

Reference: *Biol. Bull.* **214**: 233–254. (June 2008)  
© 2008 Marine Biological Laboratory

# Genomic Survey of Candidate Stress-Response Genes in the Estuarine Anemone *Nematostella vectensis*

ADAM M. REITZEL<sup>1,\*</sup>, JAMES C. SULLIVAN<sup>2,\*</sup>, NIKKI TRAYLOR-KNOWLES<sup>2,\*</sup>,  
AND JOHN R. FINNERTY<sup>2,†</sup>

<sup>1</sup>*Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543; and*

<sup>2</sup>*Department of Biology, Boston University, 5 Cummington Street, Boston, Massachusetts 02215*

**Abstract.** Salt marshes are challenging habitats due to natural variability in key environmental parameters including temperature, salinity, ultraviolet light, oxygen, sulfides, and reactive oxygen species. Compounding this natural variation, salt marshes are often heavily impacted by anthropogenic insults including eutrophication, toxic contamination, and coastal development that alter tidal and freshwater inputs. Commensurate with this environmental variability, estuarine animals generally exhibit broader physiological tolerances than freshwater, marine, or terrestrial species. One factor that determines an organism's physiological tolerance is its ability to upregulate "stress-response genes" in reaction to particular stressors. Comparative studies on diverse organisms have identified a number of evolutionarily conserved genes involved in responding to abiotic and biotic stressors. We used homology-based scans to survey the sequenced genome of *Nematostella vectensis*, the starlet sea anemone, an estuarine specialist, to identify genes involved in the response to three kinds of insult—physiochemical insults, pathogens, and injury. Many components of the stress-response networks identified in triploblastic animals have clear orthologs in the sea anemone, meaning that they must predate the cnidarian-triploblast split (*e.g.*, xenobiotic receptors, biotransformative genes, ATP-dependent transporters, and genes involved in responding to reactive oxygen species, toxic metals, osmotic shock, thermal stress, pathogen exposure, and wounding). However, in some instances, stress-response genes known from triploblasts appear to be absent from the

*Nematostella* genome (*e.g.*, many metal-complexing genes). This is the first comprehensive examination of the genomic stress-response repertoire of an estuarine animal and a member of the phylum Cnidaria. The molecular markers of stress response identified in *Nematostella* may prove useful in monitoring estuary health and evaluating coastal conservation efforts. These data may also inform conservation efforts on other cnidarians, such as the reef-building corals.

## Introduction

To survive and reproduce, organisms must deploy an array of molecular, physiological, and behavioral responses to avoid or counteract detrimental environmental conditions. These responses to environmental stress are often complex, involving many genes and integrating events at cellular and organismal levels. Except in cases where it is possible to simply flee stressful environmental conditions, an organism's response to most environmental stressors—both natural and anthropogenic stressors—will typically involve changes in baseline gene expression that alter physiology, development, or behavior. Therefore, to study organismal stress responses, we must integrate environmental data, gene expression data, and organismal data (Ankley *et al.*, 2006; Lee and Mitchell-Olds, 2006). The ongoing proliferation of genomic data is enabling comparative stress-response studies on a phylogenetically wider sampling of taxa occupying a broader range of habitat types (Lettieri, 2006). It is now possible to compare evolved differences between species occupying diverse environments at both shallow and deep evolutionary divergences.

*Nematostella vectensis* as an estuarine sentinel species

*Nematostella vectensis* Stephenson, the starlet sea anemone, is one of the relatively small number of macroscopic

Received 13 November 2007; accepted 14 February 2008.

\* These authors contributed equally to this work.

† To whom correspondence should be addressed. E-mail: [jrf3@bu.edu](mailto:jrf3@bu.edu)

*Abbreviations:* ECM, extracellular matrix; EST, expressed sequence tag; ROS, reactive oxygen species.

animals that are year-round residents in estuarine environments. It is a small infaunal anemone (typically <1 cm) inhabiting salt marshes, saline lagoons, and other sheltered brackish environments. The species' range includes the eastern Pacific, western Atlantic, northern English Channel, and western North Sea (Hand and Uhlinger, 1994).

In recent years, *Nematostella* has emerged as a model system for developmental (Finnerty *et al.*, 2004; Matus *et al.*, 2006) and genomic studies (Sullivan *et al.*, 2006; Putnam *et al.*, 2007), and it has proven particularly informative for reconstructing the functional evolution of developmental regulatory genes whose origins can be traced to the cnidarian-triploblast common ancestor. It is useful for such studies because (1) it is an outgroup to the superphylum Triploblastica (=Bilateria; Fig. 1), (2) it is easy to culture in the laboratory through its entire sexual and asexual life history, and (3) its genome appears to have evolved in a relatively conservative fashion. A general finding of comparative genomic studies involving *Nematostella* is that the sea anemone exhibits remarkable genomic complexity and a surprising degree of genomic conservation relative to vertebrates, despite its extreme evolutionary distance from vertebrates (Technau *et al.*, 2005; Sullivan *et al.*, 2006; Putnam *et al.*, 2007; Sullivan and Finnerty, 2007).

*Nematostella* is also useful for studying organismal responses to environmental stress because, in addition to extensive genomic resources, (1) it lives in a spatially and temporally variable habitat that is easily accessible, (2) it is easily subjected to laboratory stress studies, and (3) it is

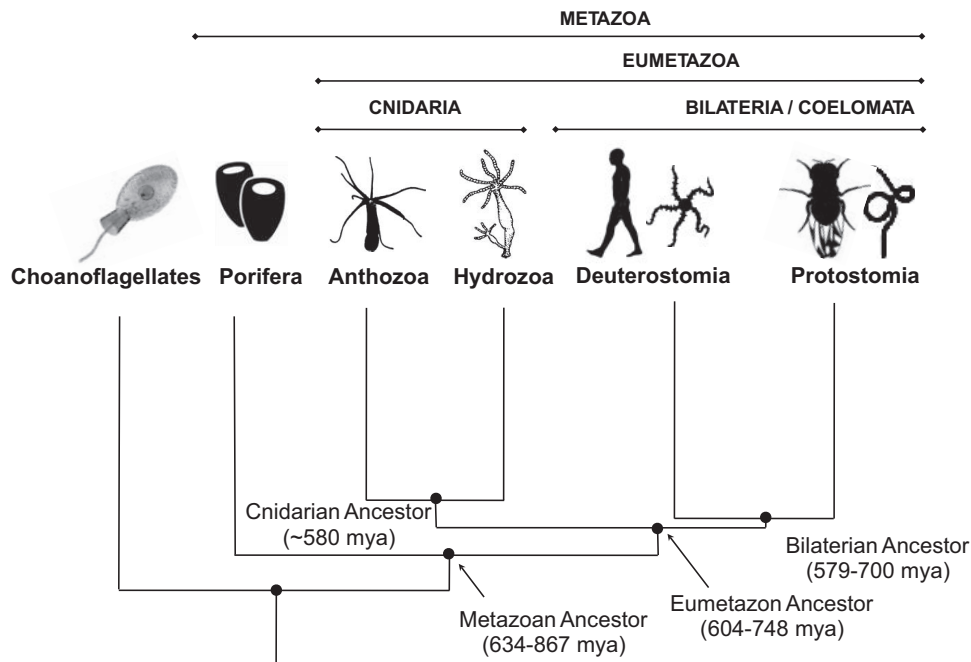
widely distributed (Hand and Uhlinger, 1994). Importantly, *Nematostella* has been collected in both pristine and anthropogenically impacted, degraded salt marsh habitats (Harter and Matthews, 2005), so it must be tolerant of a wide range of natural and anthropogenic stressors.

A first step in understanding how organisms sense and respond to stressors from the environment is to identify the genetic components of stress-response pathways. In this study, we searched the *Nematostella* genome to identify homologs of proteins known to be involved in sensing stress, counteracting particular stressors, and repairing damaged tissue. We divided the stress-response genes into three broad functional categories: (1) those that mediate chemical stresses, (2) those that counteract pathogens, (3) and those that underlie wound healing.

## Materials and Methods

### Identifying proteins implicated in stress response in other taxa

We compiled a list of proteins known to be involved in (1) chemical stress response, (2) innate immunity, and (3) wound healing. To identify genes involved in chemical stress response, we utilized a collection of "defensome" motifs compiled by Goldstone *et al.* (2006), with the following additions: zinc transporter (Pfam identifier PF02535); Ctr copper transporter family (PF04145); Cation efflux family (PF01545). To identify candidate genes involved in pathogen defense, we consulted the Innate Immu-



**Figure 1.** Metazoan phylogeny. The Cnidaria are the likely sister to the Bilateria (= Triploblastica). Dates for labeled ancestors are compiled from a number of sources (Aris-Brosou and Yang, 2002; Peterson *et al.*, 2004; Peterson and Butterfield, 2005).

nity Phylogenomic Explorer, ver. 2.0 (Krishnamurthy *et al.*, 2006). We first eliminated those protein families whose functional connection to innate immunity is relatively tenuous—*i.e.*, proteins annotated as “related to” innate immunity or that might “putatively” play a role in innate immunity. This left 128 protein families to which we could ascribe a role in innate immunity with higher confidence. To identify candidate genes involved in wound healing, we culled the recent literature on wound healing and regeneration (Kiritsy *et al.*, 1993; Galko and Krasnow, 2004; Alvarado and Tsonis, 2006; Huxley-Jones *et al.*, 2007).

#### Identifying homologous proteins in *Nematostella*

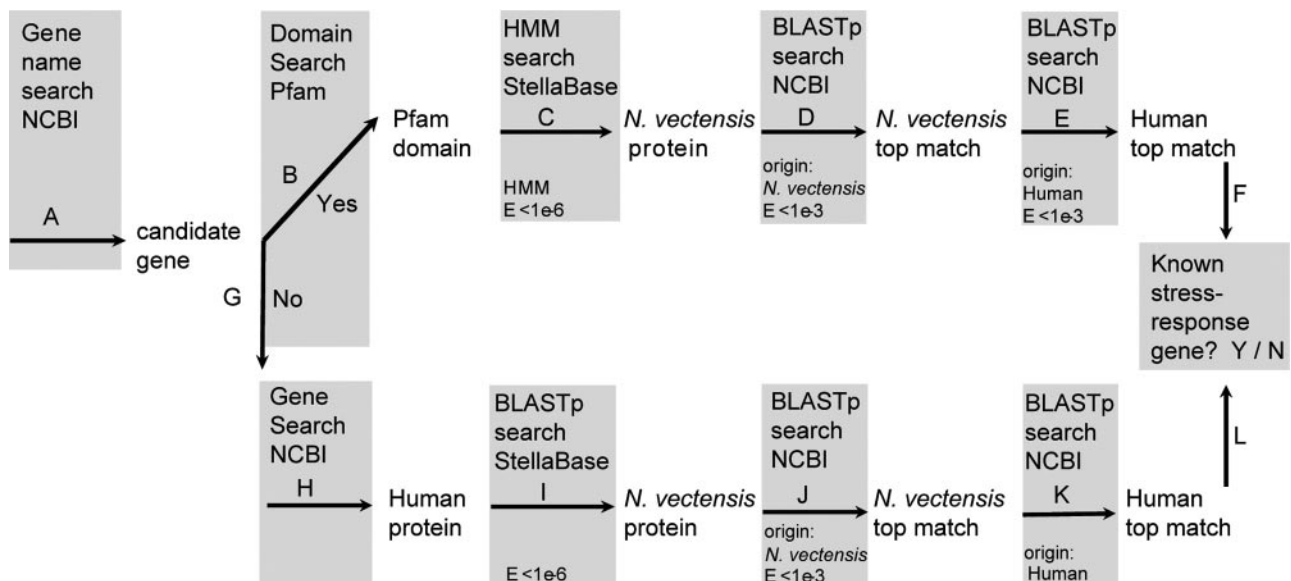
After assembling a list of stress-response query genes from other taxa (Fig. 2a), we sought to identify putative homologs in *Nematostella*. Our principal search strategy utilized conserved protein domains cataloged in the Pfam database, release 17 (Finn *et al.*, 2006). For innate immune genes and chemical defense genes, lists of conserved Pfam domains had already been compiled (Goldstone *et al.*, 2006; Krishnamurthy *et al.*, 2006). For wound-healing genes, we identified conserved domains in the query proteins using the conserved domain search function at Pfam (Fig. 2b). We then identified all *Nematostella* proteins that contained one of these conserved Pfam domains (Finn *et al.*, 2006) by searching a database of predicted proteins at StellaBase 1.0 (Sullivan *et al.*, 2008) using a Hidden Markoff Model search algorithm (Fig. 2c; Durbin *et al.*, 1998). *Nematostella* proteins were scored as possessing a particular Pfam domain if the match to a query sequence received an Expect value  $\leq 1e-6$ . These proteins predicted by using StellaBase were

then cross-referenced with predicted *Nematostella* proteins at NCBI using BLASTp (Fig. 2d). Since this search identifies all proteins matching a particular domain, even those that are not implicated in stress response, the proteins identified in our search were compared against the human Ref Seq database (downloaded 1 Oct. 2007, containing 24306 proteins) using BLASTp to determine whether a known stress-response gene is the best match (Fig. 2e–f).

In cases where no discernible Pfam domain could be located or where only certain members of large gene families are involved in stress response, human homologs of the query sequences were identified in the non-redundant database at NCBI (Fig. 2g–h) and compared against StellaBase using BLASTp (Fig. 2i), again specifying an Expect value cut-off of  $e^{-6}$  to identify putative *Nematostella* homologs. These StellaBase-predicted proteins were then cross-referenced with predicted *Nematostella* proteins at NCBI using BLASTp (Fig. 2j). Each *Nematostella* protein identified as a best match in this search was then compared against the human genome, using tBLASTn to determine whether a known stress-response gene is the best match (Fig. 2k–l).

#### Comparison of protein motifs counts across seven species

We used the Genome Comparison tool in StellaBase to compare the occurrence of each conserved Pfam motif in the sequenced genomes of seven different taxa. The taxa compared were *Escherichia coli*, *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, *Nematostella vectensis*, *Caenorhabditis elegans*, *Drosophila melanogaster*, and *Homo sapiens*. Using the Pfam database (release 17), we identified Pfam domains in each genome by using a hidden Markov model



**Figure 2.** Workflow for identifying stress-response proteins in *Nematostella*, based on Pfam domain searches and BLAST similarity.

search algorithm on the NCBI dataset for each of these species. For these searches, we queried all Pfam motifs for which we recovered at least one matching *Nematostella* protein. When we tallied the results across species, we included all proteins that matched a Pfam motif at an Expect value cut-off of  $1e^{-6}$  (Suppl. Table 1; <http://www.biolbull.org/supplemental/>). To facilitate comparisons among taxa, we calculated the ratio of the number of each Pfam motif found in each taxon divided by the number found in human (Fig. 3).

## Results

Overall, we screened the *Nematostella* genome with 137 different Pfam motifs that are known to be associated with stress-response genes in other taxa. We identified over 2400 *Nematostella* proteins that were significant matches to 123 different Pfam motifs (Suppl. Table 1; <http://www.biolbull.org/supplemental/>). Below, we discuss individual stress-response proteins that we identified or failed to identify, categorized according to the type of stressor.

### Chemical stress response

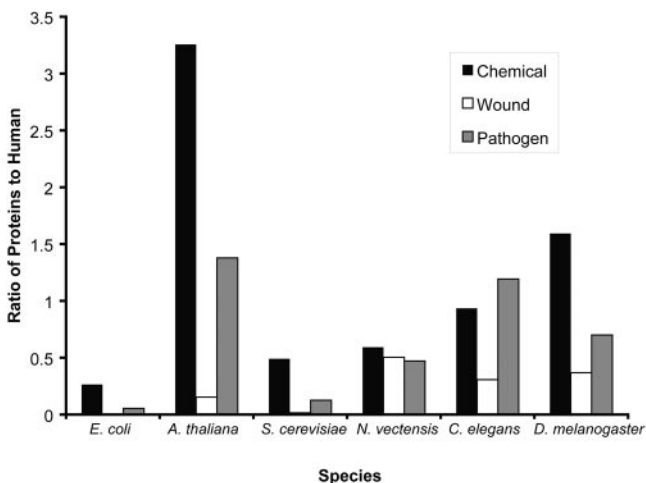
The chemical stressors that impact biodiversity in aquatic ecosystems include a diverse array of naturally and anthropogenically produced compounds (Goldstone *et al.*, 2006). Naturally produced chemical compounds, generated by organisms or by solar radiation, include reactive oxygen species, phytotoxins, and microbial metabolites. Anthropogenic chemical insults include toxic pollutants such as heavy metals, polychlorinated biphenyls (PCBs), pesticides, herbicides, pharmaceuticals, polycyclic aromatic hydrocarbons, flame retardants (polybrominated diphenyl ethers),

and organochlorine compounds (DDT, endosulfan). Estuaries are particularly susceptible to anthropogenic chemical contaminants since they are often located in the vicinity of dense coastal population centers, and their freshwater inputs tend to concentrate runoff and point sources of pollution from upstream sources. At the same time, chemical stressors generated by organisms or by solar radiation can often achieve extremely high levels in shallow estuarine pools that are only infrequently flushed by the tides.

The combination of natural and anthropogenic chemical stressors can have considerable consequences for the biodiversity in estuaries (Pennings and Bertness, 2001; Weis *et al.*, 2004, 2005). Two biotic metrics that reveal much about the health of estuarine ecosystems are (1) the community composition—in particular, the presence and abundance of key indicator species, and (2) the physiological status of resident organisms. Increasingly, molecular and genomic methods are being employed to assay both the community composition of estuaries, particularly the microbial communities, and the physiological status of resident organisms. For example, gene expression assays have recently been employed to characterize the chemical stress response of *Fundulus heteroclitus*, the mummichog, an estuarine fish (Meyer *et al.*, 2005; Schulte, 2007). The mummichog is a highly mobile animal, living in the surface waters of estuarine pools and creeks. It can migrate extensively throughout an estuary to avoid locally unfavorable conditions. However, many animals commonly found in estuaries are sessile, benthic species with limited ability to escape locally stressful conditions. Due to their lack of mobility and their habit of living in or on sediments that tend to accumulate toxicants, the gene expression profiles of such species (*e.g.*, *Nematostella*) are likely to more accurately reflect the recent history of chemical stressors at a particular geographic locale.

*Nematostella* is an ideal estuarine sentinel species for genomic studies on chemical stress. As a sessile benthic animal and a basal metazoan, it can provide results to complement the ongoing work on *Fundulus*, a motile vertebrate. It is an easily collected and common inhabitant of estuaries all along the east coast of North America from Nova Scotia to Georgia and along the Gulf coast. Introduced populations can also be found in England and on the Pacific coast of North America, from central California to British Columbia (Reitzel *et al.*, 2008). It occupies both pristine and highly contaminated habitats. For example, *Nematostella* has been collected in the Hackensack Meadowlands, a restored marsh located along the Hackensack River in northeastern New Jersey. This well-studied site is known to be contaminated with heavy metals (Weis *et al.*, 2004, 2005; Barrett and McBrien, 2007), chlorinated hydrocarbons (Bopp *et al.*, 1998), and pesticides (Barrett and McBrien, 2007).

As *Nematostella*'s apparently low dispersal ability probably renders it unable to escape toxic contamination in the



**Figure 3.** Comparison of the abundance of Pfam motifs associated with responses to chemical stressors, wound, and pathogens in seven model systems. Bars depict the ratio of motifs recovered from each model system relative to the number recovered from human.

sediment through migration, it is likely that some populations have evolved greater tolerance to some chemical contaminants. Population genetic surveys conducted throughout the animal's extensive range have identified sharp genetic breaks between neighboring estuaries, and even between adjacent pools within single estuaries (Darling *et al.*, 2004; Reitzel *et al.*, 2008). Although the swimming larva represents a potential dispersal vehicle, these genetic studies suggest that the effective dispersal potential for this animal may be quite low. We have observed that larvae in culture exhibit positive geotaxis, a behavioral tendency that would contribute to their apparently limited dispersal (Reitzel, Darling, Sullivan, and Finnerty, unpubl. data).

Goldstone *et al.* (2006) characterized the chemical "defense" of the sea urchin. Their classification, which we follow here, included (1) stress-activated receptors, signal transduction proteins, and transcription factors; (2) efflux pumps; (3) oxidizing enzymes; (4) reducing and conjugating enzymes; (5) antioxidant proteins; (6) metal detoxicants; and (7) heat-shock proteins. The first stage in chemical defense is environmental sensing, which involves two principal groups of genes, the PAS (Per-ARNT-SIM) family of transcription factors that respond to oxygen and small molecules, and the nuclear receptor superfamily that bind a variety of ligands, some involved in stress. Both these classes of genes regulate transcription of a variety of effector genes including cytochrome p450s (CYPs), conjugating enzymes, and transporters. Some chemical compounds that breach the cell membrane are removed by efflux proteins such as the ATP Binding Cassette (ABC) and other ion transporters such as the organic anion and cation transporters (OAT and OCT, respectively). Other chemical stressors are biotransformed to inactivate and eliminate them. Biotransformation involves two phases: oxidation, then reduction or conjugation. Oxidation is typically carried out by two families of genes, the flavoprotein monooxygenases (FMO) and the CYPs. After oxidation, compounds are reduced or conjugated by a large suite of gene families including glutathione-S-transferases (GSTs), sulfotransferases (SULTs), and aldo-keto reductases (AKRs). The final group of genes in chemical defense summarized by Goldstone *et al.* (2006) is antioxidant defenses. Reactive oxygen species (ROS), including superoxide, H<sub>2</sub>O<sub>2</sub>, and hydroxyl radical, are a product of both metabolism and exogenous processes (pollutants, ultraviolet radiation, hypoxia). Regardless of their source, ROS affect signal transduction cascades, transcription factors, DNA integrity, and lipids, to name a few, resulting in effects on differentiation, apoptosis, stress-responsive genes, and aging (Adler *et al.*, 1999). A variety of antioxidant proteins including superoxide dismutase (SOD), catalases (CATs), and peroxidases have been identified in diverse metazoans that regulate ROS in the cell.

#### *Stress-activated receptors, signal transduction pathways, and transcription factors*

*Nematostella* possesses a number of receptor and transduction genes involved in chemical defense in sea urchin and human (Table 1). For example, a recent phylogenetic study revealed that most of the gene families in the basic helix-loop-helix (bHLH) superfamily had evolved prior to the cnidarian-triploblast divergence. *Nematostella* was found to possess 68 genes representing 29–32 families, including two proteins that also possess a second domain, the Per-Arnt-Sim, or PAS, domain: (1) the hypoxia inducible factor 1 $\alpha$  (HIF-1 $\alpha$ ) and (2) the aryl hydrocarbon receptor nuclear translocator (ARNT; Simionato *et al.*, 2007). In our search of the *Nematostella* genome, we identified three PAS proteins—ARNT, HIF-1 $\alpha$ , and the aryl-hydrocarbon receptor (AHR). AHR was not identified in the earlier study, and given its presence in *Hydra*, its apparent absence in *Nematostella* was ascribed to a lineage-specific loss in the sea anemone (Simionato *et al.*, 2007). The putative AHR homolog we identified (SB\_56923; gi156394392), whose expression has been confirmed through an expressed sequence tag (EST) (JGI\_CAGN20098.fwd), was not represented in the bHLH proteins culled from *Nematostella* in this previous study. However, a single "AHR related" gene, presumably the same protein, was identified in the publication of the *Nematostella* genome (Putnam *et al.*, 2007). Both AHR and HIF-1 $\alpha$  form complexes with ARNT and then regulate transcription of downstream targets through recognition of xenobiotic-responsive elements (XREs) or hypoxia-responsive elements (HREs), respectively. Further bioinformatic searches for these elements in upstream re-

**Table 1**

*Xenobiotic receptors and conditional transcription factors in Nematostella genome*

Gene Family	Gene name	<i>Nematostella</i> homolog	Human gene ID	<i>E</i> value
bHLH-PAS	<i>AHR</i>	SB_50387	gi 4502003	e-33
	<i>ARNT</i>	SB_15721	gi 2702319	e-35
	<i>Hif-1<math>\alpha</math></i>	SB_27525	gi 4504385	e-32
CNC-bZIP & related	<i>Cnc/NEF2</i>	SB_56664	gi 5453774	e-7
	<i>Keap</i>	SB_8385	gi 32425813	e-53
	<i>Maf</i>	SB_34218	gi 3068761	e-14
		SB_16005		e-12
Nuclear receptors	<i>Hnf4</i>	SB_4250	gi 31077207	e-34
	<i>Rxr</i>	SB_14472	gi 5902068	e-58
Metal & heat response	<i>Mtf</i>	SB_8762	gi 112790163	e-65
	<i>Hsf</i>	SB_766	gi 5031767	e-45
		SB_49214		e-34

gions of potential effector genes would provide a fruitful avenue of research in constructing gene networks.

Our search recovered likely homologs to the bZIP transcription factors NF-E2 and Maf, which heterodimerize and activate gene expression in response to oxidative and xenobiotic stress. Both were recently identified in a number of cnidarians including *Nematostella* (Amoutzias *et al.*, 2007). We also identified a KEAP1, an oxidative stress and electrophile sensor protein that binds NF-E2 in the cytoplasm, preventing NF-E2 translocation to the nucleus when the cell is not under stress (Jaiswal, 2004).

Predicted *Nematostella* proteins also exhibit strong matches to the metal-responsive transcription factor (MTF1) and the heat-shock factor HSF1. Because each of these transcription factors regulates gene expression by binding to well-characterized recognition motifs in target genes—metal-responsive elements for MTF: TGCRCNC (Saydam *et al.*, 2001; Zhang *et al.*, 2001); heat-shock elements for HSF: repeats of AGAAN and its complement recognized by trimers of HSF1 (Orosz *et al.*, 1996)—it will be possible to screen the *Nematostella* genome for candidate target genes.

*Nematostella*, like other cnidarians surveyed (Grasso *et al.*, 2001; Thornton, 2003; Bertrand *et al.*, 2004), lacks many of the nuclear receptors traditionally studied in organismal stress (*e.g.*, estrogen receptors, NR1 family). However, the receptors most relevant to environmental perturbation, hepatocyte nuclear factor 4 (Hnf4) and the retinoid X receptor (RXR), both members of the ancestral nuclear receptor subfamily 2, appear to be present in *Nematostella* (Reitzel and Tarrant, unpubl. data) and other cnidarians (*e.g.*, Grasso *et al.*, 2001). The functional role of these genes awaits further study, but some evidence suggests that cnidarians may be susceptible to a condition resembling endocrine disruption (Tarrant, 2005). Furthermore, understand-

ing the interaction of nuclear receptors and other genes involved in endocrine-like function in *Nematostella* may reveal how the endocrine system evolved (Tarrant, 2007).

#### Efflux pumps

ATP-binding cassette (ABC) superfamily proteins are efflux transporters that pump compounds across cellular membranes against their concentration gradient (Dean and Annilo, 2005). ABC transporters are grouped into eight subfamilies (A–H) of which B, C, and G are known to expel toxic substances (Goldstone *et al.*, 2006). Our search recovered 42 candidate ABC transporters. The ABC transporter complement from *Nematostella* is similar to that of human ( $n = 48$ ), greater than *Ciona* ( $n = 31$ ), and less than urchin ( $n = 65$ ). Phylogenetic studies of these genes are ongoing and will allow a comparison of the ABC gene distribution in each subfamily.

*Nematostella* also has strong matches to other toxicant efflux proteins involved in removal of herbicides and toxic metals such as mercury and cadmium. We observed significant matches to the organic anion transporter polypeptide (OATP) family ( $n = 4$ , SB\_12293, 7514, 12005, and 47307), the organic cation transporter (OCT) family ( $n = 1$ , SB\_18740), and zinc transporters ( $n = 4$ , SB\_34021, 4313, 28001, and 34688).

#### Oxidative biotransformation

*Nematostella* has putative homologs for all but one of the gene families involved in oxidative biotransformation (Table 2). The *Nematostella* genome encodes 48 predicted cytochrome p450s (Table 2, Supp. Table 1; <http://www.biolbull.org/supplemental/>), a gene family involved in oxidation of xenobiotic compounds including polycyclic aromatic hydrocarbons. Two other enzymes involved in

**Table 2**

Comparison of the number of proteins in the *Nematostella*, urchin (*Strongylocentrotus purpuratus*), and human genomes that exhibit significant matches to Pfam domains found in biotransformative genes

Classification	Description	Pfam accession	Version	<i>Nematostella</i>	Urchin*	Human*
Oxidative	CYP	PF00067	11	48	120	57
	FMO	PF00743	8	2	16	6
	ALDH	PF00171	10	13	20	19
	PTGS/COX	PF01124	7	0	0	2
Conjugative	GST	PF00043/02798	5/8	8	38	21
	MGST	PF01124	7	1†	12	3
	SULT	PF00685	14	5	36	13
	UGT	PF00201	8	1	49	13
	NAT	PF00797	7	0	1	2 (10)
Reductive	AKR-like	PF00248	10	2	10	8
	EPHX	PF06441	1	0	5	2

\* Data for urchin and human from Goldstone *et al.*, 2006.

† A search of *Nematostella* expressed sequence tags using BLAST at NCBI resulted in a significant match (GI:156227921).

oxidation of exogenous compounds—flavin-containing monooxygenase (FMO) and aldehyde dehydrogenase (ALDH)—are also present in *Nematostella*'s genome. We did not identify a clear homolog for the prostaglandin-endoperoxide synthase (PTGS/COX) in *Nematostella*. This result is surprising given that PTGS genes have been identified in corals (Varvas *et al.*, 1999; Järving *et al.*, 2004). Although the PTGS protein of the coral *Gersemia fruticosa* is 50% identical to the human version, the best *Nematostella* match is only 23% identical to human PTGS and exhibits equivalent similarity to human peroxidases. Overall, for those gene families known to be involved in oxidative biotransformation, *Nematostella* appears to have fewer genes than human or urchin (Table 2).

*Nematostella* also has a diverse set of genes involved in reductive and conjugative biotransformation. We identified *Nematostella* homologs for the conjugative gene families that transfer glutathione (GSTs,  $n = 8$ ), that metabolize xenobiotics (SULTs,  $n = 5$ ), and that detoxify contaminants by addition of a glycosyl group to form hydrophobic molecules (UGTs,  $n = 1$ ). A larger number of proteins in each of these families was detected with BLASTp searches of the *Nematostella* genome (footnotes in Table 2). We did not detect a *Nematostella* homolog of the NQO-like aldo-keto reductase superfamily or the NATs, which detoxify compounds by acetylation of amines. NQO-like AKRs have only been reported in vertebrates and appear to be a vertebrate-specific gene family (Vasiliou *et al.*, 2006). Our Pfam-based search of predicted proteins at StellaBase failed to identify a microsomal-membrane-bound glutathione transferase (MGST) or a microsomal epoxide hydrolase (EPHX). However, using BLASTp queries against the assembled *Nematostella* genome and ESTs produced by the Joint Genome Institute (Putnam *et al.*, 2007), we identified one

significant MGST also represented among the *Nematostella* ESTs (GI:156227921) and one EPHX (GI: XP\_001622323). The presence of only a single MGST in *Nematostella* may reflect the condition found in the cnidarian-triploblast ancestor; if so, this gene family appears to have undergone a significant expansion in the Triploblastica. Overall, for the shared reductive and conjugative genes, *Nematostella* seems to possess fewer representatives than sea urchin or human, suggesting potential expansion after the cnidarian-triploblast split (Table 2).

#### *Antioxidants and metal complexing*

The *Nematostella* genome appears to encode a number of proteins involved in responding to ROS, including superoxide dismutase (SOD), various peroxidases, and the glutathione pathway (Table 3). Previous research with cnidarians, primarily anthozoans, has characterized the expression of SODs and catalase (CAT) in response to environmental stress (Richier *et al.*, 2003, 2005; Yakovleva *et al.*, 2004; Dash *et al.*, 2006, 2007; Merle *et al.*, 2007).

We also identified putative homologs to the iron-storage protein ferritin, the heme-detoxification enzyme heme oxygenase 1 (HMOX1), and the metal-binding phytochelatins. Interestingly, *Nematostella* appears to lack any member of the evolutionarily conserved metal-binding genes, the class I and II metallothioneins, which are found in diverse eukaryotes (Hamer, 1986) including sponges (Schröder *et al.*, 2000). Metallothioneins are also reported to be absent from *Hydra* (Andersen *et al.*, 1988), suggesting that these proteins may be absent in all cnidarians. Although cnidarians seemingly lack metallothioneins, a few studies have shown that various species accumulate metals in tissue and that species exhibit differential sensitivity to particular metals

**Table 3**

*Pfam domains found in proteins that respond to oxidative damage and metal toxicity that are represented in the Nematostella genome*

Classification	Pfam Name	Description	Pfam Accession	Version	Count
Oxidative damage	Sod_Cu	Superoxide dismutase	PF00080	8	1
	Sob_Fe_N	Superoxide dismutase	PF00081	11	1
	An_peroxidase	Peroxidase	PF03098	5	4
	Catalase	Catalase	PF00199	8	1 <sup>1</sup>
	GSHPx	Glutathione peroxidase	PF00255	9	6
	Thioredoxin	Thioredoxin	PF00085	8	13
Metal complexing	Metallothio	Metallothionein	PF00131	8	0
	Cu-oxidase	Multicopper oxidase	PF00394	11	1
	Transferrin	Transferrin	PF00405	7	1
	Ferritin	Ferritin	PF00210	11	1
	Heme_oxygenase	Heme_oxygenase	PF01126	10	1
	Phytochelatin	Phytochelatin synthase	PF05023	4	2
Glutathione cycling	GCS	Glutamate cysteine ligase	PF03074	5	1
	GSH_synthase	Eukaryotic glutathione synthase	PF03199	5	1
	Glutaredoxin	Glutathione reductase	PF00462	10	4

<sup>1</sup> Significant BLAST match to *Nematostella* predicted protein in GenBank (GI: 156386091).

(Karntanut and Pascoe, 2000, 2002, 2007; Mitchelmore *et al.*, 2003a). Studies on the anemone *Anthopleura elegantissima* have suggested that the accumulation of metals may be influenced by glutathione, a metal-binding antioxidant, as well as by endosymbionts (Mitchelmore *et al.*, 2003a, b). *Nematostella* has a number of glutathione peroxidases (Table 3), enzymes that use glutathione as a cofactor, and glutathione *S*-transferases (Table 2) that catalyze conjugation and reduction reactions with glutathione as a substrate. *Nematostella* is not known to harbor endosymbionts.

### Heat-shock proteins

HSF, identified above, is a transcription factor that regulates expression of the heat-shock proteins (HSPs), which are involved in responses to temperature fluctuations and a wide variety of other stressors. HSPs are categorized into families on the basis of the molecular weight of the protein. *Nematostella* has multiple representatives in the hsp20 (7), hsp70 (6), and hsp90 (3) families. HSPs from various cnidarians have been studied in a variety of stress responses including temperature (Hayes and King, 1995; Kingsley *et al.*, 2003; Schroth *et al.*, 2005), coral bleaching (Downs *et al.*, 2002), and aggression (Rossi and Snyder, 2001).

### Innate immunity and biological stress

Broadly, the immune system of vertebrates with jaws can be divided into adaptive and innate immunity. Through the adaptive immune system, jawed vertebrates are capable of mounting a “learned response” against infectious agents *via* the generation of pathogen-specific antibodies. This complex, acquired response relies on B-cells (humoral system) and T-cells (cell-mediated systems), other lymphocytes, and an ability to model an extensive number of unique antibodies *via* light-chain variability of immunoglobulin proteins.

However, even in those organisms that possess an adaptive immune response, an important component of organismal health is the innate immune system. The innate immune system comprises a number of disparate host-defense mechanisms that interfere with the efficacy of infection by pathogens. The first line of defense, and an integral part of the innate immune system, is a physical barrier that impedes the entrance of infectious agents. Epithelium, mucous membranes, waxy cuticles, shells, and chitinous skeletons perform such a function in phylogenetically diverse organisms. Once an infectious agent gains access, vertebrate hosts may deploy numerous components of the innate immune system, including an inflammatory response, activation of the complement system, deployment of nonspecific macrophages and other leukocytes, production of free radicals and peroxide, and the production of antimicrobial peptides. Although such antimicrobial peptides lack the degree of specificity exhibited by immunoglobulins, they bind cell

structural elements or macromolecules that are specific to major clades of infectious organisms.

A comparison of nonvertebrate deuterostomes (*e.g.*, sea urchin) and basal vertebrates (lamprey, hagfish) reveals that the adaptive immune system is a vertebrate invention (Sodergren *et al.*, 2006). In contrast, the evolutionary origin of the innate immune system remains obscure. Its presence in both deuterostomes (*e.g.*, human and sea urchin) and protostomes (*e.g.*, *Drosophila* sp. and *Limulus polyphemus*) indicates that this system originated prior to the radiation of triploblasts. We found that 63 of the 128 Pfam motifs specifically implicated in innate immune function appear to have homologs in the genome of *Nematostella*. These 63 Pfam motifs identify 1039 unique predicted proteins housed in StellaBase that may be involved in innate immune function (Table 4).

### Toll-receptor-mediated innate immunity

The Toll-Receptor pathway is a central signaling pathway in the innate immune system. Originally identified in *Drosophila*, Toll receptors, and their nine vertebrate homologs, the Toll-like receptors, respond to numerous bacterial cell-wall constituents and the presence of viruses by initiating a signaling cascade that culminates in the activation of the transcription factors NF- $\kappa$ B, AP-1, and/or interferon regulatory factors. These transcription factors then upregulate the expression of effector genes required for the innate response (Armant and Fenton, 2002).

Five Toll-like domains (Pfam motif: TIR) have been identified in the *Nematostella* genome (Miller *et al.*, 2007; Sullivan *et al.*, 2007). Four of these domains are present in transmembrane proteins, suggesting a role in extracellular signaling. Three of these four transmembrane proteins are associated with two or more immunoglobulin motifs, suggesting a structure reminiscent of vertebrate interleukin receptors. The remaining transmembrane-linked TIR motif is associated with numerous leucine-rich motifs, which, when combined with phylogenetic analyses, suggests a structure conserved with true Toll- and Toll-like receptors in fly and vertebrates. The final TIR domain is not associated with a transmembrane motif, but rather with a DEATH domain, indicating that this protein may act as the first step in the intracellular signaling cascade (*e.g.*, it may be a MyD88 homolog). Like *Nematostella*, *Hydra magnipapillata* was also found to possess predicted proteins combining Toll-like receptor domains in association with transmembrane motifs (Zheng *et al.*, 2005).

Other key elements of the innate immune system have also been identified in *Nematostella*, including several members of the Toll and interleukin signaling cascades such as IRAK, TRAF, ECSIT, IKK, MKK, JNK, IKB, p38, and the transcription factors NF- $\kappa$ B and AP-1 (Table 4; Miller *et al.*, 2007;



Table 4

Pfam domains associated with innate immunity that are present in *Nematostella vectensis*

Pfam name <sup>1</sup>	Description	Accession	Version <sup>2</sup>	Count <sup>3</sup>
AMP-binding	AMP-binding enzyme	PF00501	14	16
Ank	Ankyrin repeat	PF00023	16	104
Asp	Eukaryotic aspartyl protease	PF00026	13	1
BIR	Inhibitor of Apoptosis domain	PF00653	10	3
BTB	BTB/POZ domain	PF00651	18	42
bZIP_1	bZIP transcription factor	PF00170	10	1
bZIP_2	Basic region leucine zipper	PF07716	3	5
C_tripleX	Cysteine rich repeat	PF02363	8	1
CARD	Caspase recruitment domain	PF00619	10	7
cNMP_binding	Cyclic nucleotide-binding domain	PF00027	16	11
DEAD_2	DEAD_2	PF06733	5	1
Death	Death domain	PF00531	11	6
DSL	Delta serrate ligand	PF01414	8	9
Dsrn	Double-stranded RNA binding motif	PF00035	11	8
EGF	EGF-like domain	PF00008	14	121
EGF_2	EGF-like domain	PF07974	1	44
EGF_CA	Calcium binding EGF domain	PF07645	4	91
F-box	F-box domain	PF00646	19	3
FYVE	FYVE zinc finger	PF01363	11	9
Gln-synt_C	Glutamine synthetase, catalytic domain	PF00120	12	1
Glyco_hydro_1	Glycosyl hydrolase family 1	PF00232	8	2
Glycolytic	Fructose-bisphosphate aldolase class-I	PF00274	8	2
HMG_box	HMG (high mobility group) box	PF00505	8	14
Ig	Immunoglobulin domain	PF00047	14	76
Ion_trans	Ion transport protein	PF00520	17	53
Laminin_EGF	Laminin EGF-like (Domains III and V)	PF00053	12	17
LEA_4	Late embryogenesis abundant protein	PF02987	6	1
Lipoxygenase	Lipoxygenase	PF00305	7	2
LRR_1	Leucine Rich Repeat	PF00560	19	35
MATH	MATH domain	PF00917	14	6
MMR_HSR1	GTPase of unknown function	PF01926	9	5
NACHT	NACHT domain	PF05729	2	2
NMT	Myristoyl-CoA:protein N-myristoyltransferase, N-terminal	PF01233	8	3
NMT_C	Myristoyl-CoA:protein N-myristoyltransferase, C-terminal	PF02799	4	1
Pam16	Pam16	PF03656	3	2
PAZ	PAZ domain	PF02170	10	1
Pentapeptide	Pentapeptide repeats (8 copies)	PF00805	11	2
Piwi	Piwi domain	PF02171	7	4
Pkinase	Protein kinase domain	PF00069	13	85
Pkinase_Tyr	Protein tyrosine kinase	PF07714	3	88
PLAT	PLAT/LH2 domain	PF01477	9	17
PP2C	Protein phosphatase 2C	PF00481	10	5
PSI	Plexin repeat	PF01437	12	2
RCC1	Regulator of chromosome condensation (RCC1)	PF00415	8	8
RdRP	RNA dependent RNA polymerase	PF05183	2	4
Ribonuclease_3	RNase3 domain	PF00636	12	2
SAM_1	SAM domain (Sterile alpha motif)	PF00536	16	12
SAM_2	SAM domain (Sterile alpha motif)	PF07647	4	13
SapB_1	Saposin-like type B, region 1	PF05184	4	1
SapB_2	Saposin-like type B, region 2	PF03489	4	2
TIR	TIR domain	PF01582	8	2
TPR_1	Tetratricopeptide repeat	PF00515	14	74
TPR_2	Tetratricopeptide repeat	PF07719	3	74
TPR_4	Tetratricopeptide repeat	PF07721	2	2
Tsg101	Tumor susceptibility gene 101 protein (TSG101)	PF05743	2	1
Ubiquitin	Ubiquitin family	PF00240	12	5
UCH	Ubiquitin carboxyl-terminal hydrolase	PF00443	16	9
WAP	WAP-type (Whey Acidic Protein) 'four-disulfide core'	PF00095	10	3
WD40	WD domain, G-beta repeat	PF00400	18	83
zf-C2H2	Zinc finger, C2H2 type	PF00096	14	111
zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	PF00097	11	11
zf-TRAF	TRAF-type zinc finger	PF02176	7	5
zf-UBP	Zn-finger in ubiquitin-hydrolases and other protein	PF02148	7	1

<sup>1</sup> Pfam motifs may be retrieved by placing the accession after the following URL: <http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?> (e.g., <http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00501> will retrieve information for Pfam motif 'AMP-binding.'

<sup>2</sup> 'Version' refers to the version / update of the Pfam motif utilized in the hidden Markoff model search.

<sup>3</sup> 'Count' refers to the number of predicted proteins in StellaBase, ver. 1.0, which match the hidden Markoff model of each domain with an *E*-value less than or equal to 1e-6.

Sullivan *et al.*, 2007). Additionally, the transcription Interferon-regulatory factor, also inducible by the Toll-like signaling cascade, may have a homolog in this basal animal (SB 44157, Genbank Accession gil156396757reflXP\_001637559.1; BLASTp *E*-value versus human interferon regulatory factor 2 [gil153082752reflNP\_002 190.2] =  $5e^{-40}$ ). Interestingly, although *Nematostella* was found to possess a complex repertoire of innate immune genes, a relatively reduced complement of innate immune genes was identified in the freshwater hydrozoan *Hydra magnipapillata* (Miller *et al.*, 2007).

In cnidarians, it has been suggested that Toll-like receptors might be performing an ancestral function unrelated to innate immunity (Kanzok *et al.*, 2004; Zheng *et al.*, 2005). Kanzok *et al.* (2004) argued that animals lacking a coelom, including all diploblastic animals such as cnidarians, would not benefit from production of anti-microbial peptides—a hallmark of Toll-receptor-mediated innate immunity—because diploblasts lack an internal body compartment (namely, the coelom) into which these proteins could be secreted. For a time, this hypothesis was bolstered by an inability to recover an NF- $\kappa$ B or other Rel-homology domain containing transcription factor from any basal animal.

Now that many of the components of innate immunity have been identified in cnidarians, we can begin to assess whether they are performing a role homologous to the role in triploblastic animals. If functional Toll- and Interleukin-signaling cascades are indeed present in *Nematostella*, downstream targets of the cascade may be conserved between cnidarians and bilaterians (*e.g.*, canonical targets of NF- $\kappa$ B). Homologs of known targets of NF- $\kappa$ B in vertebrates have been identified in *Nematostella* through sequence homology (Table 4; Sullivan *et al.*, 2007). Downstream components present in *Nematostella* include numerous apoptosis regulators (*i.e.*, caspases [Eckhart *et al.*, 2007], p53, and p63), domains associated with antimicrobial peptides, antioxidant enzymes, and a putative homolog of a complement component protein (Table 4). We are currently searching the 5' regions of these putative targets to identify loci that are enriched for  $\kappa$ B sites, and therefore potentially regulated by NF- $\kappa$ B, perhaps *via* the Toll-signaling cascade.

#### *C3-related complement system proteins*

Among cnidarians, a homolog for the complement system protein C3 was first identified in the gorgonian *Swiftia exserta* (Dishaw *et al.*, 2005). In *Nematostella*, a systematic search of the genome for components of the complement pathway identified a C3 protein in addition to integrins, suggesting that a simplified complement system may be present (Nonaka and Kimura, 2006). Such a system, however, would have to be extremely simplified relative to the complement system of vertebrates, whose complexity re-

quired a number of gene and genome duplications over the course of vertebrate evolution. Interestingly, *Hydra magnipapillata* appears to lack a C3 protein (Miller *et al.*, 2007), although evolutionarily related A2M-domain-containing proteins are expressed in the endoderm of *Hydra* with an expression pattern similar to that of a C3 protein of the coral *Acropora millepora*.

#### **Wound healing**

Along with chemical insults and infection by pathogens, physical wounding is a major source of organismal stress. Indeed, different categories of stressors interact to threaten the well-being of organisms: for example, pathogens may be introduced upon wounding, and the inability to recover from wounds may be the proximate cause of death in animals that are already stressed by chemical insults or pathogens. All animal phyla exhibit some ability to recover from wounds. In those animals where it has been studied in detail, wound healing is a complex and dynamic process that requires the interaction of many factors including cytokines, intracellular signaling pathways, transcription factors, and extracellular matrix (ECM) proteins. The extent to which the wound-healing process of “basal animals” might be homologous to the wound healing of triploblastic animals is not yet known.

We expect that some genes in the wound-healing repertoire of cnidarians may be unique to this phylum and thus identifiable only through empirical studies on cnidarians. However, as the “epithelial level of organization” exhibited by *Nematostella* is an extremely ancient invention that dates to the common ancestor of the Eumetazoa (if not earlier), it is likely that epithelial repair mechanisms are similarly ancient, and core elements of these pathways may be deeply conserved across animals. If so, searching for homologs of known wound-healing genes should be useful to identify candidate wound-healing genes in *Nematostella*.

#### *Wound healing in vertebrates*

Traditionally in vertebrate models, the process of wound healing has been broken down into three distinct but overlapping phases (Singer and Clark, 1999): (1) inflammation, (2) proliferation, and (3) matrix rebuilding and remodeling. During the inflammatory stage, the coagulation pathway is activated and platelets appear at the wound site (Kiritsy *et al.*, 1993). Under the influence of cytokines such as platelet-derived growth factor (PDGF) and transforming growth factor beta (TGF- $\beta$ ), these platelets form a plug at the site of injury (Franchini and Ottaviani, 2000). This phase is characterized by the activation of neutrophils, macrophages, and mast cells, which help to remove debris and bacteria from the wound site. Simultaneously, cytokines stimulate circulating cells to express integrins, and the integrins enable these cells to interact with the ECM. The proliferative

phase is marked by the presence of fibroblasts, which produce a number of ECM components including collagen, proteoglycans, and fibronectin. Keratinocytes proliferate and migrate out to the edge of the wound site, creating a natural defense barrier. Additional cytokines are produced from fibroblasts, keratinocytes, and macrophages, and these intercellular signaling factors help to coordinate this complicated series of events. Lastly, the ECM is remodeled by the modification of collagen fibers, which results in the contraction of the wound. Collagen reorganization appears to be a very ancient component of metazoan wound healing shared among protostomes and deuterostomes (Tettamanti *et al.*, 2005).

#### *Wound healing in cnidarians*

In cnidarians, wound healing is often followed by regeneration, and these two processes tend to be conflated in the cnidarian literature (Henry and Hart, 2005). However, two lines of evidence reveal that wound healing and regeneration are evolutionarily and mechanistically separable processes. First, there are many animal species whose regenerative abilities are limited but whose wound-healing abilities are quite well developed—for example, if the tail of a mouse is surgically removed, the wound will heal, but the tail will not regenerate. Second, even in animals with extensive regenerative ability, such as cnidarians, wounds that do not result in the loss of a body region or structure do not necessarily trigger regeneration. In *Nematostella*, for example, complete bisection through the body column always results in the regeneration of a missing oral crown or a missing physa (Reitzel *et al.*, 2007). However, incomplete bisection of the body column only rarely triggers the development of an ectopic oral crown or physa. Far more often, it triggers wound healing without regeneration. Here, we focus on wound healing and not regeneration, because the former is clearly a response to acute stress.

In cnidarians, the cellular events underlying wound healing have not been characterized at the same level of detail as in vertebrates, but at least one potential parallel has been noted. In vertebrates, the inflammatory stage is characterized by the influx of specialized cells to the wound site, including phagocytes (Gillitzer and Goebeler, 2001). Similarly, in some cnidarians, including gorgonians, scleractinian corals, and sea anemones, amoebocytes are recruited from the mesoglea to the wound site, and these cells engage in phagocytosis (Hutton and Smith, 1996; Meszaros and Bigger, 1999; Olano and Bigger, 2000; Henry and Hart, 2005; Mydlarz *et al.*, 2006). These amoebocytes that migrate to the wound site also produce ROS (Hutton and Smith, 1996), as occurs in the circulating amoebocytes of other invertebrates (Nakamura *et al.*, 1985; Coteur *et al.*, 2002), suggesting that these cells are playing a role in an inflammatory response.

#### *Identifying homologs of known wound-healing genes in Nematostella*

In a first step toward determining whether the wound healing process in cnidarians is homologous to that of vertebrates, we can determine whether the same genes are involved in coordinating and regulating the process. Although a complex network of genes underlies the process of wound healing in vertebrates, none of the genes so far identified appears to be uniquely involved in wound healing. Thus, it is not possible to strictly categorize any gene as a “wound healing gene.” However, we can determine which genes from the entire complement of genes active during wound healing in vertebrates are present in *Nematostella*, and subsequent investigations can determine whether these putative homologs are deployed in a homologous fashion during wound healing in the anemone.

The genes involved in vertebrate wound healing can be broken down into the following broad functional categories: (1) intercellular signaling molecules; (2) cell surface receptors and intracellular signaling molecules, (3) transcription factors, and (4) ECM components. On the basis of a survey of the literature, we identified 31 Pfam domains in proteins associated with wound healing, including signaling molecules, cell surface receptors, intracellular signaling factors, transcription factors, and ECM components. Every one of these Pfam domains scored a significant match to at least one predicted protein in StellaBase (Table 5, Suppl. Table 1; <http://www.biolbull.org/supplemental/>). Overall, the 31 domains scored significant matches to 485 different anemone proteins.

Another 26 genes involved in vertebrate wound healing could not be associated with particular Pfam domains. The human homolog for each of these genes was retrieved from NCBI and a BLAST search was conducted against predicted *Nematostella* genes at NCBI to identify putative homologs in the sea anemone (Table 6). These human proteins included 22 ECM components in addition to TGF-beta receptor and Jun N-terminal kinase (JNK). Most of these human genes matched one or more *Nematostella* sequences with an Expect value at or below the threshold of  $e^{-6}$ , and in 10 cases, a reciprocal tBLASTn search, back against the human genome, provided evidence of genuine orthology between the human and sea anemone proteins (Table 6). Below, we discuss particular wound-healing genes that appear to have homologs in *Nematostella*.

#### *Intercellular signaling proteins involved in wound healing—the cytokines*

In vertebrates, cytokines function in a number of important biological processes including chemostasis and cell proliferation (Faler *et al.*, 2006), and they play key roles during both the inflammatory and proliferative stages of wound healing. Our search of the *Nematostella* genome

Table 5

*Pfam domains associated with wound healing that are represented in the Nematostella genome*

Classification	Pfam name <sup>1</sup>	Pfam description	Pfam number	Count <sup>2</sup>	Version <sup>3</sup>	
Extracellular matrix components	Collagen	Collagen triple helix	PF01391	25	7	
	Fn2	Fibronectin type II	PF00040	9	8	
	Fn3	Fibronectin type III	PF00041	107	10	
	TB	TB domain	PF00683	1	7	
	glypican	glypican	PF01153	1	9	
	Laminin_EGF	Laminin EGF-like	PF00053	17	12	
	Laminin_G_1	Laminin G domain	PF00054	14	12	
	Laminin_G_2	Laminin G domain	PF02210	17	10	
	Laminin_B	Laminin B (Domain IV)	PF00052	1	8	
	Laminin_N	Laminin N-terminal	PF00055	5	7	
	NIDO	Nidogen-like	PF06119	1	2	
	LRR_1	Leucine Rich Repeat	PF00560	55	19	
	LRRNT	Leucine Rich Repeat N	PF01462	1	8	
	TSP_1	Thrombospondin type 1	PF00090	54	8	
	TSP_3	Thrombospondin type 3	PF02412	6	7	
	TSP_C	Thrombospondin C-	PF05735	4	1	
	hemopexin	Hemopexin	PF00045	3	8	
	Somatomedin_B	Somatomedin B domain	PF01033	3	7	
	VWA	von Willebrand factor	PF00092	98	14	
	VWC	von Willebrand factor	PF00093	1	7	
	VWD	von Willebrand factor	PF00094	2	13	
	Signaling molecules and receptors	TGF_beta	Transforming growth	PF00019	5	8
		Activin_recp	Activin type I and II	PF01064	3	11
FGF		Fibroblast growth factor	PF00167	3	8	
Lectin_leg_like		Legume-like lectin	PF03388	1	3	
Gal_lectin		Galactose binding lectin	PF02140	25	7	
Lectin_C		Lectin C-type domain	PF00059	24	9	
TNF		TNF (Tumor Necrosis)	PF00229	1	8	
Integrin_beta		Integrin, beta chain	PF00362	3	7	
FG-GAP		FG-GAP repeat	PF01839	7	11	
Integrin_B_tail		Integrin beta tail domain	PF07965	2	7	
Transcription factors	CP2	CP2 domain	PF04516	1	4	
	MH2	MH2 domain	PF03166	1	4	
	MH1	MH1 domain	PF03165	4	6	

<sup>1</sup> Pfam motifs may be retrieved by placing the accession number after the following URL: <http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?> (e.g., <http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00501> will retrieve information for Pfam motif 'AMP-binding').

<sup>2</sup> 'Count' refers to the number of predicted proteins in StellaBase, ver. 1.0, which match the hidden Markoff model of each domain with an *E* value less than or equal to  $1e-6$ .

<sup>3</sup> 'Version' refers to the version / update of the Pfam motif utilized in the hidden Markoff model search.

using a "transforming growth factor beta-like" domain (PF00019) identified five TGF-beta ligands. Two TGF-beta factors were previously reported in this species, including a GDF5-like gene and a DPP-like gene; on the basis of their developmental expression, they are thought to be involved in axial patterning (Finnerty *et al.*, 2004). Both of these genes were found in our Pfam domain search, as well as other TGF-beta factors identified previously (Matus *et al.*, 2006).

There are two types of TGF-beta receptors: type I and type II. In the absence of a ligand, these receptors exist as homodimers on the surface of the cell, but they typically bind to ligands as heterotetramers (Derynck and Zhang, 2003). Activation of this signaling cascade by the binding of

a TGF-beta ligand to a receptor complex can result in the nuclear translocation of SMAD proteins, which can impact cell proliferation and migration. Also, the TGF-beta signaling pathway can be Smad-independent and result in the activation of MAPK pathways, some of which have a role in the Smad activation (Derynck and Zhang, 2003). In *Nematostella*, we identified a single TGF-beta receptor that best matches human TGF-beta receptor I.

### Integrins

Integrins are a superfamily of integral membrane proteins that function in cell migration, cell-cell adhesion, cell-ECM adhesion, and signal transduction (Takada *et al.*, 2007). In

Table 6

Reciprocal tBLASTn searches (human vs. anemone) using human proteins implicated in wound healing as the initial query

Human query	ID	<i>Nematostella</i> hit	E value	Reciprocal human hit	E value	Human gene name
<b>Signaling</b>						
Integrin, alpha 8	NM_003638	XM_001641385	1e-79	NM_003638	1e-74	integrin, alpha 8
Integrin, alpha 4	NM_000885	XM_001641385	1e-74	NM_003638	1e-74	integrin, alpha 8
TGF-beta type I receptor	AAD02042	XM_001622484	7e-138	NM_004612	9e-136	TGF-beta type I receptor
<i>JNK1-B1</i>	NM_139046	XM_001637538	3e-133	HSU35004	4e-131	<i>JNK1-B1</i>
<b>SLRPs (small leucine-rich-repeat proteins)</b>						
Biglycan	AAA52287	XM_001639122	9e-17	AY166584	5e-25	Vasorin
Chondroadherin	AAK51556	XM_001639122	4e-22	AY166584	5e-25	Vasorin
Decorin	AAB60901	XM_001633690	2.7	AC104667.5	6e-44	BAC clone RP11-810D14
Fibromodulin	CAA53233	XM_001631425	7e-14	NM_020873	1e-25	leucine rich repeat neuronal 1
Keratocan	NP_008966	XM_001625510	0.0	NP_002009	0.0	Flightless I
Lumican	AAA91639	XM_001640230	1e-16	NM_017768	8e-34	leucine rich repeat cont. 40
Opticin	CAB53459	XM_001623719	4e-08	NM_012293	6e-32	peroxidasin homolog
Osteoglycin	CAI16695	XM_001638871	0.15	AY280614	2e-19	synleurin mRNA
<i>PRELP</i>	CAG47066	XM_001630052	2e-15	NM_018214	3e-43	leucine rich repeat cont. 1
<b>Other extracellular matrix proteins</b>						
Agrin	CAI15575	XM_001632399	3e-75	NM_198576	1e-71	agrin
Asporin	CAI16697	XM_001638871.1	3e-19	AY280614	2e-19	synleurin
Dermatan Sulf. Proteoglycan	AAC50945	XM_001623719.1	1e-09	XM_001127072	6e-32	Peroxidasin homolog
Dystroglycan	AAH12740	XM_001621519.1	8e-25	AK291692	5e-21	similar to Dystroglycan I
Fibromodulin	CAA51418	XM_001629115	2e-15	BC117180	3e-166	Leucine-rich repeat kinase 2
Fibronectin 1	NM_212476	XM_001635202.1	1e-26	NM_019064.3	2e-61	Sidekick 2
Hemicentin	NP_114141.2	XM_001625345.1	4e-116	NM_031935	1e-99	Hemicentin 1
Matrilin 1	NP_002370	XM_001618434	1e-42	BC131710	2e-40	Matrilin 1
Matrilin 2	AAH10444	XM_001636411	2e-91	AB209735	3e-107	Fibrillin 2
Netrin 1	NP_004813	XM_001628849	3e-158	NM_004822	4e-160	Netrin 1
Netrin G1	NP_055732	XM_001628514	1e-145	BC113455	7e-119	Laminin, beta 1
Osteomodulin	NP_005005	XM_001638987	4e-10	BC128989.1	2e-118	Leucine rich repeat cont. 7
Osteonectin	AAA60993	XM_001626392	1e-09	NM_003118	2e-08	Osteonectin
Podocan	AAP79898	XM_001631425	6e-16	NM_130830.2	1e-25	Leucine rich repeat cont. 15
Reelin	AAC51105	XM_001635471	2e-04	NM_001410	0.0	multiple EGF-like-domains 8
Syndecan 1	EAX00831	XM_001629466	0.24	HSU78181	3e-25	Human sodium channel 2
Syndecan 2	NM_002998	XM_001637329	1.8	AF163151	1e-22	Dentin sialophosphoprotein
Syndecan 3	NP_055469	XM_001629466	1.9	HSU78181	3e-25	Human sodium channel 2
Syndecan 4	NM_002999	XM_001641418	0.12	BC020057	2e-160	Integrin, beta 1
Tenascin	CAA39628	XM_001638256	6e-64	XM_001133349	0.0	Notch homolog 2
Vitronectin	EAW51082	XM_001633180	1e-13	BC064803	7e-62	matrix metalloproteinase 14
von Willebrand fac.	NM_000552	XM_001636212	1e-118	XM_001130382	2e-152	similar to Mucin-5AC

the cell, integrins are found as heterodimers, consisting of different alpha and beta subunits. The extracellular domains of integrins interact with a variety of ligands depending on the combination of receptor units. The cytoplasmic domain can interact with signaling proteins, focal adhesions, and cytoskeletal proteins, and it is through this interaction that integrin-initiated intracellular signaling occurs (Takada *et al.*, 2007). It has been proposed that integrins can play a role in the nuclear translocation of the transcription factor NF- $\kappa$ B, as well as in activation of the Jun-N-terminal Protein Kinase (JNK) (Nikolopoulos *et al.*, 2005). The JNK pathway is activated during the proliferative phase of wound healing in vertebrates and during epidermal wound repair in *Drosophila* (Rämet *et al.*, 2002; Galko and Krasnow, 2004).

Both alpha and beta subunits of integrins were previously

cloned in the jellyfish *Podocoryne carnea* (Reber-Muller *et al.*, 2001). The beta subunit has also been cloned from sponge and coral (Brower *et al.*, 1997). A search of the *Nematostella* genome using the “integrin, beta chain” and “integrin beta tail domain” identifies two integrin beta loci. A tBLASTn search using human integrin alpha proteins identifies two integrin alpha loci in *Nematostella*.

#### The transcription factor, grainyhead

*Grainyhead* plays a role in maintaining the tension of the cuticle in *Drosophila*, and similarly, it is implicated in the epithelial integrity of mammals, where a mutation in this gene causes defects in wound healing and epithelial barrier formation (Uv *et al.*, 1994; Kudryavtseva *et al.*, 2003; Mace

*et al.*, 2005; Ting *et al.*, 2005). *Grainyhead* can be recognized by the possession of a highly conserved CP2 domain, and this transcription factor has been identified in a wide range of triploblastic animals (Moussian and Uv, 2005). A search of the *Nematostella* domain using the “CP2” domain identified a single significant match. Reciprocal tBLASTn searches between the human genome and the sea anemone genome identifies this predicted protein as the *Nematostella* ortholog of human *grainyhead*.

### Smads

Smads are transcription factors that are downstream of TGF-beta signaling. They are implicated in cell proliferation and migration. In general, all smads share sequence similarity in the amino-terminal and carboxy-terminal domains, MH1 and MH2 respectively (Xu, 2006). Searching the *Nematostella* genome for the MH1 and MH2 domains identifies four significant matches to MH1 and one to MH2. According to a BLASTp search, four of the five MH1-containing proteins and the single MH2-containing protein are matches to human Smads. The remaining MH1-containing protein matches a human BMP.

### Extracellular matrix proteins important in wound healing

During the proliferative stage of wound healing, an epithelium begins to form over the wound site, and the ECM is deposited. The ECM is a complex netting of glycoproteins, collagens, and proteoglycans. It is important for cellular remodeling during wound healing, and it is implicated in cell attachment and growth, cellular differentiation, and structural support of tissues (Huxley-Jones *et al.*, 2007). Many ECM proteins exhibit strong evolutionary conservation across the animal kingdom. This high degree of conservation is thought to show the importance of collagen-based extracellular matrices in animals (Tettamanti *et al.*, 2005; Aouacheria *et al.*, 2006; Huxley-Jones *et al.*, 2007).

### Collagens

The “collagen triple helix repeat” motif was a significant match for 19 predicted proteins in the *Nematostella* genome. When a BLAST search was performed against the human genome, these candidate *Nematostella* collagens exhibited best matches to multiple human collagen genes, including alpha 1 collagens (types V, XIII, XV, XXV), an alpha 2 collagen (type V), and an alpha 5 collagen (type IV). The BLAST searches also identified matches to a human contactin 4 (also known as BIG-2) and contactin 1 associated protein. The contactins are neuronal cell adhesion molecules that belong to the immunoglobulin superfamily (Yoshihara *et al.*, 1995).

### Fibronectins

Searches using the “fibronectin type II domain” identified nine matches in the *Nematostella* genome, though none of these identified a fibronectin as a best match in the human genome. For example, three *Nematostella* hits to “fibronectin type II” were best matches to human matrix metalloproteinases (MMP-2 and MMP-9), enzymes involved in the breakdown of extracellular matrix proteins. MMP-2 and MMP-9 are unique among metalloproteinases in harboring three repeats homologous to fibronectin II (Steffensen *et al.*, 2001). The six additional *Nematostella* loci that exhibit significant similarity to the “fibronectin type II” domain are best matches to human notch 2 preprotein, relaxin, vitrin, neuronal pentaxin, neuropilin, and semaphorin.

Searches using the “fibronectin type III domain” identified 101 significant matches in the *Nematostella* genome. Two of the genes identified are best matches to the same human fibronectin type III domain containing 3B gene, but none of the best matches are to human fibronectins. A large fraction of the *Nematostella* genes that were identified in the search for fibronectin type III domains are best matches to human protein tyrosine phosphatase receptors.

In an attempt to identify a *bona fide* fibronectin protein in *Nematostella*, we used tBLASTn to query the genome with the human fibronectin 1. When the best hit in the *Nematostella* genome was used to search the human genome, the best human match proved to be a sidekick protein (SDK2), not fibronectin 1 (Table 6).

### Small leucine-rich-repeat protein family

The small leucine rich repeat proteins (SLRPs) are a family of proteoglycans that are known to play important roles in organizing structural elements in the ECM, thereby impacting cell adhesion, wound healing, and the ability to withstand mechanical stress (Matsushima *et al.*, 2000). The proteins are grouped into four classes based on the number and type of repeats and super-repeats they possess. Class I (*e.g.*, decorin and biglycan) and class II (*e.g.*, fibromodulin, lumican, keratocan, and PRELP) have 12 LRRs. Class III SLRPs (*e.g.*, opticin and osteoglycin) have 7 LRRs. Class IV consists of chondroadherin. In vertebrates, specific LRR proteins are involved in distinct extracellular matrix types. For example, biglycan and decorin (class I) act in synergy to promote skin and bone integrity in mice (Corsi *et al.*, 2002), while the class II protein lumican plays a role in wound healing in the cornea (Kao *et al.*, 2006).

Leucine-rich repeats are found in a wide range of proteins, including bacterial virulence factors, cell adhesion molecules, enzymes, hormone receptors, and tyrosine kinase receptors, in addition to extracellular matrix-binding glycoproteins. A search of the *Nematostella* genome using the “leucine-rich-repeat” domain identified 54 matching genes, but none of these genes appears to be orthologous to

any of the vertebrate SLRPs. When the human genome was queried using these 54 anemone sequences, the top BLAST hits included a number of predicted “leucine rich repeat containing” proteins but no members of the SLRP family. Using tBLASTn, we queried the *Nematostella* genome with human SLRP proteins (decorin, biglycan, fibromodulin, lumican, keratocan, PRELP, opticin, osteoglycin, and chondroadherin). We then queried the human genome with the top BLAST hits from *Nematostella*, and not one of these reciprocal BLAST searches identified the human SLRP protein originally used to query *Nematostella* as the top hit. In all cases, other human LRR-containing proteins were identified (Table 6). These same SLRPs also seem to be lacking in protostome animals (Table 7), suggesting that they are deuterostome or vertebrate inventions (Huxley-Jones *et al.*, 2007). The complexity of this gene family in vertebrates may be correlated with the diversity of vertebrate ECM types (cartilage, bone, tendon, cornea, tooth, *etc.*)

#### Other extracellular matrix genes

On the basis of reciprocal tBLASTn searches (Table 6), we were able to identify *Nematostella* orthologs of the human ECM proteins matrilin 1, netrin 1, osteonectin, and von Willebrand factor, but not reelin, syndecan-1, syndecan-2, syndecan-3, syndecan-4, tenascin, or vitronectin. A syndecan was previously reported in *Nematostella* (Chakravarti and Adams, 2006). A single predicted protein in StellaBase (SB\_50645) does exhibit greater resemblance to syndecan than to any other protein family represented in Pfam. However, the resemblance is weak ( $E$ -value = 0.02), and it is based primarily upon similarities in the short “cytoplasmic domain 1” (C1). Within this domain, the predicted *Nematostella* protein (RLRKRDEGSY) is identical to human syndecans (RMK/RKKDEGSY) at 7 or 8 of 10 residues (Chakravarti and Adams, 2006). However, a tBLASTn search of human genome using the anemone protein as a query did not return any hits. This suggests that

Table 7

Reciprocal tBLASTn searches (human vs. protostome) using human genes implicated in wound healing in the initial query

Extracellular matrix proteins	ID	Protostome hit	$E$ value	Reciprocal human hit	$E$ value	Human gene name
Human query SLRPs (small leucine-rich-repeat proteins)						
Biglycan	AAA52287	XM_001649763	3e-21	NM_003061	1e-44	<i>SLIT1</i>
Chondroadherin	AAK51556	XM_970227	3e-26	AK125112	2e-39	Similar to platelet glycoprotein
Decorin	AAB60901	No match	n/a	n/a	n/a	V precursor n/a
Fibromodulin	CAA53233	AJ549813	9e-21	NM_018490	1e-121	Leucine-rich-repeat-containing G protein-coupled receptor 4
Keratocan	NP_008966	XM_001650288	8e-16	BC101065	1e-40	Leucine-rich-repeat cont. 15
Lumican	AAA91639	XM_395331	3e-17	BC094737	9e-31	Toll-like receptor 3
Opticin	CAB53459	XM_001606218	4e-11	BC094737	1e-28	Toll-like receptor 3
Osteoglycin	CAI16695	NM_206253	4e-04	EF090903	0	VPO mRNA
<i>PRELP</i>	CAG47066	XM_001360484	2e-15	BC126169	1e-105	Leucine-rich repeats & immunoglobulin-like domains
Other extracellular matrix proteins						
Asporin	CAI16697	XM_967172	2e-28	AF055585	0.0	Slit2
Dermatan Sulf. Proteoglycan	AAC50945	XM_001650288	9e-14	BC101065	1e-40	Leucine-rich-repeat cont. 15
Fibronectin 1	NM_212476	XM_001601320	3e-33	X54131	3e-145	Prot. tyrosine phosphatase beta
Matrilin 2	NP_002370	AB159149	2e-26	EU085556	1e-34	Collagen XXIX alpha 1
<b>Netrin G1</b>	<b>NP_055732</b>	<b>EF384215</b>	<b>6e-40</b>	<b>HSU75586</b>	<b>0</b>	<b>Human netrin-1</b>
Osteomodulin	NP_005005	XM_001658446	2e-16	AK292182	3e-27	Similar to insulin-like growth factor binding protein, acid labiles subunit (IGFALS)
Podocan	AAP79898	XM_001650288	5e-29	BC101065	1e-40	Leucine-rich-repeat cont. 15
Reelin	AAC51105	XM_320234	9e-04	AB051390	0	<i>VSGPF</i> -spondin
<b>Syndecan 1</b>	<b>EAX00831</b>	<b>XM_001361667</b>	<b>1e-06</b>	<b>NM_001006946</b>	<b>1e-10</b>	<b>Syndecan 1</b>
<b>Syndecan 2</b>	<b>NM_002998</b>	<b>XM_966400</b>	<b>5e-6</b>	<b>AK130131</b>	<b>2e-15</b>	<b>Syndecan 2</b>
<b>Syndecan 3</b>	<b>NP_055469</b>	<b>XM_001664162</b>	<b>5e-3</b>	<b>NM_002998</b>	<b>3e-25</b>	<b>Syndecan 2 (SDC2)</b>
<b>Syndecan 4</b>	<b>NM_002999</b>	<b>XM_001361667</b>	<b>4e-10</b>	<b>NM_001006946</b>	<b>1e-10</b>	<b>Syndecan 1</b>
<b>Tenascin</b>	<b>CAA39628</b>	<b>XM_310633</b>	<b>8e-62</b>	<b>XM_945786</b>	<b>0</b>	<b>Odd Oz/ten-m homolog 2 s</b>
Vitronectin	EAW51082	XM_308849	9e-30	X90925	1e-78	Matrix metalloproteinase
von Willebrand fac.	NM_000552	XM_001606600	2e-168	NM_198445.2	0	SCO-spondin homolog

Entries in bold type indicate human proteins that appear, on the basis of tBLASTn searches, to have orthologs in *Nematostella*.

the C1 domain may have predated the cnidarian-triploblast ancestor, but the ancestral syndecan may be a triploblast invention.

To determine whether the ECM proteins apparently missing from *Nematostella* might be exclusive to deuterostomes or vertebrates, we used tBLASTn to search for them in protostome genomes (Table 7). We then used tBLASTn to compare the top protostome hits back against the human genome to determine whether the best match would correspond to the original human query sequence. These reciprocal tBLASTn searches produced clear evidence for protostome syndecan and tenascin, suggesting that these ECM proteins may have evolved on the triploblastic stem lineage. However, we failed to identify asporin, fibronectin, matrilin, osteomodulin, podocan, reelin, or vitronectin in any protostome animals, suggesting that these ECM proteins may have originated within the Deuterostomia.

### Comparison of protein counts among seven diverse species

We compared the number of predicted proteins bearing Pfam motifs associated with chemical stress, innate immunity, or wound healing in seven diverse taxa (Fig. 3, Suppl. Table 1; <http://www.biolbull.org/supplemental/>). As expected, *E. coli* encodes the fewest proteins in all three categories. For motifs associated with chemical stress response, *Saccharomyces cerevisiae*, *Nematostella*, and *Caenorhabditis elegans* encode fewer proteins than human, while *Drosophila melanogaster* and *Arabidopsis thaliana* encode more. *A. thaliana* outnumbered human for proteins containing motifs associated with chemical stress by more than 3 to 1 due to extremely abundant representation of some Pfam motifs (e.g., p450s), despite lacking representatives for four of the assessed motifs.

For motifs associated with wound repair, the *Nematostella* genome encodes more proteins than any other taxon except human (~50% as many as human), including representatives of every motif queried. However, although *Nematostella* harbors a large number of proteins containing motifs associated with wound healing, it actually lacks homologs for many of the vertebrate proteins specifically implicated in wound healing, including fibronectin, many ECM proteins, and SLRPs. However, all other taxa lack representatives of two or more of the Pfam motifs associated with wound-healing proteins. *E. coli* and *S. cerevisiae* contained the fewest matches for the wound-repair motifs, which is unsurprising given that they are single-celled organisms.

For pathogen response, the overall number of proteins is larger in *A. thaliana* and *C. elegans* than in humans. With respect to pathogen defense, *Nematostella* exhibits fewer hits than other metazoans, but the anemone is the only taxon to have at least one protein representing each

of the Pfam motifs associated with pathogen defense. For example, human contained no matches for the PF05183 (RNA-dependent RNA polymerase), but *Nematostella* had four hits.

## General Discussion and Conclusions

Characterizing the molecular stress-response repertoire of *Nematostella* can be informative at four levels: (1) comparisons to triploblastic animals and non-animal models will help us to reconstruct the early evolution of stress-response pathways; (2) comparisons to species with narrower environmental tolerances, especially closely related marine anemones, will help us to understand the genomic basis for *Nematostella*'s hardiness and ability to survive in estuarine habitats; (3) comparisons among genetically and phenotypically distinct populations of the starlet sea anemone can reveal the microevolution of stress-response pathways; and (4) identifying molecular markers of stress will augment our ability to utilize *Nematostella* and other Cnidaria as environmental sentinels.

### Early evolution of animal stress-response pathways

*Nematostella* is a member of the phylum Cnidaria, a closely related outgroup to the superphylum Triploblastica (=Bilateria; Fig. 1). The Triploblastica encompasses the vast majority of animal phyla and species, and a great deal of research has been conducted in an attempt to explain the remarkable evolutionary radiation of this group (e.g., Peterson *et al.*, 2000). Many human developmental regulatory genes and disease-related genes have clear orthologs in *Nematostella*, indicating that they must have originated prior to the cnidarian-triploblast split (Putnam *et al.*, 2007; Sullivan and Finnerty, 2007). Comparing the genomic complement of stress genes in *Nematostella* and Triploblastica will help illuminate the stress-response system of early animals, perhaps revealing the importance of transcriptional control, a primary difference between stress-response pathways in triploblastic animals and other eukaryotes (e.g., Roelofs *et al.*, 2008).

### Evolution of stress tolerance

*Nematostella* occurs over a wide range of salinities (a range of practical salinity values from 2 to 54 (Sheader *et al.*, 1997) and temperatures (-1°C to 28 °C; Frank and Bleakney, 1978; Williams, 1983; Sheader *et al.*, 1997), and its remarkable environmental tolerances appear to be recently evolved. *Nematostella* is the only known estuarine specialist in the family Edwardsiidae, a clade of coastal marine anemones mainly restricted to temperate and polar seas where neither temperature nor salinity vary much. By contrast, within a single estuary, *Nematostella* may occupy isolated high marsh pools or tidal streams that flush with



each tide—habitats that can differ substantially in temperature, dissolved oxygen, and salinity (e.g. Smith and Able, 1994). The range of *Nematostella* also encompasses dramatic latitudinal variation in temperature, from Nova Scotia, where marsh waters undergo seasonal freezing, to the subtropical waters of the Gulf of Mexico (Hand and Uhlinger, 1994).

The basis for *Nematostella*'s broad environmental tolerances is not known. It could be that *Nematostella*'s genome encodes more stress-response genes than the genomes of related marine species do, and these loci might be expressed at different environmental thresholds. Our analysis suggests this is not the case, because the number of stress-response genes in *Nematostella* is comparable to or even fewer than those in other surveyed metazoans. However, as our search relied on known stress-response genes from other animals, we could have failed to identify “novel” stress-response genes that may be present in the *Nematostella* genome. Identification of such genes will require global analysis of gene expression in *Nematostella* while subjecting the animal to various stressors. Similar techniques for profiling gene expression have been employed in characterizing transcriptional responses to environmental stressors in marine animals, including other cnidarians (Edge *et al.*, 2005; Auslander *et al.*, 2008).

Alternatively, or in concert, *Nematostella* populations could harbor an unusual degree of functional polymorphism in stress-response genes. Such genetic diversity could be maintained by fluctuating selection pressures. We have begun to characterize the geographic distribution and functional significance of some evolutionarily unusual protein-coding polymorphisms that have been identified in stress-response genes (e.g., NF- $\kappa$ B and deiodinase; Sullivan, Reitzel, and Finnerty, unpubl. data).

### Microevolution of stress response

Commensurate with the environmental variation it encounters, *Nematostella* is known to harbor extensive genetic variation (Putnam *et al.*, 2007; Sullivan *et al.*, 2008; Reitzel *et al.*, 2008), evident even at fine spatial scales (Darling *et al.*, 2004; Reitzel *et al.*, 2007). This suggests that the natural dispersal ability of the animals may be quite limited, that local adaptation may be driving genetic differentiation, or a combination of both. At the phenotypic level, we have observed significant differences among clonal stocks in their tolerance to temperature variation and oxidative stress (unpubl. results). With the identification of numerous candidate stress-response genes, we can begin to investigate the genetic basis for *Nematostella*'s remarkable stress tolerance.

### *Nematostella vectensis* as a sentinel species

#### *The need for estuarine sentinel species*

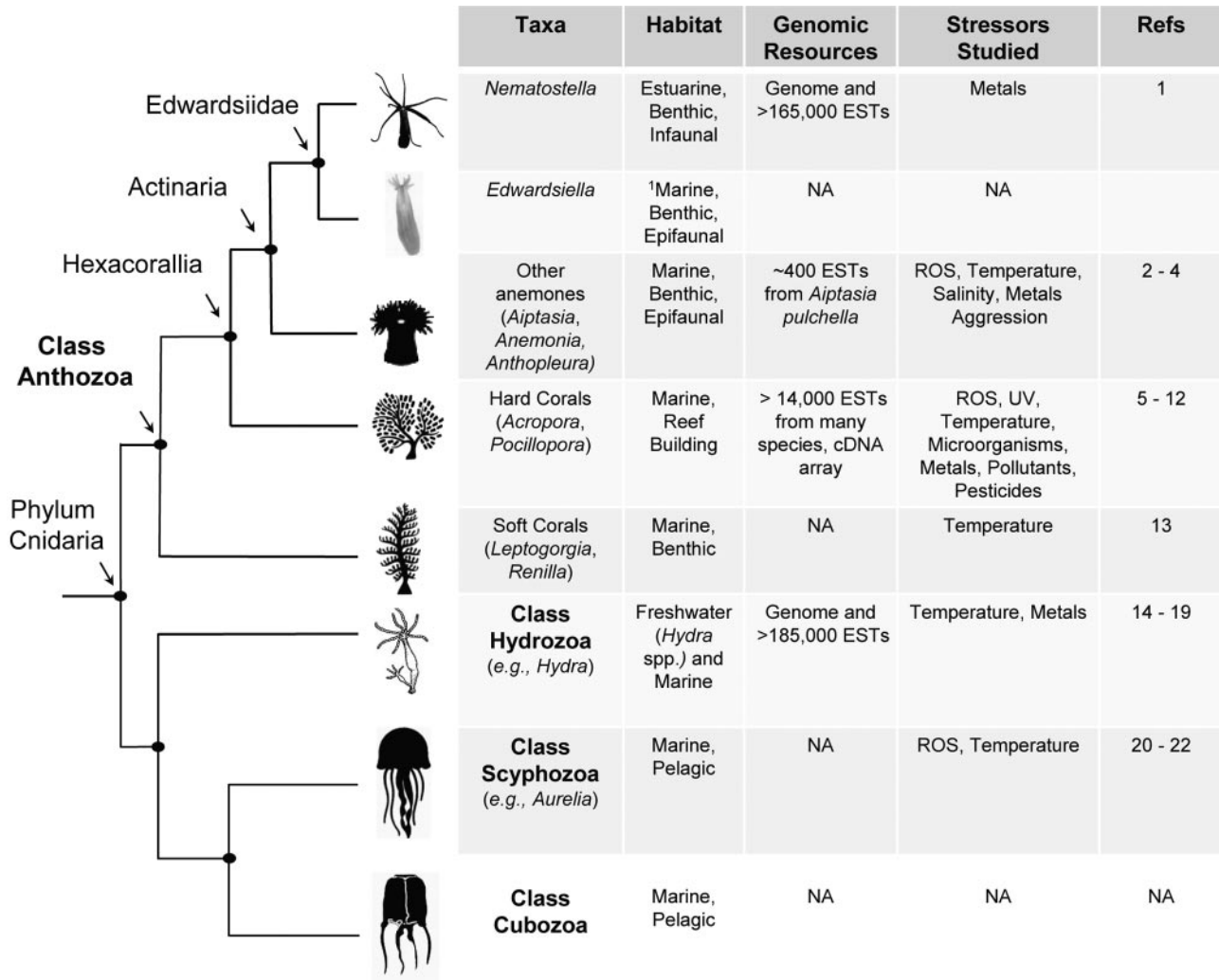
While estuaries are highly variable with respect to key environmental variables, they are also extremely important to marine biodiversity and heavily threatened by increasing human encroachment. Many marine species, including commercially important fishes and crustaceans, utilize estuarine habitats during pre-adult stages, and several endangered shore birds depend upon estuaries for essential forage.

Over the past century, estuarine environments have been severely degraded by increasing human encroachment (Pennings and Bertness, 2001). Estuaries are heavily impacted by anthropogenic contaminants, including heavy metals and endocrine disruptors that are released from point sources of pollution. Estuaries have also become increasingly fragmented by extensive coastal development, some of which interrupts the natural hydrological circulation, rendering the habitat unsuitable for estuarine flora and fauna.

In addition, global climate change represents a particular threat to estuarine natives, which, like other species occupying “island” habitats, may be limited in their ability to disperse to more suitable habitat islands (Stocks and Grassle, 2003; Harley *et al.*, 2006). Global warming will affect not only temperatures in the estuaries, but also the range of other key physiochemical variables (Harley *et al.*, 2006). For example, increasing temperatures can indirectly alter the salinity of salt marshes by impacting the frequency and severity of storms, by melting polar ice caps and increasing sea levels, by altering ocean currents, and by increasing evaporation. The geographic range of estuarine species will respond differently depending upon their physiological tolerances and dispersal ability—some may expand, some may contract, and others will merely shift. This process will alter the membership of estuarine communities, including pathogens, which may increase the incidence of disease in many estuaries. The fate of estuarine communities may depend upon the ability of resident organisms to tolerate and adapt to this combination of physical, chemical, and biotic stressors and decreasing habitat availability. It is therefore urgent that we develop more sophisticated tools to monitor the health of estuarine ecosystems.

#### *Collateral benefits for other cnidarian sentinel species*

Data on *Nematostella* will inform stress-response studies on other cnidarian species occupying different habitats (Fig. 4). The phylum Cnidaria is good source of environmental sentinel species because cnidarians are practically ubiquitous in aquatic habitats, and they are easily cultured *via* asexual propagation. Cnidarians occupy tropical, temperate, and polar latitudes, from the sea surface to the sea floor, in deep and shallow waters, and in freshwater, estuarine, ma-



**Figure 4.** A selection of cnidarian model systems useful for investigating the ecological and evolutionary genomics of stress response. Most of these organisms have a coupled benthic-pelagic life history, so successive life-history stages likely encounter different stressors. Representative stress-response studies are provided for each species: 1. Harter and Matthews (2005); 2. Smith and Watt (1994); 3. Merle *et al.* (2007); 4. Richier *et al.* (2003); 5. Yakovleva *et al.* (2004); 6. Morgan and Snell (2002); 7. Downs *et al.* (2002); 8. Hayes and King (1995); 9. Jones (2005); 10. Edge *et al.* (2005); 11. Varvas *et al.* (1999); 12. Rosenberg *et al.* (2007); 13. Kingsley *et al.* (2003); 14. Bosch *et al.* (1988); 15. Dash *et al.* (2007); 16. Karntanut and Pascoe (2007); 17. Karntanut and Pascoe (2000); 18. Child (1947); 19. Dash *et al.* (2006); 20. (Berking *et al.*, 2005); 21. (Schroth *et al.*, 2005); 22. (Black and Bloom, 1984). <sup>1</sup> *Edwardsiella* is benthic as an adult, but pelagic as a larva and during its parasitic phase in the ctenophore *Mnemiopsis leidyi*.

rine, and hypersaline environments. Their ability to reproduce asexually makes it possible to generate clonal stocks through regeneration—carefully controlled laboratory studies of organismal stress response can then eliminate genetic variation or exploit it, as desired. Repeated trials on the same genotypes (*i.e.*, clonal lineages) exposed to the same environmental conditions can also be used to quantify stochastic variation in gene expression.

A number of cnidarian species are already being used for stress-response studies (Fig. 4), and *Nematostella* comple-

ments these models ecologically and phylogenetically. For example, sea anemones (class Anthozoa, subclass Hexacorallia, order Actinaria) belong to the same subclass as the scleractinians, the main group of reef-building corals. Due to this relatively close evolutionary distance and the apparent similarity of coral and *Nematostella* genes (Miller *et al.*, 2007), the identification of stress-response genes in the starlet sea anemone will inform ongoing genomic stress-response research on corals, which are more limited in experimental tractability (Fig. 4).

### Acknowledgments

AMR was supported by a Postdoctoral Scholar Program at the Woods Hole Oceanographic Institution, with funding provided by The Beacon Institute for Rivers and Estuaries, and the J. Seward Johnson Fund. NTK was supported by a graduate research training grant from the National Institutes of Health. This research was also supported by NSF grant FP-91656101-0 to JCS and JRF, EPA grant F5E11155 to AMR and JRF, and a grant from the Conservation International Marine Management Area Science Program to JRF.

### Literature Cited

- Adler, V., Z. Yin, K. D. Tew, and Z. Ronai. 1999. Role of redox potential and reactive oxygen species in stress signaling. *Oncogene* **18**: 6104–6111.
- Alvarado, A. S., and P. A. Tsonis. 2006. Bridging the regeneration gap: genetic insights from diverse animal models. *Nat. Rev. Genet.* **7**: 873–884.
- Amoutzias, G. D., A. S. Veron, J. Weiner III, M. Robinson-Rechavi, E. Bornberg-Bauer, S. G. Oliver, et al. 2007. One billion years of bZIP transcription factor evolution: conservation and change in dimerization and DNA-binding site specificity. *Mol. Biol. Evol.* **24**: 827–835.
- Andersen, R. L., W. Wiger, H. L. Daae, and K. D. H. Eriksen. 1988. Is the metal binding protein metallothionein present in the coelenterate *Hydra attenuata*? *J. Comp. Physiol.* **91**: 553–557.
- Ankley, G. T., G. P. Daston, S. J. Degitz, N. D. Denslow, R. A. Hoke, S. W. Kennedy, A. L. Miracle, E. J. Perkins, J. Snape, D. E. Tillitt, et al. 2006. Toxicogenomics in regulatory ecotoxicology. *Environ. Sci. Technol.* **40**: 4055–4065.
- Aouacheria, A., C. Geourjon, N. Aghajari, V. Navratil, G. Deleage, C. Lethias, and J.-Y. Exposito. 2006. Insights into early extracellular matrix evolution: spongin short chain collagen-related proteins are homologous to basement membrane type IV collagens and form a novel family widely distributed in invertebrates. *Mol. Biol. Evol.* **23**: 2288–2302.
- Aris-Brosou, S., and Z. Yang. 2002. Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. *Syst. Biol.* **51**: 703–714.
- Armant, M. A., and M. J. Fenton. 2002. Toll-like receptors: a family of pattern-recognition receptors in mammals. *Genome Biol.* **3**: REVIEWS3011.
- Auslander, M., Y. Yudkovski, V. Chalifa-Caspi, B. Herut, R. Ophir, R. Reinhardt, P. M. Neumann, and M. Tom. 2008. Pollution-affected fish hepatic transcriptome and its expression patterns on exposure to cadmium. *Mar. Biotechnol.* **10**: 250–261.
- Barrett, K., and M. McBrien. 2007. Chemical and biological assessment of an urban, estuarine marsh in northeastern New Jersey USA. *Environ. Monit. Assess.* **124**: 63–88.
- Berking, S., N. Czech, M. Gerharz, K. Herrmann, U. Hoffmann, H. Raifer, G. Sekul, B. Siefker, A. Sommerei, and F. Vedder. 2005. A newly discovered oxidant defence system and its involvement in the development of *Aurelia aurita* (Scyphozoa, Cnidaria): reactive oxygen species and elemental iodine control medusa formation. *Int. J. Dev. Biol.* **49**: 969–976.
- Bertrand, S., F. G. Brunet, H. Escriva, G. Parmentier, V. Laudet, and M. Robinson-Rechavi. 2004. Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. *Mol. Biol. Evol.* **21**: 1923–1937.
- Black, R. E., and L. Bloom. 1984. Heat shock proteins in *Aurelia* (Cnidaria, Scyphozoa). *J. Exp. Zool.* **230**: 303–307.
- Bopp, R. F., S. N. Chillrud, E. L. Shuster, H. J. Simpson, and F. D. Estabrooks. 1998. Trends in chlorinated hydrocarbon levels in Hudson River basin sediments. *Environ. Health Perspect.* **106**: 1075–1081.
- Bosch, T. C. G., S. M. Krylow, H. R. Bode, and R. E. Steele. 1988. Thermotolerance and synthesis of heat shock proteins: These responses are present in *Hydra attenuata* but absent in *Hydra oligactis*. *Proc. Natl. Acad. Sci. USA* **85**: 7927–7931.
- Brower, D. L., S. M. Brower, D. C. Hayward, and E. E. Ball. 1997. Molecular evolution of integrins: genes encoding integrin beta subunits from a coral and a sponge. *Proc. Natl. Acad. Sci. USA* **94**: 9182–9187.
- Chakravarti, R., and J. Adams. 2006. Comparative genomics of the syndecans defines an ancestral genomic context associated with matrixins in vertebrates. *BMC Genomics* **7**: 83.
- Child, C. M. 1947. Oxidation and reduction of indicators by hydra. *J. Exp. Zool.* **104**: 153–195.
- Corsi, A., T. Xu, X. D. Chen, A. Boyde, J. Liang, M. Mankani, B. Sommer, R. V. Iozzo, I. Eichstetter, P. G. Robey, et al. 2002. Phenotypic effects of biglycan deficiency are linked to collagen fibril abnormalities, are synergized by decorin deficiency, and mimic Ehlers-Danlos-like changes in bone and other connective tissues. *J. Bone Miner. Res.* **17**: 1180–1189.
- Coteur, G., M. Warnau, M. Jangoux, and P. Dubois. 2002. Reactive oxygen species (ROS) production by amoebocytes of *Asterias rubens* (Echinodermata). *Fish Shellfish Immun.* **12**: 187–200.
- Darling, J. A., A. M. Reitzel, and J. R. Finnerty. 2004. Regional population structure of a widely introduced estuarine invertebrate: *Nematostella vectensis* Stephenson in New England. *Mol. Ecol.* **13**: 2969–2981.
- Dash, B., R. Metz, H. J. Huebner, W. Porter, and T. D. Phillips. 2006. Molecular characterization of phospholipid hydroperoxide glutathione peroxidases from *Hydra vulgaris*. *Gene* **381**: 1–12.
- Dash, B., R. Metz, H. J. Huebner, W. Porter, and T. D. Phillips. 2007. Molecular characterization of two superoxide dismutases from *Hydra vulgaris*. *Gene* **387**: 93–108.
- Dean, M., and T. Annilo. 2005. Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. *Annu. Rev. Genomics Hum. Genet.* **6**: 123–142.
- Derynck, R., and Y. E. Zhang. 2003. Smad-dependent and Smad-independent pathways in TGF- $\beta$  family signalling. *Nature* **425**: 577–584.
- Dishaw, L. J., S. L. Smith, and C. H. Bigger. 2005. Characterization of a C3-like cDNA in a coral: phylogenetic implications. *Immunogenetics* **57**: 535–548.
- Downs, C. A., J. E. Fauth, J. C. Halas, P. Dustan, J. Bemiss, and C. M. Woodley. 2002. Oxidative stress and seasonal coral bleaching. *Free Radic. Biol. Med.* **33**: 533–543.
- Durbin, R., S. Eddy, A. Krogh, and G. Mitchison. 1998. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, Cambridge.
- Eckhart, L., H. Fischer, and E. Tschachler. 2007. Phylogenomics of caspase-activated DNA fragmentation factor. *Biochem. Biophys. Res. Comm.* **356**: 293–299.
- Edge, S. E., M. B. Morgan, D. F. Gleason, and T. W. Snell. 2005. Development of a coral cDNA array to examine gene expression profiles in *Montastraea faveolata* exposed to environmental stress. *Mar. Pollut. Bull.* **51**: 507–523.
- Faler, B. J., R. A. Macsata, D. Plummer, L. Mishra, and A. N. Sidawy. 2006. Transforming growth factor- $\beta$  and wound healing. *Perspect. Vasc. Surg. Endovasc. Ther.* **18**: 55–62.
- Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, et al. 2006. Pfam: clans, web tools and services. *Nucleic Acids Res.* **34**: D247–251.

- Finnerty, J. R., K. Pang, P. Burton, D. Paulson, and M. Q. Martindale. 2004.** Origins of bilateral symmetry: *Hox* and *dpp* expression in a sea anemone. *Science* **304**: 1335–1337.
- Franchini, A., and E. Ottaviani. 2000.** Repair of molluscan tissue injury: role of PDGF and TGF- $\alpha$ . *Tissue Cell* **32**: 312–321.
- Frank, P. G., and J. S. Bleakney. 1978.** Asexual reproduction, diet, and anomalies of the anemone *Nematostella vectensis* in Nova Scotia. *Can. Field-Nat.* **92**: 259–263.
- Galko, M. J., and M. A. Krasnow. 2004.** The molecular biology of wound healing. *PLoS Biol.* **2**: e278.
- Gillitzer, R., and M. Goebeler. 2001.** Chemokines in cutaneous wound healing. *J. Leukoc. Biol.* **69**: 513–521.
- Goldstone, J. V., A. Hamdoun, B. J. Cole, M. Howard-Ashby, D. W. Nebert, M. Scally, M. Dean, E. Epel, M. E. Hahn, and J. J. Stegeman. 2006.** The chemical defenseome: environmental sensing and response genes in the *Strongylocentrotus purpuratus* genome. *Dev. Biol.* **300**: 366–384.
- Grasso, L. C., D. C. Hayward, J. W. H. Trueman, K. M. Hardie, P. A. Janssens, and E. E. Ball. 2001.** The evolution of nuclear receptors: evidence from the coral *Acropora*. *Mol. Phylogenet. Evol.* **21**: 93–102.
- Hamer, D. H. 1986.** Metallothionein. *Annu. Rev. Biochem.* **55**: 913–951.
- Hand, C., and K. Uhlinger. 1994.** The unique, widely distributed sea anemone, *Nematostella vectensis* Stephenson: a review, new facts, and questions. *Estuaries* **17**: 501–508.
- Harley, C. D. G., A. Randall Hughes, K. M. Hultgren, B. G. Miner, C. J. B. Sorte, C. S. Thornber, L. F. Rodriguez, L. Tomanek, and S. L. Williams. 2006.** The impacts of climate change in coastal marine systems. *Ecol. Lett.* **9**: 228–241.
- Harter, V. L., and R. A. Matthews. 2005.** Acute and chronic toxicity test methods for *Nematostella vectensis* Stephenson. *Bull. Environ. Contam. Toxicol.* **74**: 830–836.
- Hayes, R. L., and C. M. King. 1995.** Induction of 70-kD heat shock protein in scleractinian corals by elevated temperature: significance for coral bleaching. *Mol. Mar. Biol. Biotechnol.* **4**: 36–42.
- Henry, L., and M. W. Hart. 2005.** Regeneration from injury and resource allocation in sponges and corals—a review. *Int. Rev. Hydrobiol.* **90**: 125–158.
- Hutton, D. M. C., and V. J. Smith. 1996.** Antibacterial properties of isolated amoebocytes from the sea anemone *Actinia equina*. *Biol. Bull.* **191**: 441–451.
- Huxley-Jones, J., D. L. Robertson, and R. P. Boot-Handford. 2007.** On the origins of the extracellular matrix in vertebrates. *Matrix Biol.* **26**: 2–11.
- Jaiswal, A. K. 2004.** Nrf2 signaling in coordinated activation of anti-oxidant gene expression. *Free Radic. Biol. Med.* **36**: 1199–1207.
- Järving, R., I. Järving, R. Kurg, A. R. Brash, and N. Samel. 2004.** On the evolutionary origin of cyclooxygenase (COX) isozymes: Characterization of marine invertebrate *Cox* genes points to independent duplication events in vertebrate and invertebrate lineages. *J. Biol. Chem.* **279**: 13624–13633.
- Jones, R. 2005.** The ecotoxicological effects of Photosystem II herbicides on corals. *Mar. Pollut. Bull.* **51**: 495–506.
- Kanzok, S. M., N. T. Hoa, M. Bonizzoni, C. Luna, Y. Huang, A. R. Malacrida, and L. Zheng. 2004.** Origin of Toll-like receptor-mediated innate immunity. *J. Mol. Evol.* **58**: 442–448.
- Kao, W. W., J. L. Funderburgh, Y. Xia, C. Y. Liu, and G. W. Conrad. 2006.** Focus on molecules: lumican. *Exp. Eye Res.* **82**: 3–4.
- Karntanut, W., and D. Pascoe. 2000.** A comparison of methods for measuring acute toxicity to *Hydra vulgaris*. *Chemosphere* **41**: 1543–1548.
- Karntanut, W., and D. Pascoe. 2002.** The toxicity of copper, cadmium and zinc to four different *Hydra* (Cnidaria: Hydrozoa). *Chemosphere* **47**: 1059–1064.
- Karntanut, W., and D. Pascoe. 2007.** A comparison of metal accumulation by the cnidarian *Hydra vulgaris* directly from water or through contaminated prey and effects upon reproduction and regeneration. *Songklanakarin J. Sci. Technol.* **29**: 869–880.
- Kingsley, R. J., E. Afif, B. C. Cox, S. Kothari, K. Kriechbaum, K. Kuchinsky, A. T. Neill, A. F. Puri, and V. M. Kish. 2003.** Expression of heat shock and cold shock proteins in the gorgonian *Leptogorgia virgulata*. *J. Exp. Zool. Comp. Exp. Biol.* **296A**: 98–107.
- Kiritsy, C. P., A. B. Lynch, and S. E. Lynch. 1993.** Role of growth factors in cutaneous wound healing: a review. *Crit. Rev. Oral Biol. Med.* **4**: 729–760.
- Krishnamurthy, N., D. Brown, D. Kirshner, and K. Sjölander. 2006.** PhyloFacts: an online structural phylogenomic encyclopedia for protein functional and structural classification. *Genome Biol.* **7**: R83.
- Kudryavtseva, E. I., T. M. Sugihara, N. Wang, R. J. Lasso, J. F. Gudnason, S. M. Lipkin, and B. Andersen. 2003.** Identification and characterization of Grainyhead-like epithelial transactivator (GET-1), a novel mammalian Grainyhead-like factor. *Dev. Dyn.* **226**: 604–617.
- Lee, C. E., and T. Mitchell-Olds. 2006.** Preface to the special issue: ecological and evolutionary genomics of populations in nature. *Mol. Ecol.* **15**: 1193–1196.
- Lettieri, T. 2006.** Recent applications of DNA microarray technology to toxicology and ecotoxicology. *Environ. Health Perspect.* **114**: 4–9.
- Mace, K. A., J. C. Pearson, and W. McGinnis. 2005.** An epidermal barrier wound repair pathway in *Drosophila* is mediated by *grainy head*. *Science* **308**: 381–385.
- Matsushima, N., T. Ohyanagi, T. Tanaka, and R. H. Kretsinger. 2000.** Super-motifs and evolution of tandem leucine-rich repeats within the small proteoglycans—biglycan, decorin, lumican, fibromodulin, PRELP, keratocan, osteoadherin, epiphycan, and osteoglycin. *Proteins Struct. Funct. Genet.* **38**: 210–225.
- Matus, D. Q., K. Pang, H. Marlow, C. W. Dunn, G. H. Thomsen, and M. Q. Martindale. 2006.** Molecular evidence for deep evolutionary roots of bilaterality in animal development. *Proc. Natl. Acad. Sci. USA* **103**: 11195–11200.
- Merle, P.-L., C. Sabourault, S. Richier, D. Allemand, and P. Furla. 2007.** Catalase characterization and implication in bleaching of a symbiotic sea anemone. *Free Radic. Biol. Med.* **42**: 236–246.
- Meszaros, A., and C. Bigger. 1999.** Qualitative and quantitative study of wound healing processes in the coelenterate, *Plexaurella fusifera*: spatial, temporal, and environmental (light attenuation) influences. *J. Invertebr. Pathol.* **73**: 321–331.
- Meyer, J. N., D. C. Volz, J. H. Freedman, and R. T. D. Giulio. 2005.** Differential display of hepatic mRNA from killifish (*Fundulus heteroclitus*) inhabiting a Superfund estuary. *Aquat. Toxicol.* **73**: 327–341.
- Miller, D. J., G. Hemmrich, E. E. Ball, D. C. Hayward, K. Khalturin, N. Funayama, K. Agata, and T. C. G. Bosch. 2007.** The innate immune repertoire in Cnidaria—ancestral complexity and stochastic gene loss. *Genome Biol.* **8**: R59.
- Mitchelmore, C. L., A. H. Ringwood, and V. M. Weis. 2003a.** Differential accumulation of cadmium and changes in glutathione levels as a function of symbiotic state in the sea anemone *Anthopleura elegantissima*. *J. Exp. Mar. Biol. Ecol.* **284**: 71–85.
- Mitchelmore, C. L., E. A. Verde, A. H. Ringwood, and V. M. Weis. 2003b.** Differential accumulation of heavy metals in the sea anemone *Anthopleura elegantissima* as a function of symbiotic state. *Aquat. Toxicol.* **64**: 317–329.
- Morgan, M. B., and T. W. Snell. 2002.** Characterizing stress gene expression in reef-building corals exposed to the mosquitocide dibrom. *Mar. Pollut. Bull.* **44**: 1206–1218.
- Moussian, B., and A. E. Uv. 2005.** An ancient control of epithelial barrier formation and wound healing. *BioEssays* **27**: 987–990.
- Mydlarz, L. D., L. E. Jones, and C. D. Harvell. 2006.** Innate immu-

- nity, environmental drivers, and disease ecology of marine and freshwater invertebrates. *Annu. Rev. Ecol. Evol. Syst.* **37**: 251–288.
- Nakamura, M., K. Mori, S. Inooka, and T. Nomura. 1985.** *In vitro* production of hydrogen peroxide by the amoebocytes of the scallop, *Patinopecten yessoensis* (Jay). *Dev. Comp. Immunol.* **9**: 407–417.
- Nikolopoulos, S. N., P. Blaikie, T. Yoshioka, W. Guo, C. Puri, C. Tacchetti, and F. G. Giancotti. 2005.** Targeted deletion of the integrin  $\beta 4$  signaling domain suppresses laminin-5-dependent nuclear entry of mitogen-activated protein kinases and NF- $\kappa$ B, causing defects in epidermal growth and migration. *Mol. Cell. Biol.* **25**: 6090–6102.
- Nonaka, M., and A. Kimura. 2006.** Genomic view of the evolution of the complement system. *Immunogenetics* **58**: 701–713.
- Olano, C. T., and C. H. Bigger. 2000.** Phagocytic activities of the gorgonian coral *Swiftia exserta*. *J. Invertebr. Pathol.* **76**: 176–184.
- Orosz, A., J. Wisniewski, and C. Wu. 1996.** Regulation of *Drosophila* heat shock factor trimerization: global sequence requirements and independence of nuclear localization. *Mol. Cell. Biol.* **16**: 7018–7030.
- Pennings, S. C., and M. D. Bertness. 2001.** Salt marsh communities. Pp. 289–316 in *Marine Community Ecology*, M. D. Bertness, S. D. Gaines, and M. Hay, eds. Sinauer Associates, Sunderland, MA.
- Peterson, K., R. Cameron, and E. Davidson. 2000.** Bilaterian origins: significance of new experimental observations. *Dev. Biol.* **219**: 1–17.
- Peterson, K. J., and N. J. Butterfield. 2005.** Origin of the Eumetazoa: testing ecological predictions of molecular clocks against the Proterozoic fossil record. *Proc. Natl. Acad. Sci. USA* **102**: 9547–9552.
- Peterson, K. J., J. B. Lyons, K. S. Nowak, C. M. Takacs, M. J. Wargo, and M. A. McPeck. 2004.** Estimating metazoan divergence times with a molecular clock. *Proc. Natl. Acad. Sci. USA* **101**: 6536–6541.
- Putnam, N. H., M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, A. Salamov, A. Terry, H. Shapiro, E. Lindquist, V. V. Kapitonov, et al. 2007.** Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* **317**: 86–94.
- Rämte, M., R. Lanot, D. Zachary, and P. Manfrulli. 2002.** JNK signaling pathway is required for efficient wound healing in *Drosophila*. *Dev. Biol.* **241**: 145–156.
- Reber-Muller, S., R. Studer, P. Muller, N. Yanze, and V. Schmid. 2001.** Integrin and talin in the jellyfish *Podocoryne carnea*. *Cell Biol. Int.* **25**: 753–769.
- Reitzel, A. M., P. M. Burton, C. Krone, and J. R. Finnerty. 2007.** Comparison of developmental trajectories in the starlet sea anemone *Nematostella vectensis*: embryogenesis, regeneration, and two forms of asexual fission. *Invertebr. Biology* **126**: 99–112.
- Reitzel, A. M., J. D. Darling, J. C. Sullivan, and J. R. Finnerty. 2008.** Global population genetic structure of the starlet anemone *Nematostella vectensis*: multiple introductions and implications for conservation policy. *Biol. Invasions*. (In press).
- Richier, S., P. L. Merle, P. Furla, D. Pigozzi, F. Sola, and D. Allemand. 2003.** Characterization of superoxide dismutases in anoxia- and hyperoxia-tolerant symbiotic cnidarians. *Biochim. Biophys. Acta* **1621**: 84–91.
- Richier, S., P. Furla, A. Plantivaux, P.-L. Merle, and D. Allemand. 2005.** Symbiosis-induced adaptation to oxidative stress. *J. Exp. Biol.* **208**: 277–285.
- Roelofs, D., M. G. M. Aarts, H. Schat, and N. M. Van Straalen. 2008.** Functional ecological genomics to demonstrate general and specific responses to abiotic stress. *Funct. Ecol.* **22**: 8–18.
- Rosenberg, E., O. Koren, L. Reshef, R. Efrony, and I. Zilber-Rosenberg. 2007.** The role of microorganisms in coral health, disease and evolution. *Nat. Rev. Microbiol.* **5**: 355–362.
- Rossi, S., and M. J. Snyder. 2001.** Competition for space among sessile marine invertebrates: changes in HSP70 expression in two Pacific cnidarians. *Biol. Bull.* **201**: 385–393.
- Saydam, N., O. Georgiev, M. Y. Nakano, U. F. Greber, and W. Schaffner. 2001.** Nucleo-cytoplasmic trafficking of metal-regulatory transcription factor 1 is regulated by diverse stress signals. *J. Biol. Chem.* **276**: 25487–25495.
- Schröder, H. C., K. Shostak, V. Gamulin, M. Lacorn, A. Skorokhod, V. Kavsan, and W. E. G. Müller. 2000.** Purification, cDNA cloning and expression of a cadmium-inducible cysteine-rich metallothionein-like protein from the marine sponge *Suberites domuncula*. *Mar. Ecol. Prog. Ser.* **200**: 149–157.
- Schroth, W., A. Ender, and B. Schierwater. 2005.** Molecular biomarkers and adaptation to environmental stress in moon jelly (*Aurelia* spp.). *Mar. Biotechnol. (NY)* **7**: 449–461.
- Schulte, P. M. 2007.** Responses to environmental stressors in an estuarine fish: interacting stressors and the impacts of local adaptation. *J. Therm. Biol.* **32**: 152–161.
- Shearer, M., A. M. Suwailem, and G. A. Rowe. 1997.** The anemone, *Nematostella vectensis*, in Britain: considerations for conservation management. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **7**: 13–25.
- Simionato, E., V. Ledent, G. Richards, M. Thomas-Chollier, P. Kerner, D. Coornaert, B. M. Degnan, and M. Vervoort. 2007.** Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics. *BMC Evol. Biol.* **7**: 33.
- Singer, A. J., and R. A. Clark. 1999.** Cutaneous wound healing. *N. Engl. J. Med.* **341**: 738–746.
- Smith, B. L., and W. B. Watt. 1994.** Adenylate levels and environmental stress in the sea anemone *Anthopleura elegantissima*. *Mol. Mar. Biol. Biotechnol.* **3**: 261–269.
- Smith, K. J., and K. W. Able. 1994.** Salt-marsh tide pools as winter refuges for the mummichog, *Fundulus heteroclitus*, in New Jersey. *Estuaries* **17**: 226–234.
- Sodergren, E., G. M. Weinstock, E. H. Davidson, R. A. Cameron, R. A. Gibbs, R. C. Angerer, L. M. Angerer, M. I. Arnone, D. R. Burgess, R. D. Burke, et al. 2006.** The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* **314**: 941–952.
- Steffensen, B., L. Hakkinen, and H. Larjava. 2001.** Proteolytic events of wound-healing—coordinated interactions among matrix metalloproteinases (MMPs), integrins, and extracellular matrix molecules. *Crit. Rev. Oral Biol. Med.* **12**: 373–398.
- Stocks, K. I., and J. F. Grassle. 2003.** Benthic macrofaunal communities in partially impounded salt marshes in Delaware: comparisons with natural marshes and responses to sediment exposure. *Estuaries* **26**: 777–789.
- Sullivan, J. C., and J. R. Finnerty. 2007.** A surprising abundance of human disease genes in a simple, basal animal, the starlet sea anemone (*Nematostella vectensis*). *Genome* **50**: 689–692.
- Sullivan, J. C., A. M. Reitzel, and J. R. Finnerty. 2006.** A high percentage of introns in human genes were present early in animal evolution: evidence from the basal metazoan *Nematostella vectensis*. *Genome Inform.* **17**: 219–229.
- Sullivan, J. C., D. Kalaitzidis, T. D. Gilmore, and J. R. Finnerty. 2007.** Rel homology domain-containing transcription factors in the cnidarian *Nematostella vectensis*. *Dev. Genes Evol.* **217**: 63–72.
- Sullivan, J. C., A. M. Reitzel, and J. R. Finnerty. 2008.** Upgrades to *StellaBase* facilitate medical and genetic studies on the starlet sea anemone, *Nematostella vectensis*. *Nucleic Acids Res.* **36**: D607–D611.
- Takada, Y., X. Ye, and S. Simon. 2007.** The integrins. *Geome Biol.* **8**: 215.
- Tarrant, A. 2007.** Hormonal signaling in cnidarians: Do we understand the pathways well enough to know whether they are being disrupted? *Ecotoxicology* **16**: 5–13.
- Tarrant, A. M. 2005.** Endocrine-like signaling in cnidarians: current understanding and implications for ecophysiology. *Integr. Comp. Biol.* **45**: 201–214.
- Technau, U., S. Rudd, P. Maxwell, P. M. K. Gordon, M. Saina, L. C. Grasso, D. C. Hayward, C. W. Sensen, R. Saint, et al. 2005.**

- Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. *Trends Genet.* **21**: 633–639.
- Tettamanti, G., A. Grimaldi, T. Congiu, G. Perletti, M. Raspanti, R. Valvassori, and M. de Eguileor. 2005.** Collagen reorganization in leech wound healing. *Biol. Cell* **97**: 557–568.
- Thornton, J. W. 2003.** Nonmammalian nuclear receptors: evolution and endocrine disruption. *Pure Appl. Chem.* **75**: 1827–1839.
- Ting, S. B., J. Caddy, N. Hislop, T. Wilanowski, A. Auden, L.-L. Zhao, S. Ellis, P. Kaur, Y. Uchida, W. M. Holleran, et al. 2005.** A homolog of *Drosophila* grainy head is essential for epidermal integrity in mice. *Science* **308**: 411–413.
- Uv, A. E., C. R. Thompson, and S. J. Bray. 1994.** The *Drosophila* tissue-specific factor Grainyhead contains novel DNA-binding and dimerization domains which are conserved in the human protein CP2. *Mol. Cell. Biol.* **14**: 4020–4031.
- Varvas, K., I. Järving, R. Koljak, K. Valmsen, A. R. Brash, and N. Samel. 1999.** Evidence of a cyclooxygenase-related prostaglandin synthesis in coral. The allene oxide pathway is not involved in prostaglandin biosynthesis. *J. Biol. Chem.* **274**: 9923–9929.
- Vasiliou, V., D. Ross, and D. W. Nebert. 2006.** Update of the NAD(P)H:quinone oxidoreductase (NQO) gene family. *Hum. Genomics* **2**: 329–335.
- Weis, J. S., J. Skurnick, and P. Weis. 2004.** Studies of a contaminated brackish marsh in the Hackensack Meadowlands of Northeastern New Jersey: benthic communities and metal contamination. *Mar. Pollut. Bull.* **49**: 1025–1035.
- Weis, P., K. R. Barrett, T. Proctor, and R. F. Bopp. 2005.** Studies of a contaminated brackish marsh in the Hackensack Meadowlands of northeastern New Jersey: an assessment of natural recovery. *Mar. Pollut. Bull.* **50**: 1405–1415.
- Williams, R. B. 1983.** Starlet sea anemone: *Nematostella vectensis* Stephenson 1935. Pp. 43–46 in *The IUCN Invertebrate Red Data Book*. International Union for Conservation of Nature and Natural Resources, Gland, Switzerland.
- Xu, L. 2006.** Regulation of Smad activities. *Biochim. Biophys. Acta* **1759**: 503–513.
- Yakovleva, I., R. Bhagooli, A. Takemura, and M. Hidaka. 2004.** Differential susceptibility to oxidative stress of two scleractinian corals: antioxidant functioning of mycosporine-glycine. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **139**: 721–730.
- Yoshihara, Y., M. Kawasaki, A. Tamada, S. Nagata, H. Kagamiyama, and K. Mori. 1995.** Overlapping and differential expression of BIG-2, BIG-1, TAG-1, and F3: four members of an axon-associated cell adhesion molecule subgroup of the immunoglobulin superfamily. *J. Neurobiol.* **28**: 51–69.
- Zhang, B., D. Egli, O. Georgiev, and W. Schaffner. 2001.** The *Drosophila* homolog of mammalian zinc finger factor MTF-1 activates transcription in response to heavy metals. *Mol. Cell. Biol.* **21**: 4505–4514.
- Zheng, L., L. Zhang, H. Lin, M. T. McIntosh, and A. R. Malacrida. 2005.** Toll-like receptors in invertebrate innate immunity. *Invertebr. Survival J.* **2**: 105–113.