	—			
Hox11/13a	—	—	—	—	—	—	—			
Hox11/13b	FKK7YRX02NUT6F	3E-17	6E-17	9E-16	7E-16	2E-12	2E-14	Sp-Hox11/13b	5E-18	3E-17
Hox11/13c	—	—	—	—	—	—	—			
Lox	—	—	—	—	—	—	—			
Mox	—	—	—	—	—	—	—			
Not	—	—	—	—	—	—	—			

Table 3.8: Atypical Homeobox Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Hnf1	—	—	—	—	—	—	—			
IrxA	—	—	—	—	—	—	—			
IrxB	FKK7YRX02LHGCF	2E-05	1E-08	—	—	6E-06	5E-06	Sp-IrxB	6E-08	2E-05
Meis	FKK7YRX02K5EES	5E-38	8E-38	—	—	—	—	Sp-Meis	1E-43	5E-38
Pknox	FKK7YRX02ND68E	6E-10	2E-16	6E-12	2E-20	6E-09	7E-09	Sp-Pknox	4E-17	6E-10
Prox1	—	—	—	—	—	—	—			
Six1/2	—	—	—	—	—	—	—			
Six3	contig00851	2E-75	2E-72	4E-65	2E-64	3E-95	4E-95	Sp-Six3	1E-88	2E-75
Six4	—	—	—	—	—	—	—			
Tgif	FKK7YRX02KZVQF	1E-27	9E-29	—	—	2E-36	3E-37	*Ixodes scapularis* TGIF	2E-33	4E-27

Table 3.9: Hox/Lim Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Awh	—	—	—	—	—	—	—			
Isl	—	—	—	—	—	—	—			
Lhx2	FKK7YRX02MZFIC	3E-10	4E-11	—	—	—	—	*Mus musculus* Lhx2	7E-14	3E-12
Lhx3-4	FKK7YRX02MWWP3	5E-03	9E-03	—	—	—	—	—	—	—
Lim1	—	—	—	—	—	—	—			
Lmx1	—	—	—	—	—	—	—			

Table 3.10: Nk Homeobox Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Hex	FKK7YRX02O2X0B	1E-18	6E-18	—	—	1E-19	9E-20	Sp-Hex	2E-19	1E-18
Hlx	—	—	—	—	—	—	—			
Lbx	—	—	—	—	—	—	—			
Nk1	—	—	—	—	—	—	—			
Nk2.2	contig03829	2E-49	2E-58	6E-65	1E-75	2E-68	5E-79	Sp-Nk2.2	6E-55	2E-49
Nk7	—	—	—	—	—	—	—			
Nkx2.1	FKK7YRX02OFE4V	4E-22	7E-22	—	—	2E-33	6E-34	*Saccoglossus kowalevskii* NK2.1 *	1E-30	1E-33
	FKK7YRX02K456Z	5E-20	8E-20			3E-34	7E-35	*Saccoglossus kowalevskii* NK2.1 *	3E-28	1E-33
Nkx2.5	—	—	—	—	—	—	—			
Nkx3.2	—	—	—	—	—	—	—			
Nkx6.1	—	—	—	—	—	—	—			

* *Saccoglossus kowalevskii* NK2.1 matches back to contig03829 (Nk2.2)

Table 3.11: Dl, Cut, Pou, Barx, Zinc Finger Homeobox Transcription Factor Families

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Atbf1	FKK7YRX02M98L4	1E-30	1E-30	—	—	—	—	Sp-Atbf1	1E-33	1E-30
Brn1-2-4	FKK7YRX02NG6O4	1E-07	8E-12	—	—	1E-16	9E-17	Sp=Brn1-2-4	9E-17	1E-07
Cutl	FKK7YRX02N0HRH	4E-34	4E-34	—	—	—	—	*Branchiostoma floridae* hyp protein	9E-39	6E-33
Dlx	—	—	—	—	—	—	—			
Engrailed	—	—	—	—	—	—	—			
Hnf6	contig01249	1E-45	6E-56	3E-38	3E-52	6E-23	3E-51	Sp-Hnf6	1E-64	1E-45
Lass6	contig01288	5E-55	6E-51	—	—	2E-11	5E-13	*Drosophila melanogaster* schlank isoform B	1E-146	1E-125
Msx	—	—	—	—	—	—	—			
Msxl	FKK7YRX02NW07M	5E-17	1E-03	—	—	1E-11	2E-10	*Saccoglossus kowalevskii* Msxl	8E-03	7E-11
Oct1.2	—	—	—	—	—	—	—			
Pou4f2	—	—	—	—	—	—	—			
Pou6	—	—	—	—	—	—	—			
Rough	—	—	—	—	—	—	—			
Sip, Smad_ip	FKK7YRX02LBCB0	3E-23	5E-24	—	—	—	—	Sp-Smad_ip	6E-26	3E-23

Table 3.12: Sox/Hmg Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Bbx	FKK7YRX02OF6RJ	—	—	5E-02	4E-02	—	—	—	—	—
Cic	—	—	—	—	—	—	—			
Lef1	—	—	—	—	—	—	—			
SoxB1	contig13264	3E-70	3E-83	7E-70	4E-88	2E-89	1E-82	Sp-SoxB1	2E-91	3E-70
SoxB2	—	—	—	—	—	—	—			
SoxC	—	—	—	—	—	—	—			
SoxD1	FKK7YRX02L0FKT	2E-29	4E-29	—	—	—	—	Sp-SoxD1	1E-32	2E-29
	FKK7YRX02L6YD9	2E-16	4E-16	—	—	—	—	Sp-SoxD1	3E-16	6E-16
	contig06454	4E-11	2E-11	—	—	—	—	Sp-SoxD1	9E-11	3E-11
SoxE	—	—	—	—	—	—	—			
SoxF	—	—	—	—	—	—	—			
SoxH	—	—	—	—	—	—	—			

Table 3.13: Smad Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Smad1/5/8	contig10543	1E-50	5E-36	5E-49	1E-59	4E-58	1E-64	*Saccoglossus kowalevskii* Smad1/5	3E-50	1E-47
	FKK7YRX02L3QMQ	2E-41	4E-41	1E-48	2E-49	4E-50	7E-51	Sp-Smad1/5/8	6E-48	2E-41
	contig06372	5E-33	2E-32	2E-34	3E-35	3E-38	2E-38	Sp-Smad1/5/8	1E-35	5E-33
	FKK7YRX02MYLR3	4E-32	7E-32	6E-34	2E-34	2E-43	8E-44	Sp-Smad1/5/8	3E-36	4E-32
	FKK7YRX02KKZCA	3E-15	2E-16	1E-20	3E-21	2E-21	5E-23	*Branchiostoma floridae* Smad1	6E-19	2E-15
Smad2/3	—	—	—	—	—	—	—			
Smad4	FKK7YRX02KKDPL	7E-27	1E-26	—	—	—	—	Sp-Smad4	2E-19	7E-27
Smad6/7	FKK7YRX02OAVFY	5E-19	6E-22	2E-19	4E-20	9E-32	2E-32	Sp-Smad6/7	2E-16	5E-19

Table 3.14: Tbox Transcription Factor Family

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Bra	—	—	—	—	—	—	—			
Tbr	FKK7YRX02LPLLP	—	—	—	—	2E-01	1E-01	—	—	—
Tbx1	—	—	—	—	—	—	—			
Tbx2-3	FKK7YRX02OWR5N	2E-14	2E-14	2E-14	2E-15	1E-14	2E-15	Sp-Tbx2-3	3E-10	2E-14
Tbx6/16	—	—	—	—	—	—	—			
Tbx20	—	—	—	—	—	—	—			

Table 3.15: Other Transcription Factor Families

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Af9	FKK7YRX02N4DEO	2E-35	4E-35	—	—	—	—	***Saccoglossus kowlevskii*** **MII**	**4E-31**	**5E-39**
Ap2 (AP2)	—	—	—	—	—	—	—			
Ash1 (trxG)	FKK7YRX02NZPR1	2E-14	1E-14	—	—	3E-10	7E-11	Sp-Ash1	6E-15	2E-14
Ash2 (trxG)	FKK7YRX02LIF4B	9E-25	3E-14	—	—	3E-27	2E-27	*Drosophila yakuba* GE23503	6E-36	2E-38
	FKK7YRX02LH8HS	3E-24	5E-24	—	—	4E-29	2E-29	Sp-Ash2	2E-16	5E-24
	FKK7YRX02MFKC3	3E-24	5E-24	—	—	1E-28	1E-28	Sp-Ash2	5E-26	5E-24
Cp2 (CP2)	contig05027	1E-07	2E-06	—	—	—	—	*Drosophila melanogaster* Gemini isoform D	1E-81	6E-81
Dac (Ski-Sno)	FKK7YRX02NQL24	4E-39	8E-39	—	—	1E-46	1E-46	*Ixodes scapularis* Dachshund	5E-35	2E-38

continued

Table 3.15: Other Transcription Factor Families *(continued)*

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Dmtf (myb)	—	—	—	—	—	—	—			
Dp1 (E2F)	contig02393	—	—	5E-02	1E-01	—	—	—	—	—
Dri (bright)	FKK7YRX02NU4PZ	3E-34	5E-34	2E-39	4E-40	5E-41	2E-41	Sp-Dri	2E-38	3E-34
E2f3 (E2F)	**contig06767 ***	2E-19	**6E-20**	—	—	—	—	*Drosophila melanogaster* E2f2	8E-110	3E-90
E2f4 (E2F)	**contig06767 ***	**4E-16**	6E-20	3E-06	1E-04	—	—	*Drosophila melanogaster* E2f2	8E-110	3E-90
	FKK7YRX02K0L30	3E-13	6E-13	—	—	8E-15	3E-15	*Saccoglossus kowalevskii* E2f4	9E-12	4E-18
Enz1 (pcg)	FKK7YRX02M7OI0	2E-32	3E-32	1E-24	3E-24	6E-42	2E-42	*Aedes aegypti* Enhancer of zeste, Ezh	1E-39	1E-35
Enz2 (pcg)	—	—	—	—	—	—	—			
GataC	—	—	—	—	—	—	—			
GataE	—	—	—	—	—	—	—			
Gcm (gcm)	FKK7YRX02N0PUC	5E-31	8E-31	4E-36	3E-37	5E-36	4E-36	*Saccoglossus kowalevskii* Gcm	5E-20	4E-31
Gro, Groucho	FKK7YRX02MS79L	5E-18	7E-18	7E-21	1E-21	6E-22	6E-22	*Ixodes scapularis* Groucho	3E-19	2E-18
Irf1 (IRF)	—	—	—	—	—	—	—			
Irf4 (IRF)	—	—	—	—	—	—	—			
L3mbt (pcg)	FKK7YRX02MBZBN	8E-34	1E-33	5E-40	2E-40	—	—	Sp-L3mbt_1	4E-37	8E-34
Ldb2 (lim)	FKK7YRX02N36PD	—	—	1E-02	1E-02	—	—	—	—	—
Lmo2 (lim)	—	—	—	—	—	—	—			
Lmo4 (lim)	contig01552	5E-29	2E-30	1E-11	5E-11	—	—	*Drosophila melanogaster* CG5708 isoform A	4E-120	2E-96
	FKK7YRX02MBU2P	1E-36	2E-10	1E-10	4E-11	—	—	*Drosophila sechellia* GM11938	2E-43	1E-36
Lmpt, Fhl2 (lim)	—	—	—	—	—	—	—			
Mef2 (mads)	FKK7YRX02OTQI6	4E-14	6E-14	—	—	4E-17	5E-17	*Hydra magnipapillata* Mef2	4E-16	3E-14
Mll3 (trxG)	FKK7YRX02LW0TU	—	—	—	—	4E-15	2E-15	*Ixodes scapularis* Mll	3E-33	3E-28
Mta1 (myb)	FKK7YRX02LS90G	—	—	5E-09	2E-09	4E-06	2E-06	*Saccoglossus kowalevskii* Mta1	1E-26	1E-24
Myb (myb)	FKK7YRX02LBET9	3E-09	5E-09	—	—	—	—	*Drosophila melanogaster* Myb isoform A	5E-38	4E-41
NfiA (NFI)	FKK7YRX02N238H	2E-07	2E-08	—	—	—	—	Sp-NfiA	2E-01	2E-07
NfkB (NFI)	FKK7YRX02M7BNC	3E-11	5E-11	—	—	3E-29	1E-29	Sp-NfkB	1E-21	3E-11
P3a2	contig02928	8E-49	1E-44	1E-50	1E-49	—	—	Sp-P3a2	3E-45	8E-49
Pric, Prkl2 (lim)	FKK7YRX02OM5EK	5E-05	6E-04	—	—	—	—	—	—	—
Rfx3	FKK7YRX02MU6TJ	—	—	6E-04	1E-10	—	—	*Saccoglossus kowalevskii* Rfx	1E-30	3E-28
Runt1	FKK7YRX02MVIHC	6E-08	8E-04	9E-10	2E-11	3E-09	6E-11	*Danio rerio* Runt1	8E-08	2E-10
	FKK7YRX02KKU0C	8E-06	3E-05	6E-13	4E-08	5E-11	3E-08	Sp-Runt1	6E-04	8E-06
Runt2, Runx1	—	—	—	—	—	—	—			
Scml1 (pcg)	FKK7YRX02MIB2G	3E-18	6E-18	—	—	—	—	*Saccoglossus kowalevskii* consrv hyp protein	6E-35	4E-31
Shrls, Su(H) (IPT)	FKK7YRX02ONPEM	5E-23	7E-23	2E-39	1E-39	—	—	*Drosophila melanogaster* Shrls	5E-24	1E-38
	FKK7YRX02KT78U	1E-32	1E-26	2E-32	7E-33	—	—	*Acyrthosiphon pisum* Shrls	4E-28	**
Srf (mads)	—	—	—	—	—	—	—			
Tead4	FKK7YRX02MPMDC	4E-03	7E-03	—	—	1E-13	2E-13	*Drosophila melanogaster* scalloped isoform A	2E-49	3E-42
Trx, Nsd1 (trxG)	contig01748	1E-131	2E-98	—	—	1E-102	8E-99	—	—	—
Trx2 (trxG)	—	—	—	—	—	—	—			
Tulp4L, Tubby (tulp)	FKK7YRX02MD5K1	7E-16	1E-15	—	—	—	—	*Ornithorhynchus anatinus* Tubby-related 4	4E-30	1E-32

* Sp-E2f3 matches contig06767 better than Sp-E2f4 does, but contig06767 matches back to Sp-E2f4 better.

** FKK7YRX02KT78U matches best to *Acyrthosiphon pisum* Shrls, which matches back to FKK7YRX02ONPEM better.

Chapter Four:
Conclusion

SUMMARY

This study has sequenced and characterized the gastrula transcriptome of the brittle star *Ophiocoma wendtii,* and begun an examination of its developmental gene regulatory network in relation to those of other echinoderms. Figures 4.1 and 4.2 show the genes in the sea urchin GRN for which homologs were found in the brittle star gastrula transcriptome. At this early stage of analysis, it appears that brittle stars employ a conserved developmental genetic toolkit, with various important features shared with other echinoderms, with other deuterostomes, and with the entire animal kingdom. Most of the differences in terms of the genes present and absent in the brittle star gastrula transcriptome are consistent with morphological differences between brittle stars and other echinoderms. In regard to the brittle star embryonic skeleton, the most plausible explanation is that it was acquired through an independent early activation of the adult skeletogenic program, with an unknown gene other than *Pmar1* acting as the master switch.

AREAS FOR FURTHER INVESTIGATION

This study represents only the first few steps toward a full understanding of the molecular players and events behind brittle star development. Much work

remains to be done in detailing the locations of gene expression, the patterns of expression at other developmental stages, and the regulatory interactions between genes, as well as sequencing of the full brittle star genome. Documentation of the genes active at the important stage of gastrulation provides a map to guide further research.

Tables 4.1 and 4.2 list nine transcription factor genes which warrant further inquiry because they were found to be expressed only in brittle star, or expressed in brittle star and starfish gastrulae but not in sea urchins. *FoxF* does not appear in urchins until almost the end of gastrulation, when it is seen in the coelomic pouches [Tu *et al* 2006], but its function there is unknown. Brittle stars form the coelomic pouches earlier, so its presence in the gastrula transcriptome is not surprising. Of the remainder, *Ebf3* is seen in the adult urchin nervous system, while *Meis* and *NfiA* are involved in vertebrate neural development. This is consistent with observations that neurogenesis appears to proceed faster in brittle stars than in urchins. Most of the key genes seen in urchin apical neuroectoderm [Burke *et al* 2006, Wei *et al* 2009] and in early neural development generally [Angerer *et al* 2011, Nomaksteinsky *et al* 2009, Vopalensky and Kozmik 2009, Lagutin *et al* 2003] were also found in brittle star, including *Six3, FoxQ2, SoxB1, Pax6, Dac, Otx, Mitf, Hlf, BMP2/4,* and *Chordin.* Three important genes which function early in urchin neural development, *Rx, Ac/Sc,* and *Hbn* were not found in brittle star, possibly because they have already completed their roles in the accelerated neural program prior to gastrulation.

The three genes in Table 4.2 are found in starfish gastrula, but none of these, nor the remaining genes in Table 4.1, have been seen to be significantly expressed at all during sea urchin embryogenesis, and nothing is known of their roles in echinoderm development. Their presence in the brittle star gastrula thus invites further investigation of their spatial distributions and regulatory functions.

In the last few years, several related processes collectively termed RNA interference (RNAi) have gained considerable attention. RNAi appears to serve several important functions in virtually all eukaryotic taxa, most notably both pre- and post-transcriptional silencing of gene expression, but also promotion of transcription and antiviral defense [Hannon 2002]. RNAi thus has the potential to play a vital role in developmental gene regulatory networks, and much evidence is accumulating for this [Stefani and Slack 2008].

Examination of RNAi participation in echinoderm development has thus far been very preliminary. Both urchins and starfish possess all components of the RNAi pathway, and both dynamically express a variety of known and suspected miRNAs with significant similarities and differences between the two organisms [Rodriguez *et al* 2005, Song and Wessel 2007, Kadri *et al* 2011]. Perturbation studies have shown that many genes central to echinoderm GRNs are influenced by RNAi processes. Urchin embryos in which RNAi is inhibited show major defects in germ-layer and axis specification, skeletogenesis, and gastrulation [Okamitsu *et al* 2010, Song *et al* 2012]. Brittle stars possess homologs for all components of the RNAi pathway [Table 4.3]. Investigation of the roles of specific

interfering RNAs and their contributions to development in the various echinoderm classes is another major area deserving further study.

The exact position of the brittle stars within the echinoderm family tree remains uncertain—but perhaps slightly less so. An independent acquisition of the embryonic skeleton would cast into doubt the main evidence for a clade of urchins and brittle stars. At the same time, the major argument for placing brittle stars with starfish—that they both have five arms—is weak on several morphological fronts, as discussed earlier. The brittle star gastrula transcriptome has slightly more matches to starfish than to either urchin species for transcription factor genes [Figure 3.1], but only slightly (57% vs 49% and 52%). Thus no firm phylogenetic conclusions can be drawn at this point. However, this first step toward GRN comparisons may still be a step forward in this regard. In the same way that specific mutations and chromosomal rearrangements can be used as phylogenetic characters to untangle evolutionary relationships, specific conservations and alterations in the structure of gene regulatory networks could potentially be used in cladistic analyses as well. This will require exploring the spatial distributions and definite regulatory interactions between the brittle star GRN components, as they have been studied in sea urchins, and are currently being examined in starfish. Drawing the GRN structures for the various echinoderm groups—and many other groups of organisms as well—will be another tool for drawing the tree of life.

FINAL THOUGHTS

The twin mottos of evo-devo might be stated as, “It’s *déjà vu* all over again, but everything old is new again.” On the one hand, many of the transcription factors and signalling pathways are as old as the animal kingdom itself, if not older, and found in virtually every phylum. Many of these transcription factors and signals have functional roles and interactions that have also been extremely widely conserved. On the other hand, evolution comes from doing something differently. There is nothing absolute that requires gene X to perform function Y in the process of developing organ Z. All that is really required is that a functional body gets built and is able to pass on more copies of its genes than other variations on that body plan currently around in that species’ ecological niche. So throughout evolutionary history, transcription factors, signalling molecules, and even large pieces of GRN circuitry have frequently been redeployed to appear at different times during the lifespan (such as the early—and independent—activation of the skeleton program in sea urchin and brittle star embryos), in different parts of the body (as with the six legs of insects and the many legs of centipedes), or to do completely different jobs (as with *Tbrain* in the starfish gut, in the sea urchin skeleton, and in whatever it does—if anything—in the brittle star) [Ettensohn *et al* 2007, Gao and Davidson 2008, Galant and Carroll 2002, Ronshaugen *et al* 2002, Hinman *et al* 2003a]. While currently only informed speculations about the exact structure of the brittle star GRN are possible, it is certain that, as in all animals, some combination of conservation and innovation is at work in brittle star development.

An example of this entanglement of sameness and difference is the role of BMP/Chordin signalling in axis specification. BMPs are diffusible secreted TGFβ signalling proteins, while Chordin binds to BMPs and thereby blocks them from binding to their receptors. BMP induces the ectoderm to become skin. Where BMP signals are blocked by high levels of Chordin, ectoderm proceeds to its default fate as neural tissue. In the embryos of most bilateral animals, including insects, a BMP is expressed on the dorsal side and diffuses toward the ventral side where it is inhibited by Chordin, thereby setting up a gradient of BMP signalling that patterns the dorsoventral axis and allows formation of a ventral nerve cord.

In vertebrates, tunicates, and amphioxus, the polarity is reversed, with BMP ventral and Chordin dorsal along the notochord (hence the name), thereby leading eventually to a dorsal nerve cord [De Robertis and Sasai 1996, Yu *et al* 2007]. The chordate heart is moved from dorsal to ventral, and the locations of expression for all the other genes along the DV axis are inverted as well, as are those for the left-right axis. It has been suggested that the event leading to this was a relocation of the mouth from its ancestral place on the Chordin-expressing side (as in insects) to the BMP-expressing side in the chordate common ancestor [De Robertis and Sasai 1996].

In hemichordate acorn worms, the BMP/Chordin DV polarity is the same as in insects and opposite that of chordates, but neural development is not repressed by BMP, and hemichordates exhibit a more diffuse nervous system, though the nerve net still displays some degree of centralization and

anteroposterior molecular patterning very similar to chordates [Lowe *et al* 2003 & 2006, Nomaksteinsky *et al* 2009, Röttinger and Martindale 2011].

In cnidarians, the nervous system is thoroughly diffuse and unaffected by BMP. Cnidarians are traditionally considered radial, but do possess a subtle secondary anatomical axis (the "directive axis") perpendicular to the oral-aboral axis. BMP and Chordin are expressed on the same side of the body, but with BMP in the endoderm and Chordin in the ectoderm [Matus *et al* 2006]. They also interact in further complex ways along both axes, so that there is no simple homology to the bilaterian DV axis [Saina *et al* 2009].

Sea urchins put yet another strange twist on this story. Here, both BMP and Chordin are expressed on the ventral (oral) side. Bound together, they diffuse to the dorsal (aboral) side, where Chordin is cleaved off and BMP binds to its receptors to trigger expression of other dorsal specification genes [Lapraz *et al* 2009]. This is likely an urchin-specific trait, since both sea cucumbers (the sister taxon to sea urchins), and hemichordates (the sister taxon to echinoderms) express BMP on the dorsal side [Harada *et al* 2002]. Dorsal BMP expression is therefore also the most likely pattern for brittle stars. Most of the other DV patterning genes appear in the same locations in both urchins and indirectly developing hemichordates, and most of these were also found in brittle star. And even though urchins and hemichordates express BMP on opposite sides, BMP is active only on the dorsal side in both, so the effect is the same [Röttinger and Martindale 2011].

"It's *déjà vu* all over again, but everything old is new again."

ACKNOWLEDGEMENTS

Thanks to those who provided help and support during this project: Brian Livingston, Jim Garey, Kelley Thomas, Nancy Garnhart, Mitch Ruzek, Darren Bauer, Dan Bergeron, Tiehang Wu, The Center for Genomics and Bioinformatics at Indiana University–Bloomington, USF Physicians Group, University Community Hospital, Ralph Vaughn, Marie Vaughn, Jillian Warren, and Mike Hall. This work was funded by National Science Foundation grant 0909797 to Brian Livingston.

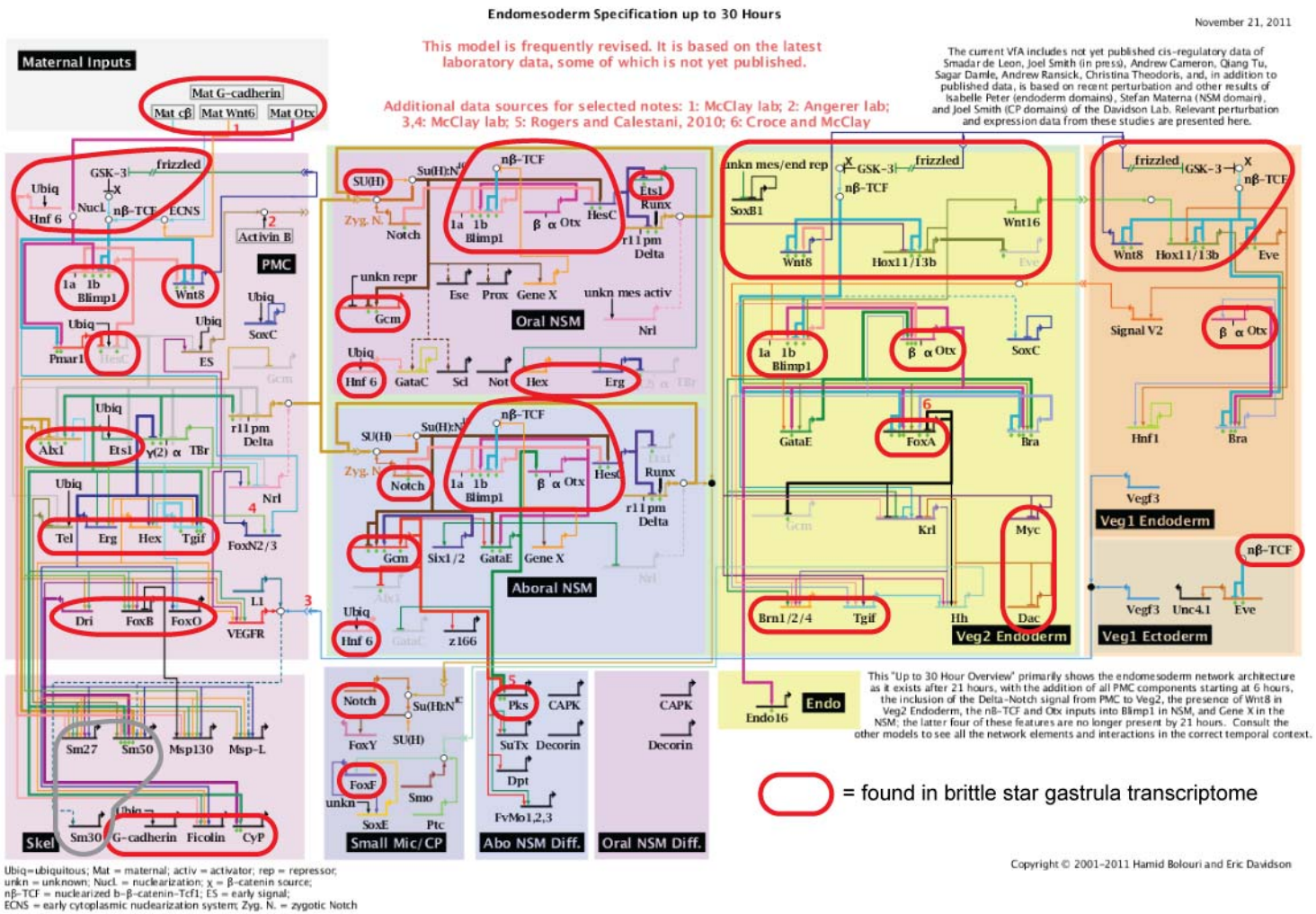

Figure 4.1: Genes from the Sea Urchin Endomesodermal Developmental Gene Regulatory Network Expressed in the Brittle Star Gastrula-Stage Embryo. Symbolism is the same as for Figure 1.2 (page 4). Genes with homologs found in the brittle star gastrula transcriptome are **circled in red**. [Figure generated with BioTapestry software – www.biotapestry.org]

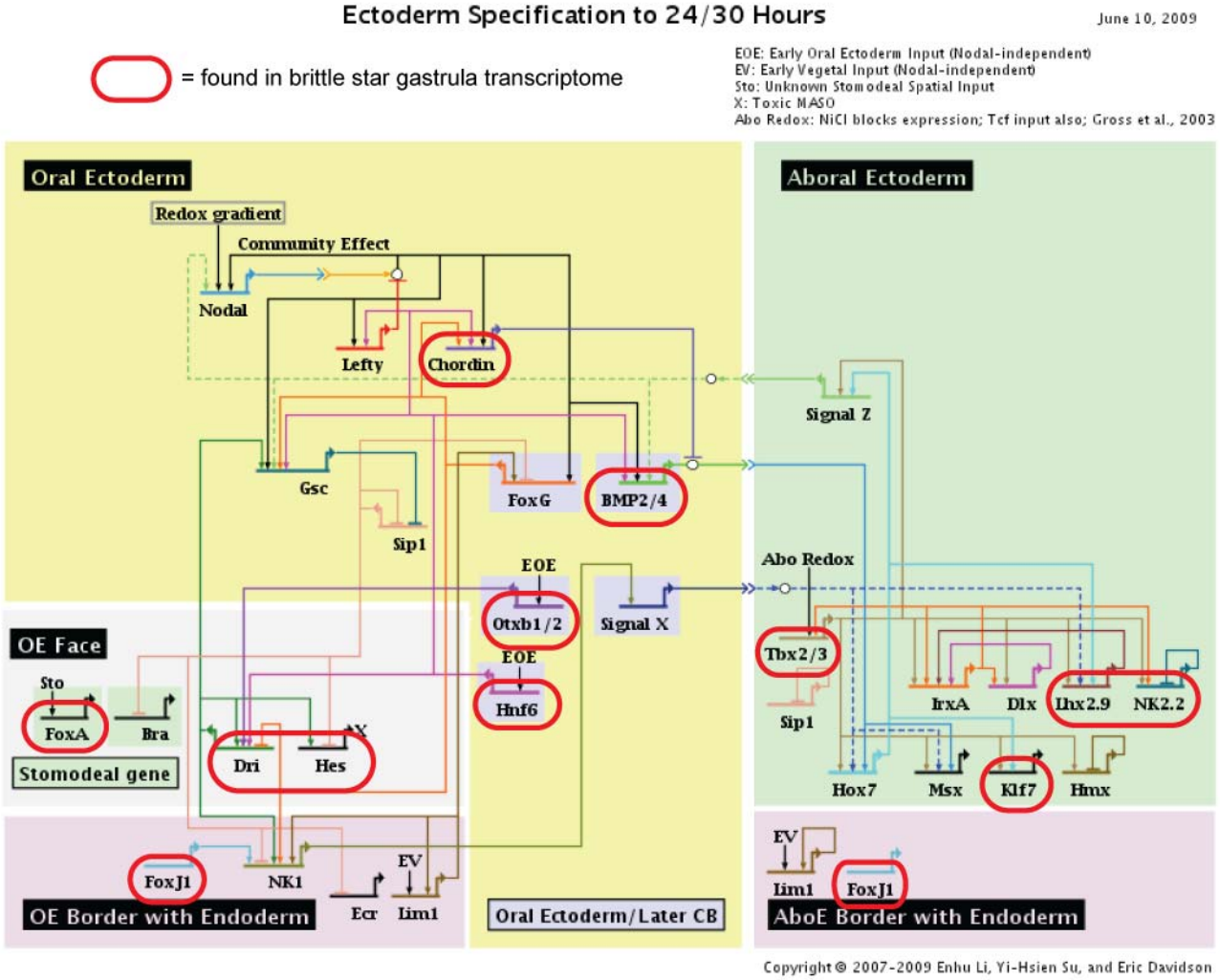

Figure 4.2: Genes from the Sea Urchin Ectodermal Developmental Gene Regulatory Network Expressed in the Brittle Star Gastrula-Stage Embryo. Symbolism is the same as for Figure 1.3 (page 5). Genes with homologs found in the brittle star gastrula transcriptome are **circled in red**. [Figure generated with BioTapestry software – www.biotapestry.org]

Table 4.1: Transcription Factors Expressed in Brittle Star Gastrula But Not in Urchin or Starfish Gastrulae

Gene	Brittle Star Sequence ID	Family	Functions in Other Taxa
Ebf3	FKK7YRX02NS71F	bHLH	Urchin — adult neural Vertebrates — tumor suppressor, neural & limb development
FoxF	FKK7YRX02OCVUU	Fox	Urchin — late gastrula in coelomic pouches Tunicate — heart Vertebrates — sexual development, insulin sensitivity, gut development Fly & Nematode — visceral mesoderm
Meis	FKK7YRX02K5EES	Homeobox	Vertebrates — neural development, restless leg syndrome Fly — *Homothorax* — pre-blastoderm divisions, head, neural, & limb development
NfiA	FKK7YRX02N238H	Other	Vertebrates — neural development
Ppar2	FKK7YRX02MESJR	NR	Vertebrates — xenobiotic defense
Scml1	FKK7YRX02MIB2G	Other	Vertebrates & Flies — repression of *Hox* genes Primates — spermatogenesis

References: ***Ebf3*** – Burke *et al* 2006, Garcia-Dominguez *et al* 2003, Mella *et al* 2004. ***FoxF*** – Tu *et al* 2006, Beh *et al* 2007, Ormestad *et al* 2006, Zaffran *et al* 2001. ***Meis*** – Larsen *et al* 2010, Salvany *et al* 2009, Kurant *et al* 1998, Noro *et al* 2006. ***NfiA*** – Gronostajski 2000. ***Ppar2*** – Goldstone 2006. ***Scml1*** – van de Vosse 1998, Wu and Su 2008.

Table 4.2: Transcription Factors Expressed in Brittle Star and Starfish Gastrulae But Not in Urchin Gastrula

Gene	Brittle Star Sequence ID	Family	Functions in Other Taxa
Lass6	contig01288	Homeobox	Vertebrates — apoptosis
Maf	contig00697	bZip	Vertebrates — oxidative stress response, hematopoiesis
Msxl	FKK7YRX02NW07M	Homeobox	Vertebrates — craniofacial, tooth & limb development

References: ***Lass6*** – Mizutani *et al* 2005. ***Maf*** – Kusakabe *et al* 2011. ***Msxl*** – Alappat *et al* 2003.

Table 4.3: Genes in the RNA Interference Pathway

Gene	*Ow* ID	*Sp* vs *Ow*	*Ow* vs *Sp*	*Lv* vs *Ow*	*Ow* vs *Lv*	*Pm* vs *Ow*	*Ow* vs *Pm*	*Ow* vs GB	*Ow* vs GB	GB vs *Ow*
Drosha	FKK7YRX02OUK17	2E-18	3E-18	—	—	—	—	Sp-Drosha	2E-18	3E-18
Dicer	FKK7YRX02K01VL	2E-19	4E-19	—	—	—	—	*Drosophila melanogaster* Dicer-2	3E-43	3E-40
Exportin5	FKK7YRX02MJIMD	1E-01	2E-01	—	—	—	—		—	—
Dgcr8/Pasha	FKK7YRX02OYHJV	3E-04	6E-04	6E-04	6E-04	—	—		—	—
Tarbp2	FKK7YRX02KMNNS	7E-09	1E-08	3E-10	9E-11	—	—	*Saccoglossus kowalevskii* TARBP2	4E-14	4E-14
Argonaute	FKK7YRX02MN77C	2E-22	6E-23	4E-07	2E-07	7E-27	2E-27	*Acyrthosiphon pisum* Argonaute-2	1E-39	2E-36
Piwi	contig04999	1E-107	1E-106	1E-120	1E-121	1E-119	1E-119	Sp-Seawi	6E-123	1E-107

References

Adell T, Muller I 2004: **Isolation and characterization of five Fox (Forkhead) genes from the sponge *Suberites domuncula.*** *Gene* 334: 35–46.

Aizenberg J, Tkachenko A, Weiner S, Addadi L, Hendler G 2001: **Calcitic microlenses as part of the photoreceptor system in brittlestars.** *Nature* 412:819-822.

Alappat S, Zhang ZY, Chen YP 2003: ***Msx* homeobox gene family and craniofacial development.** *Cell Res* 13:429-442.

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: 3389-3402.

Amore G, Yavrouian RG, Peterson KJ, Ransick A, McClay DR, Davidson EH 2003: ***Spdeadringer*, a sea urchin embryo gene required separately in skeletogenic and oral ectoderm gene regulatory networks.** *Dev Biol* 261:55-81.

Angerer LM, Oleksyn DW, Levine AM, Li X, Klein WH, Angerer RC 2001: **Sea urchin *goosecoid* function links fate specification along the animal-vegetal and oral-aboral embryonic axes.** *Development* 128:4393-4404.

Angerer LM, Yaguchi S, Angerer RC, Burke RD 2011: **The evolution of nervous system patterning: insights from sea urchin development.** *Development*. 138:3613-3623.

Arenas-Mena C, Cameron AR, Davidson EH 2000: **Spatial expression of *Hox* cluster genes in the ontogeny of a sea urchin.** *Development* 127: 4631–4643.

Arnone MI, Rizzo F, Annunciata R, Cameron RA, Peterson KJ, Martínez P 2006: **Genetic organization and embryonic expression of the ParaHox genes in the sea urchin *S purpuratus*: insights into the relationship between clustering and colinearity.** *Dev Biol* 300: 63-73.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM,

Sherlock G 2000: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium**. *Nat Genet* 25:25-29

Beh J, Shi W, Levine M, Davidson B, Christiaen L 2007: ***FoxF* is essential for FGF-induced migration of heart progenitor cells in the ascidian *Ciona intestinalis.*** *Development* 134:3297-3305.

Bradham CA, Oikonomou C, Kühn A, Core AB, Modell JW, McClay DR, Poustka AJ 2009: **Chordin is required for neural but not axial development in sea urchin embryos.** *Dev Biol* 328:221-233.

Briggs E, Wessel GM 2006: **In the beginning… Animal fertilization and sea urchin development.** *Dev Biol* 300:15-26.

Burke RD, Angerer LM, Elphick MR, Humphrey GW, Yaguchi S, Kiyama T, Liang S, Mu X, Agca C, Klein WH, Brandhorst BP, Rol M, Wilson K, Churcher AM, Taylor JS, Chen N, Murray G, Wang D, Mellott D, Olinski R, Hallböök F, Thorndyke MC 2006: **A genomic view of the sea urchin nervous system.** *Dev Biol* 300:434-460.

Cameron RA, Samanta M, Yuan A, He D, Davidson E 2009: **SpBase: the sea urchin genome database and web site.** *Nucleic Acids Res* 37: D750-754. [http://wwwspbaseorg/SpBase/]

Coffman JA, Kirchhamer CV, Harrington MG, Davidson EH 1996: **SpRunt-1 a new member of the runt-domain family of transcription factors is a positive regulator of the aboral ectoderm-specific *CyIIIa* gene in sea urchin embryos.** *Dev Biol* 174: 43-54.

Coffman JA, Kirchhamer CV, Harrington MG Davidson EH 1997: **SpMyb functions as an intramodular repressor to regulate spatial expression of *CyIIIa* in sea urchin embryos.** *Development* 124: 4717–4727.

Cooper VS, Vohr SH, Wrocklage SC, Hatcher PJ 2010: **Why genes evolve faster on secondary chromosomes in bacteria**. *PLoS Comput Biol* 6:e1000732.

Croce JC, McClay DR 2010: **Dynamics of Delta/Notch signaling on endomesoderm segregation in the sea urchin embryo.** *Development* 137:83-91.

Darras S, Gerhart J, Terasaki M, Kirschner M, Lol CJ 2011: **β-catenin specifies the endomesoderm and defines the posterior organizer of the hemichordate *Saccoglossus kowalevskii.*** *Development*. 138:959-970.

Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh CH, Minokawa T, Amore G, Hinman V, Arenas-Mena C, Otim O, Brown CT, Livi CB, Lee PY, Revilla R, Rust AG, Pan Z, Schilstra MJ, Clarke PJ, Arnone MI, Rowen L, Cameron RA, McClay DR, Hood L, Bolouri H 2002a: **A genomic regulatory network for development.** *Science* 295: 1669-1678.

Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh CH, Minokawa T, Amore G, Hinman V, Arenas-Mena C, Otim O, Brown CT, Livi CB, Lee PY, Revilla R, Schilstra MJ, Clarke PJ, Rust AG, Pan Z, Arnone MI, Rowen L, Cameron RA, McClay DR, Hood L, Bolouri H 2002b: **A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo.** *Dev Biol* 246: 162-190.

Davidson EH, Erwin DH 2006: **Gene regulatory networks and the evolution of animal body plans.** *Science* 311: 796-800.

Degnan BM, Degnan SM, Naganuma T, Morse DE 1993: **The *ets* multigene family is conserved throughout the Metazoa.** *Nucleic Acids Res* 21: 3479-3484.

De Robertis EM, Sasai Y 1996: **A common plan for dorsoventral patterning in Bilateria.** *Nature* 380:37-40.

Dobias SL, Zhao AZ, Tan H, Bell JR, Maxson R 1996: ***SpHbox7*, a new Abd-B class homeobox gene from the sea urchin *Strongylocentrotus purpuratus*: insights into the evolution of hox gene expression and function.** *Dev Dyn* 207:450-460.

Duboc V, Lepage T 2008a: **A conserved role for the nodal signaling pathway in the establishment of dorso-ventral and left–right axes in deuterostomes.** *J Exp Zool B Mol Dev Evol* 310B:41-53.

Duboc V, Lapraz F, Besnardeau L, Lepage T 2008b: **Lefty acts as an essential modulator of Nodal activity during sea urchin oral-aboral axis formation.** *Dev Biol* 320:49-59.

Ettensohn CA, Illies MR, Oliveri P, De Jong DL 2003: **Alx1, a member of the Cart1/Alx3/Alx4 subfamily of Paired-class homeodomain proteins, is an essential component of the gene network controlling skeletogenic fate specification in the sea urchin embryo.** *Development*. 130:2917-2298.

Ettensohn CA, Kitazawa C, Cheers MS, Leonard JD, Sharma T 2007: **Gene regulatory networks and developmental plasticity in the early sea**

urchin embryo: alternative deployment of the skeletogenic gene regulatory network. *Development* 134: 3077-3087.

Ettensohn CA 2009: **Lessons from a gene regulatory network: echinoderm skeletogenesis provides insights into evolution, plasticity and morphogenesis.** *Development* 136:11-21.

Flynn KM, Vohr SH, Hatcher PJ, Cooper VS 2010: **Evolutionary rates and gene dispensability associate with replication timing in the archaeon *Sulfolobus islandicus***. *Genome Biol Evol* 2:859-869.

Galant R, Carroll SB 2002: **Evolution of a transcriptional repression domain in an insect Hox protein.** *Nature* 415:910-913.

Gao F, Davidson EH 2008: **Transfer of a large gene regulatory apparatus to a new developmental address in echinoid evolution.** *Proc Natl Acad Sci USA* 105:6091-6096.

Garcia-Dominguez M, Poquet C, Garel S, Charnay P 2003: ***Ebf* gene function is required for coupling neuronal differentiation and cell cycle exit.** *Development* 130:6013-6025.

Goldstone JV, Hamdoun A, Cole BJ, Howard-Ashby M, Nebert DW, Scally M, Dean M, Epel D, Hahn ME, Stegeman JJ 2006: **The chemical defensome: environmental sensing and response genes in the *Strongylocentrotus purpuratus* genome.** *Dev Biol* 300:366-384.

Gotelli NJ, Entsminger GL 2011: **EcoSim: null models software for ecology. Version 7.** [http://garyentsminger.com/ecosim.htm]

Gronostajski RM 2000: **Roles of the *NFI/CTF* gene family in transcription and development.** *Gene* 249:31-45.

Gross JM, McClay DR 2001: **The role of Brachyury (T) during gastrulation movements in the sea urchin *Lytechinus variegatus.*** *Dev Biol* 239:132-147.

Gross JM, Peterson RE, Wu SY, McClay DR 2003: ***LvTbx2/3*: a T-box family transcription factor involved in formation of the oral/aboral axis of the sea urchin embryo.** *Development* 130: 1989–1999.

Hannon GJ 2002: **RNA interference.** *Nature* 418:244-251.

Harada Y, Shoguchi E, Taguchi S, Okai N, Humphreys T, Tagawa K, Satoh N 2002: **Conserved expression pattern of BMP-2/4 in hemichordate**

acorn worm and echinoderm sea cucumber embryos. *Zoolog Sci* 19:1113-1121.

Harmon MC 2005: **The position of the ophiuroidea within the phylum echinodermata.** MS thesis. University of South Florida, Biology Department.

Henry JQ, Tagawa K, Martindale MQ 2001: **Deuterostome evolution: early development in the enteropneust hemichordate, *Ptychodera flava.*** *Evol Dev* 3:375–390.

Hinman VF, Nguyen AT, Cameron A, Davidson EH 2003a: **Developmental gene regulatory network architecture across 500 million years of echinoderm evolution.** *Proc Natl Acad Sci USA* 100: 13356-13361.

Hinman VF, Nguyen AT, Davidson EH 2003b: **Expression and function of a starfish Otx ortholog, AmOtx: a conserved role for Otx proteins in endoderm development that predates divergence of the eleutherozoa.** *Mech Dev* 120:1165-1176.

Hinman VF, Davidson EH 2003c: **Expression of a gene encoding a Gata transcription factor during embryogenesis of the starfish *Asterina miniata.*** *Gene Expr Patterns* 3:419-422.

Hinman VF, Davidson EH 2007a: **Evolutionary plasticity of developmental gene regulatory network architecture.** *Proc Natl Acad Sci USA* 104:19404-19409.

Hinman VF, Nguyen A, Davidson EH 2007b: **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: 584-595.

Holland LZ 2002: **Heads or tails? Amphioxus and the evolution of anterior-posterior patterning in deuterostomes.** *Dev Biol* 241:209-228.

Howard EW, Newman LA, Oleksyn DW, Angerer RC, Angerer LM 2001: ***SpKrl*: a direct target of β-catenin regulation required for endoderm differentiation in sea urchin embryos.** *Development* 128:365-375.

Howard-Ashby M, Materna SC, Brown CT, Chen L, Cameron RA, Davidson EH 2006a: **Identification and characterization of homeobox transcription factor genes in *Strongylocentrotus purpuratus* and their expression in embryonic development.** *Dev Biol* 300: 74-89.

Howard-Ashby M, Materna SC, Brown CT, Chen L, Cameron RA, Davidson EH 2006b: **Gene families encoding transcription factors expressed in**

early development of *Strongylocentrotus purpuratus.* *Dev Biol* 300: 90-107.

Itoh S, Itoh F, Goumans MJ, Ten Dijke P 2000: **Signaling of transforming growth factor-β family members through Smad proteins.** *Eur J Biochem* 267:6954-6967.

Juliano CE, Voronina E, Stack C, Aldrich M, Cameron AR, Wessel GM 2006: **Germ line determinants are not localized early in sea urchin development, but do accumulate in the small micromere lineage.** *Dev Biol* 300:406-415.

Kadri S, Hinman VF, Benos PV 2011: **RNA deep sequencing reveals differential microRNA expression during development of sea urchin and sea star.** *PLoS One* 6:e29217.

Kanehisa M. Goto S 2000**: KEGG: Kyoto encyclopedia of genes and genomes**. *Nucleic Acids Res* 28:27-30.

KEGG Orthology Database [http://www.genome.jp/kegg/ko.html]

Kenny AP, Oleksyn DW, Newman LA, Angerer RC, Angerer LM 2003: **Tight regulation of SpSoxB factors is required for patterning and morphogenesis in sea urchin embryos.** *Dev Biol* 261: 412–425.

Kinoshita K, Kikuchi Y, Sasakura Y, Suzuki M, Fujii-Kuriyama Y, Sogawa K 2004: **Altered DNA binding specificity of Arnt by selection of partner bHLH-PAS proteins.** *Nucleic Acids Res* 32:3169-3179.

Kiyama T, Klein WH 2007: **SpGataE, a *Strongylocentrotus purpuratus* ortholog of mammalian Gata4/5/6: protein expression, interaction with putative target gene *spec2a*, and identification of friend of Gata factor SpFog1.** *Dev Genes Evol* 217:651-663.

Knight RD, Shimeld SM 2001: **Identification of conserved C_2H_2 zinc-finger gene families in the Bilateria.** *Genome Biol* 2: research00161-00168

Kurant E, Pai CY, Sharf R, Halachmi N, Sun YH, Salzberg A 1998: ***Dorsotonals/homothorax,* the *Drosophila* homologue of *meis1*, interacts with *extradenticle* in patterning of the embryonic PNS.** *Development* 125:1037-1048.

Kusakabe M, Hasegawa K, Hamada M, Nakamura M, Ohsumi T, Suzuki H, Tran MT, Kudo T, Uchida K, Ninomiya H, Chiba S, Takahashi S 2011: ***c-Maf* plays a crucial role for the definitive erythropoiesis that accompanies erythroblastic island formation in the fetal liver.** *Blood* 118:1374-1385.

Lagutin OV, Zhu CC, Kobayashi D, Topczewski J, Shimamura K, Puelles L, Russell HR, McKinnon PJ, Solnica-Krezel L, Oliver G 2003: **Six3 repression of Wnt signaling in the anterior neuroectoderm is essential for vertebrate forebrain development.** *Genes Dev* 17:368-379.

Laity JH, Lee BM, Wright PE 2001: **Zinc finger proteins: new insights into structural and functional diversity.** *Curr Opin Struct Biol* 11: 39–46.

Lapraz F, Besnardeau L, Lepage T 2009: **Patterning of the dorsal-ventral axis in echinoderms: insights into the evolution of the BMP-chordin signaling network.** *PLoS Biol* 7:e1000248.

Larsen KB, Lutterodt MC, Laursen H, Graem N, Pakkenberg B, Møllgård K, Møller M 2010: **Spatiotemporal distribution of *PAX6* and *MEIS2* expression and total cell numbers in the ganglionic eminence in the early developing human forebrain.** *Dev Neurosci* 32:149-162.

Lee YH, Huang GM, Cameron RA, Graham G, Davidson EH, Hood L, Britten RJ 1999: **EST analysis of gene expression in early cleavage-stage sea urchin embryos.** *Development* 126:3857-3867.

Lerat E, Daubin V, Ochman H, Moran NA 2005: **Evolutionary origins of genomic repertoires in bacteria**. *PLoS Biol* 3:e130.

Li X, Chuang CK, Mao CA, Angerer LM, Klein WH 1997: **Two Otx proteins generated from multiple transcripts of a single gene in *Strongylocentrotus purpuratus.*** *Dev Biol* 187:253-266.

Littlewood DTJ, Smith AB, Clough KA, Emson RH 1997: **The interrelationships of the echinoderm classes: morphological and molecular evidence.** *Biol J Linnean Soc* 61:409-438.

Livingston BT, Killian CE, Wilt F, Cameron A, Landrum MJ, Ermolaeva O, Sapojnikov V, Maglott DR, Buchanan AM, Ettensohn CA 2006: **A genome-wide analysis of biomineralization-related proteins in the sea urchin *Strongylocentrotus purpuratus.*** *Dev Biol* 300:335-348.

Lowe CJ, Terasaki M, Wu M, Freeman RM Jr, Runft L, Kwan K, Haigo S, Aronowicz J, Lander E, Gruber C, Smith M, Kirschner M, Gerhart J 2006: **Dorsoventral patterning in hemichordates: insights into early chordate evolution.** *PLoS Biol* 4:e291.

Lowry JA, Atchley WR 2000: **Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding domain.** *J Mol Evol* 50:103-115.

Luke NH, Killian CE, Livingston BT 1997: ***Spfkh1* encodes a transcription factor implicated in gut formation during sea urchin development.** *Dev Growth Differ* 39: 285-294.

***Lytechinus variegatus* transcriptome database**
[http://www.spbase.org/LV/index.php]

Magie CR, Pang K, Martindale MQ 2005: **Genomic inventory and expression of Sox and Fox genes in the cnidarian *Nematostella vectensis.*** *Dev Genes Evol* 215: 618–630.

Mann K, Poustka AJ, Mann M 2008: **The sea urchin (*Strongylocentrotus purpuratus*) test and spine proteomes.** *Proteome Sci* 6:22.

Mann K, Wilt FH, Poustka AJ 2010: **Proteomic analysis of sea urchin (*Strongylocentrotus purpuratus*) spicule matrix.** *Proteome Sci* 8:33.

Maruyama YK 2000: **A sea cucumber homolog of the mouse *T-Brain-1* is expressed in the invaginated cells of the early gastrula in *Holothuria leucospilota.*** *Zoolog Sci* 17:383-387.

Materna SC, Howard-Ashby M, Gray RF, Davidson EH 2006a: **The C_2H_2 zinc finger genes of *Strongylocentrotus purpuratus* and their expression in embryonic development.** *Dev Biol* 300: 108-120.

Materna SC, Berney K, Cameron RA 2006b: **The *S. purpuratus* genome: A comparative perspective.** *Dev Biol* 300: 485-495.

Materna SC, Nam J, Davidson EH 2010: **High accuracy high-resolution prevalence measurement for the majority of locally expressed regulatory genes in early sea urchin development.** *Gene Expr Patterns* 10: 177-184. [http://vanbeneden.caltech.edu/~m/cgi-bin/hd-tc/plot.cgi]

Matus DQ, Thomsen GH, Martindale MQ 2006: **Dorso/ventral genes are asymmetrically expressed and involved in germ-layer demarcation during cnidarian gastrulation.** *Curr Biol* 16:499-505.

McCauley BS, Weideman EP, Hinman VF 2010: **A conserved gene regulatory network subcircuit drives different developmental fates in the vegetal pole of highly divergent echinoderm embryos.** *Dev Biol* 340:200-208.

McClay DR 2011: **Evolutionary crossroads in developmental biology: sea urchins.** *Development* 138:2639-2648.

Mella S, Soula C, Morello D, Crozatier M, Vincent A 2004: **Expression patterns of the *coe/ebf* transcription factor genes during chicken and mouse limb development.** Gene Expr Patterns 4:537-542.

Meyer E, Aglyamova GV, Wang S, Buchanan-Carter J, Abrego D, Colbourne JK, Willis BL, Matz MV 2009: **Sequencing and *de novo* analysis of a coral larval transcriptome using 454 GSFlx.** *BMC Genomics* 10:219.

Minemura K, Yamaguchi M, Minokawa T 2009: **Evolutionary modification of *T-brain (tbr)* expression patterns in sand dollar.** *Gene Expr Patterns* 9:468-474.

Minokawa T, Rast JP, Arenas-Mena C., Franco CB, Davidson EH 2004: **Expression patterns of four different regulatory genes that function during sea urchin development.** *Gene Expr Patterns* 4:449–456.

Mitsunaga-Nakatsubo K, Harada Y, Satoh N, Shimada H, Akasaka K 2001: ***Brachyury* homolog (*HpTa*) is involved in the formation of archenteron and secondary mesenchyme cell differentiation in the sea urchin embryo.** *Zoology* 104:99-102.

Mizutani Y, Kihara A, Igarashi Y 2005: **Mammalian *Lass6* and its related family members regulate synthesis of specific ceramides.** *Biochem J* 390:263-271.

Momose T, Derelle R, Houliston E 2008: **A maternally localised Wnt ligand required for axial patterning in the cnidarian *Clytia hemisphaerica.*** *Development* 135:2105-2113.

NCBI RefSeq proteins [http://www.ncbi.nlm.nih.gov/RefSeq/]

Nomaksteinsky M, Röttinger E, Dufour HD, Chettouh Z, Lowe CJ, Martindale MQ, Brunet JF 2009: **Centralization of the deuterostome nervous system predates chordates.** *Curr Biol* 19:1264-1269.

Noro B, Culi J, McKay DJ, Zhang W, Mann RS 2006: **Distinct functions of homeodomain-containing and homeodomain-less isoforms encoded by *homothorax*.** *Genes Dev* 20:1636-1650.

Okamitsu Y, Yamamoto T, Fujii T, Ochiai H, Sakamoto N 2010: **Dicer is required for the normal development of sea urchin, *Hemicentrotus pulcherrimus.*** *Zoolog Sci* 27:477-86.

Oliveri P, Davidson EH, McClay DR 2003: **Activation of *pmar1* controls specification of micromeres in the sea urchin embryo.** *Dev Biol* 258:32-43

Oliveri P, Davidson EH 2004: **Gene regulatory network controlling embryonic specification in the sea urchin.** *Curr Opin Genet Dev* 204: 351-380.

Oliveri P, Walton K, Davidson EH, McClay DR 2006: **Repression of mesodermal fate by foxa a key endoderm regulator of the sea urchin embryo.** *Development* 133: 4173–4181.

Ormestad M, Astorga J, Landgren H, Wang T, Johansson BR, Miura N, Carlsson P 2006: ***Foxf1*** **and** ***Foxf2*** **control murine gut development by limiting mesenchymal Wnt signaling and promoting extracellular matrix production.** *Development* 133:833-843.

Otim O, Amore G, Minokawa T, McClay DR, Davidson EH 2004: ***SpHnf6*** **a transcription factor that executes multiple functions in sea urchin embryogenesis.** *Dev Biol* 273: 226-243.

Patient RK, McGhee JD 2002: **The GATA family (vertebrates and invertebrates).** *Curr Opin Genet Dev* 12:416-422.

Patiria miniata **transcriptome database** [http://www.spbase.org/PM/index.php]

Paul CRC, Smith AB 1984: **The early radiation and phylogeny of echinoderms.** *Biol Rev* 59:443–461.

Peter IS, Davidson EH 2010: **The endoderm gene regulatory network in sea urchin embryos up to mid-blastula stage.** *Dev Biol* 340: 188-199.

Peter IS, Davidson EH 2011: **A gene regulatory network controlling the embryonic specification of endoderm.** *Nature* 474:635-639.

Pederson T 2006: **The sea urchin's siren.** *Dev Biol* 300:9-14.

Peterson KJ, Harada Y, Cameron RA, Davidson EH 1999: **Expression Pattern of** ***Brachyury*** **and** ***Not*** **in the Sea Urchin: Comparative Implications for the Origins of Mesoderm in the Basal Deuterostomes.** *Dev Biol* 207:419-431.

Poustka AJ, Herwig R, Krause A, Hennig S, Meier-Ewert S, Lehrach H 1999: **Toward the gene catalogue of sea urchin development: the construction and analysis of an unfertilized egg cDNA library highly normalized by oligonucleotide fingerprinting.** *Genomics* 59:122-133.

Poustka AJ, Groth D, Hennig S, Thamm S, Cameron A, Beck A, Reinhardt R, Herwig R, Panopoulou G, Lehrach H 2003: **Generation, annotation, evolutionary analysis, and database integration of 20,000 unique sea urchin EST clusters.** *Genome Res* 13:2736-2746.

Range RC, Venuti JM, McClay DR 2005: **LvGroucho and nuclear beta-catenin functionally compete for Tcf binding to influence activation of the endomesoderm gene regulatory network in the sea urchin embryo.** *Dev Biol* 279:252-267.

Ransick A, Rast JP, Minokawa T, Calestani C, Davidson EH 2002: **New early zygotic regulators expressed in endomesoderm of sea urchin embryos discovered by differential array hybridization.** *Dev Biol* 246:132-147.

Ransick A, Davidson EH 2006: ***cis*-regulatory processing of Notch signaling input to the sea urchin *glial cells missing* gene during mesoderm specification.** *Dev Biol* 297:587-602.

Revilla-i-Domingo R, Oliveri P, Davidson EH 2007: **A missing link in the sea urchin embryo gene regulatory network: *hesC* and the double-negative specification of micromeres.** *Proc Natl Acad Sci USA* 104:**1**2383-12388.

Rho HK, McClay D 2011: **The control of *foxN2/3* expression in sea urchin embryos and its function in the skeletogenic gene regulatory network.** *Development* 138:937-945.

Rizzo F, Fernandez-Serra M, Squarzoni P, Archimandritis A, Arnone MI 2006: **Identification and developmental expression of the *ets* gene family in the sea urchin (*Strongylocentrotus purpuratus*).** *Dev Biol* 300: 35-48.

Rodriguez AJ, Seipel SA, Hamill DA, Romancino DP, Di Carlo M, Suprenant KA, Bonder EM 2005: **Seawi—a sea urchin piwi/argonaute family member is a component of MT-RNP complexes.** *RNA* 11:646-656.

Ronshaugen M, McGinnis N, McGinnis W 2002: **Hox protein mutation and macroevolution of the insect body plan.** *Nature* 415:914-917.

Röttinger E, Martindale MQ 2011: **Ventralization of an indirect developing hemichordate by $NiCl_2$ suggests a conserved mechanism of dorso-ventral (D/V) patterning in Ambulacraria (hemichordates and echinoderms).** *Dev Biol* 354:173-190.

Saina M, Genikhovich G, Renfer E, Technau U 2009: **BMPs and chordin regulate patterning of the directive axis in a sea anemone.** *Proc Natl Acad Sci USA* 106:18592-18597.

Salvany L, Aldaz S, Corsetti E, Azpiazu N 2009: **A new role for *hth* in the early pre-blastodermic divisions in drosophila.** *Cell Cycle* 8:2748-2755.

Saudemont A, Haillot E, Mekpoh F, Bessodes N, Quirin M, Lapraz F, Duboc V, Röttinger E, Range R, Oisel A, Besnardeau L, Wincker P, Lepage T 2010: **Ancestral Regulatory Circuits Governing Ectoderm Patterning Downstream of Nodal and BMP2/4 Revealed by Gene Regulatory Network Analysis in an Echinoderm.** *PLoS Genetics* 6:1-31.

Sea Urchin Genome Sequencing Consortium 2006: **The genome of the sea urchin *Strongylocentrotus purpuratus.*** *Science* 314: 941-952.

Sharma T, Ettensohn CA 2010: **Activation of the skeletogenic gene regulatory network in the early sea urchin embryo.** *Development* 137:1149-1157.

Sharrocks AD 2001: **The ETS-domain transcription factor family.** *Nat Rev Mol Cell Biol* 2: 827-837.

Shoguchi E, Satoh N, Maruyama YK 1999: **Pattern of *Brachyury* gene expression in starfish embryos resembles that of hemichordate embryos but not of sea urchin embryos.** *Mech Dev* 82:185-189.

Smith J, Theodoris C, Davidson EH 2007: **A gene regulatory network subcircuit drives a dynamic pattern of gene expression.** *Science* 318:794-797.

Smith J, Kraemer E, Liu H, Theodoris C, Davidson E 2008: **A spatially dynamic cohort of regulatory genes in the endomesodermal gene network of the sea urchin embryo.** *Dev Biol* 313: 863-875.

Song JL, Wessel GM 2007: **Genes involved in the RNA interference pathway are differentially expressed during sea urchin development.** *Dev Dyn* 236:3180-3190.

Song JL, Stoeckius M, Maaskola J, Friedländer M, Stepicheva N, Juliano C, Lebedeva S, Thompson W, Rajewsky N, Wessel GM 2012: **Select microRNAs are essential for early development in the sea urchin.** *Dev Biol* 362:104-113.

Stefani G, Slack FJ 2008: **Small non-coding RNAs in animal development.** *Nat Rev Mol Cell Biol* 9:219-230.

Swalla BJ, Smith AB 2008: **Deciphering deuterostome phylogeny: molecular, morphological and palaeontological perspectives.** *Phil Trans R Soc B* 363:1557-1568.

Su YH 2009: **Gene regulatory networks for ectoderm specification in sea urchin embryos.** *Biochim Biophys Acta* 1789: 261-267.

Tagawa K, Humphreys T, Satoh N 2000: ***T-Brain* expression in the apical organ of hemichordate tornaria larvae suggests its evolutionary link to the vertebrate forebrain.** *J Exp Zool B Mol Dev Evol* 288:23-31.

Takacs CM, Amore G, Oliveri P, Poustka AJ, Wang D, Burke RD, Peterson KJ 2004: **Expression of an NK2 homeodomain gene in the apical ectoderm defines a new territory in the early sea urchin embryo.** *Dev Biol* 269: 152-164.

Tatusov RL, Koonin EV, Lipman DJ 1997: **A genomic perspective on protein families.** *Science* 278:631-637.

Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA 2003: **The COG database: an updated version includes eukaryotes**. *BMC Bioinformatics* 4:41.

Tu Q, Brown CT, Davidson EH, Oliveri P 2006: **Sea urchin *forkhead* gene family: phylogeny and embryonic expression.** *Dev Biol* 300: 49-62.

Ullrich-Lüter EM, Dupont S, Arboleda E, Hausen H, Arnone MI 2011: **Unique system of photoreceptors in sea urchin tube feet.** *Proc Natl Acad Sci USA* 108: 8367-8372.

UniGene sea urchin embryo cDNA libraries
[http://www.ncbi.nlm.nih.gov/UniGene/lbrowse2.cgi?TAXID=7668&CUTOFF=1000]

van de Vosse E, Walpole SM, Nicolaou A, van der Bent P, Cahn A, Vaudin M, Ross MT, Durham J, Pavitt R, Wilkinson J, Grafham D, Bergen AA, van Ommen GJ, Yates JR, den Dunnen JT, Trump D 1998: **Characterization of *SCML1*, a new gene in Xp22, with homology to developmental polycomb genes.** *Genomics* 49:96-102.

Vinson C, Myakishev M, Acharya A, Mir AA, Moll JR, Bonovich M 2002: **Classification of human B-Zip proteins based on dimerization properties.** *Mol Cell Biol* 22: 6321–6335.

Vopalensky P, Kozmik Z 2009: **Eye evolution: common use and independent recruitment of genetic components.** *Philos Trans R Soc Lond B Biol Sci* 364:2819-2832.

Wei Z, Angerer LM, Angerer RC 1999: **Spatially regulated SpEts4 transcription factor activity along the sea urchin embryo animal-vegetal axis.** *Development* 126: 1729-1737.

Wei Z, Angerer RC, Angerer LM 2006: **A database of mRNA expression patterns for the sea urchin embryo.** *Dev Biol* 300:476-484.

Wei Z, Yaguchi J, Yaguchi S, Angerer RC, Angerer LM 2009: **The sea urchin animal pole domain is a Six3-dependent neurogenic patterning center.** *Development* 136: 1179-1189.

Whittaker CA, Bergeron KF, Whittle J, Brandhorst BP, Burke RD, Hynes RO 2006: **The echinoderm adhesome.** *Dev Biol* 300:252–266.

Wilt FH, Ettensohn CA 2007: **The morphogenesis and biomineralization of the sea urchin larval skeleton.** In *Handbook of Biomineralization* (ed. E. Bauerlin), 183-210.

Wray GA, McClay DR 1988: **The origin of spicule-forming cells in a "primitive" sea urchin (*Eucidares tribloides*) which appears to lack primary mesenchyme cells.** *Development* 103:305-315.

Wu HH, Su B 2008: **Adaptive evolution of *SCML1* in primates, a gene involved in male reproduction.** *BMC Evol Biol* 8:192.

Yaguchi S, Yaguchi J, Burke RD 2006: **Sp-Smad2/3 mediates patterning of neurogenic ectoderm by nodal in the sea urchin embryo.** *Dev Biol* 302: 494-503.

Yaguchi S, Yaguchi J, Angerer RC, Angerer LM 2008: **A Wnt-FoxQ2-Nodal pathway links primary and secondary axis specification in sea urchin embryos.** *Dev Cell* 14: 97–107.

Yu JK, Satou Y, Holland ND, Shin-I T, Kohara Y, Satoh N, Bronner-Fraser M, Holland LZ 2007: **Axial patterning in cephalochordates and the evolution of the organizer.** *Nature*. 445:613-617.

Yuh CW, Dorman ER, Howard ML, Davidson EH 2004: **An otx cis-regulatory module: a key node in the sea urchin endomesoderm gene regulatory network.** *Dev Biol* 269:536-551.

Zaffran S, Küchler A, Lee HH, Frasch M 2001: ***biniou (FoxF),* a central component in a regulatory network controlling visceral mesoderm development and midgut morphogenesis in *Drosophila*.** *Genes Dev* 15:2900-2915.

Zhu X, Mahairas G, Illies M, Cameron RA, Davidson EH, Ettensohn CA 2001: **A large-scale analysis of mRNAs expressed by primary mesenchyme cells of the sea urchin embryo.** *Development* 128:2615-2627.