Early evolution of animal cell signaling and adhesion genes

Scott A. Nichols*, William Dirks†, John S. Pearse‡, and Nicole King*†§¶

Departments of *Molecular and Cell Biology and §Integrative Biology and †Center for Integrative Genomics, University of California, Berkeley, CA 94720; and ‡Joseph M. Long Marine Laboratory, University of California, Santa Cruz, CA 95060

Edited by James W. Valentine, University of California, Berkeley, CA, and approved June 28, 2006 (received for review May 17, 2006)

In stark contrast to the rapid morphological radiation of eumetazoans during the Cambrian explosion, the simple body plan of sponges (Phylum Porifera) emerged from the Cambrian relatively unchanged. Although the genetic and developmental underpinnings of these disparate evolutionary outcomes are unknown, comparisons between modern sponges and eumetazoans promise to reveal the extent to which critical genetic factors were present in their common ancestors. Two particularly interesting classes of genes in this respect are those involved in cell signaling and adhesion. These genes help guide development and morphogenesis in modern eumetazoans, but the timing and sequence of their origins is unknown. Here, we demonstrate that the sponge *Oscarella carmela*, one of the earliest branching animals, expresses core components of the Wnt, transforming growth factor β, receptor tyrosine kinase, Notch, Hedgehog, and Jak/Stat signaling pathways. Furthermore, we identify sponge homologs of nearly every major eumetazoan cell-adhesion gene family, including those that encode cell-surface receptors, cytoplasmic linkers, and extracellular-matrix proteins. From these data, we infer that key signaling and adhesion genes were in place early in animal evolution, before the divergence of sponge and eumetazoan lineages.

eumetazoa | porifera | Cambrian explosion | homoscleromorpha | *Oscarella*

The fossil record from ≈540 Mya documents the abrupt appearance of fully diversified eumetazoan body plans during the "Cambrian explosion." In contrast, fossil evidence of sponges dates to ≈580 Mya and reveals that their simple body plan predates the Cambrian and has since remained relatively unchanged (1, 2). Although extrinsic environmental factors have received much attention as the driving force behind the Cambrian explosion (3), the disparate outcomes in lineages leading to modern sponges and eumetazoans point to intrinsic genetic differences. Intrinsic factors that potentially contributed to the selective morphological diversification of eumetazoans include the evolution of developmentally important gene families, signaling pathways, or regulatory elements and linkages in their ancestors. A key to illuminating how intrinsic factors contributed to the Cambrian explosion will be to reconstruct the evolutionary sequence of origination, elaboration, and assembly of developmentally important gene families into interacting pathways and networks.

Modern sponges promise to offer critical insights into the earliest events in animal genome evolution. The fossil record of sponges predates that of all other animal groups (2), and phylogenetic analyses support sponges as an outgroup of all eumetazoan phyla (4). Indeed, recent studies indicate that sponges are paraphyletic (5–9) and that eumetazoans evolved from a sponge-like ancestor (9). Among the major sponge groups, the homoscleromorphs (Fig. 1*b*) have particular relevance to the study of eumetazoan origins and evolution. Although of uncertain phylogenetic position (10, 11), homoscleromorphs are thought to be unique among sponges in their possession of eumetazoan-like ultrastructural features, including an epithelium characterized by closely apposed cells with an underlying basement membrane (Fig. 1 *c* and *d*) and regularly distributed cell–cell junctions (12, 13). The phylogenetic position of sponges and the cell biology of homoscleromorph sponges suggest that homoscleromorphs may provide an unprecedented window into the ancestry of genes important for eumetazoan development and evolution.

Two crucial classes of genes in this respect are the signaling and adhesion genes that interact to coordinate complex morphogenetic events in eumetazoans as diverse as flies, worms, and humans (14–16). Cell-signaling pathways allow cells to receive, relay, and interpret messages from the extracellular environment and are involved in developmental cell-fate specification and patterning (15). Classically, seven major signaling pathways have been identified as unique to, and ubiquitous among, eumetazoans (15): Wnt, transforming growth factor β (TGFβ), Hedgehog, receptor tyrosine kinase (RTK), Jak/STAT, Notch, and nuclear hormone receptor. Core elements from each of these pathways are present in eumetazoans ranging from cnidarians to vertebrates (17), where they have both retained their presumed ancestral roles, such as in axial patterning, and been coopted to guide the development of derived features, such as wings and eyes. With the exception of several RTKs (18–23) and a *Frizzled* receptor (part of the Wnt pathway) (24), it is not known whether the core components of any of the remaining pathways exist in sponges. Because cell signaling sits at the top of key developmental regulatory networks (15, 25), one prediction is that morphological novelties during the diversification of eumetazoan body plans stemmed, in part, from evolutionary changes in signaling pathways.

A second class of genes relevant to the search for intrinsic factors underlying the diversification of eumetazoan body plans are the cell-adhesion genes, which are often downstream targets of signaling cascades. By helping to regulate the sorting and migration of different cell populations, cell-adhesion genes act to effect pivotal morphogenetic changes during eumetazoan development (14, 16). Cell-adhesion proteins also play critical roles in eumetazoan epithelia to establish and maintain the transport and occluding functions hypothesized to be prerequisites of animal body-plan diversification (26). The functional properties of epithelia laid the foundation for the evolution of differentiated body compartments with unique physiological properties (26) and, therefore, were likely to have served as an important template for morphological evolution during the Cambrian radiation. Despite intensive study of cell adhesion in sponges, few sponge homologs of eumetazoan cell-adhesion gene families have been identified. Instead, models of cell adhesion in sponges largely emphasize the interactions between carbohydrate moi-

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: GO, Gene Ontology; NCBI, National Center for Biotechnology Information; nr, nonredundant.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. EB741367–EB741549 and EC366551–EC377543).

¶To whom correspondence should be addressed. E-mail: nking@berkeley.edu.

EVOLUTION

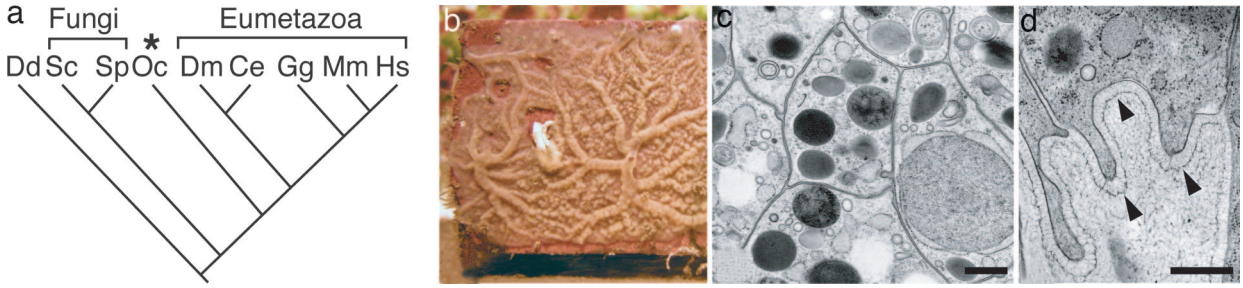

Fig. 1. An introduction to the sponge, *O. carmela*. (*a*) Phylogenetic position of *O. carmela* (asterisk) relative to model bilaterians. *O. carmela* belongs to one of the earliest branching phyla of animals and may provide insight into the biology and genome of the first multicellular animals. (*b*) Macroscopic view of *O. carmela* overgrowing a standard brick (shown for scale). As an adult, *O. carmela* has no body axes and no differentiated structures other than epithelia. (*c* and *d*) Electron micrographs of the larval epithelium of *O. carmela* demonstrate that it has closely apposed cells (*c*) and a basement membrane (*d*, arrowheads), a feature thought to be exclusive to eumetazoans and the sponge clade containing *Oscarella*, the homoscleromorph sponges. Dd, *D. discoideum*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Oc, *O. carmela*; Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*; Gg, *Gallus gallus*; Mm, *Mus musculus*; Hs, *Homo sapiens*. (Scale bars, 500 nm.)

eties of large extracellular proteoglycans (termed aggregation factors) that are hypothesized to bind to each other and to cell surface receptors that may interact with integrins (27, 28).

To identify cell-signaling and adhesion genes that evolved before the divergence of sponges and eumetazoans, we collected and annotated 11,520 ESTs from the homoscleromorph sponge *Oscarella carmela* (phylum Porifera; Fig. 1). Despite the morphological simplicity of sponges, we find that six of the seven major eumetazoan signaling pathways have homologs in *O. carmela*. Furthermore, *O. carmela* expresses an unexpected diversity of eumetazoan cell-adhesion gene homologs, representing most of the major eumetazoan cell-adhesion gene families. Our findings indicate that developmentally important signaling and adhesion gene families evolved before the divergence of sponge and eumetazoan lineages and were, therefore, in place in all animal lineages at the onset of the Cambrian explosion.

Results and Discussion

Diversity of Cell-Signaling Gene Homologs in *O. carmela*. To characterize the diversity of cell-signaling gene homologs available before the diversification of eumetazoan body plans, we collected 11,520 ESTs from embryonic, larval, and adult tissue of the sponge *O. carmela*. The ESTs were generated by 5′- and 3′-end sequencing of randomly selected clones from a directionally cloned cDNA library and were evaluated through two sequential rounds of annotation. By using contiguous sequences (contigs) built from overlapping ESTs, a phylogeny-based annotation system called Metazome, coupled with comparisons to the National Center for Biotechnology Information (NCBI) nonredundant (nr) and the Gene Ontology (GO) databases, was used to identify sponge sequences with high similarity to components of eumetazoan signaling pathways. These sequences were then manually inspected to confirm proper contig assembly and then were reannotated. Predicted sequence homology was based on annotations from the three independent databases.

We find that *O. carmela* expresses an unexpected diversity of signaling genes, including homologs of core genes in six of the seven key bilaterian signaling pathways (15): Wnt, TGF-β, Hedgehog, receptor tyrosine kinase, Jak/STAT, and Notch signaling pathways (Table 1). Although we do not find evidence of nuclear hormone receptors (NHRs) (the one remaining eumetazoan signaling pathway), we do not conclude from our EST survey that NHRs are absent from sponge genomes. To explore the evolutionary patterns of gain and loss of the signaling genes discovered in *O. carmela*, we charted their distribution in diverse animals, fungi, and the amoebozoan *Dictyostelium discoideum*. Because our analysis is restricted to organisms from which complete genome sequences are available, both presence and absence data shown in Table 1 are meaningful. Notably, pathway components that have been previously characterized only in eumetazoans [e.g., *Wnt*, *Dickkopf*, *Patched*, *Notch*, *Delta*, *Mothers against decapentaplegic* (*Mad*), and *Janus kinase* (*Jak*)] are now known from sponges, thus extending the history of these important genes back to some of the earliest animal ancestors. An unexpected outcome from comparative genetic studies of early branching animals is that they sometimes have homologs of genes previously thought to be derived within bilaterian clades. This was true of the discovery of *Dickkopf* in Cnidarians (29) (and now in sponges) and is true of the Wnt pathway component *Dixdc*, now known from deuterostomes and *O. carmela*.

The diversity of signaling-pathway elements present in *O. carmela* reveals that major animal signal-transduction mechanisms evolved before they were adopted for their sophisticated eumetazoan functions. Examples of how signaling pathways are deployed during the development of bilaterians include the roles of Wnt and Hedgehog pathways in neural circuit assembly (30) and limb development (31) and the role of Notch signaling in *Drosophila* eye development and vertebrate segmentation (32). Given that sponges lack neurons, limbs, eyes, and segments, the question emerges: Why are so many components of these and other signaling pathways conserved in *O. carmela*? One possible explanation is that early animals evolved intricate developmental programs long before they evolved diverse developmental outcomes. This hypothesis stems from the observation that sponge embryos undergo coordinated and predictable cell rearrangements that resemble gastrulation and result in larvae with clear anterior/posterior axes (33, 34). Furthermore, this hypothesis is consistent with the prediction, derived from studies of *Wnt* expression in the cnidarian *Nematostella vectensis*, that *Wnt* functioned in axial patterning and gastrulation in the last common ancestors of eumetazoa (35). Indeed, the recent discovery that homologs of eumetazoan transcription factors are developmentally deployed in the sponge *Reniera* sp. provides independent evidence that the last common metazoan ancestor exhibited eumetazoan-like developmental patterning (36).

Diversity of Cell-Adhesion Gene Homologs in *O. carmela*. From our EST survey, we find that *O. carmela* has a nearly full complement (37) of gene families that, in eumetazoans, function as cell-adhesion receptors, extracellular matrix (ECM)/basement membrane components, and cytoskeletal-linker proteins (Table 2). Notably, *O. carmela* has a homolog of *Fras1*, an ECM component otherwise known only from deuterostomes. The observation that sponges express diverse eumetazoan cell-adhesion genes suggests that current models of sponge cell adhesion, which em-

Table 1. Animal signaling-pathway components in the sponge *O. carmela*

		Phylogenetic distribution of signaling gene families*								
		Bilateria						Fungi		Amb
Pathway components†	Functional roles	Hs	Mm	Gg	Sp	Dm	Ce	Sp	Sc	Dd
Wnt										
Wnt	Ligand	●	●	●	●	●	●	○	○	○
Frizzled	Receptor	●	●	●	●	●	●	○	○	●
Dkk	Antagonist	●	●	●	○	○	○	○	○	○
Dixdc	Positive regulator	●	●	●	●	○	○	○	○	○
β-catenin	Downstream effector	●	●	●	●	●	●	○	●	●
Dvl	Downstream effector	●	●	●	●	●	●	○	○	○
Nemo	Downstream effector	●	●	●	●	●	●	○	○	○
TGFβ										
Follistatin	Antagonist	●	●	●	●	●	○	○	○	○
Mad	Downstream effector	●	●	●	●	●	●	○	○	○
Hedgehog										
Hh	Ligand	●	●	●	●	●	●	○	○	○
Ptc	Receptor	●	●	●	●	●	●	○	○	○
Disp	Ligand export	●	●	●	●	●	●	○	○	○
Sufu	Downstream effector	●	●	●	●	●	○	○	○	○
Receptor tyrosine kinase										
Egfr	Receptor	●	●	●	●	●	●	○	○	○
Igfr	Receptor	●	●	●	●	●	●	○	○	○
Fgfr	Receptor	●	●	●	●	●	●	○	○	○
Epha	Receptor	●	●	●	●	●	●	○	○	○
Ret	Receptor	●	●	●	○	●	○	○	○	○
Musk	Receptor	●	●	●	○	●	●	○	○	○
Ddr	Receptor	●	●	●	●	●	●	○	○	○
Jak/Stat										
Jak	Receptor cofactor	●	●	●	●	●	○	○	○	○
Stat	Transcription factor	●	●	●	●	●	●	○	○	●
Stam	Downstream effector	●	●	●	●	●	●	●	●	○
Pias	Antagonist	●	●	●	●	●	●	○	●	○
Notch/Delta										
Delta	Ligand	●	●	●	●	●	●	○	○	○
Notch	Receptor	●	●	●	●	●	●	○	○	○

Dixdc, Dix domain containing; Dkk, Dickkopf; Dvl, Dishevelled; Nemo, Nemo-like kinase; Mad, mothers against decapentaplegic homolog; *Hh, hedgehog* homolog; *Ptc, Patched; Sufu, suppressor of fused; Disp, Dispatched; Egfr, epidermal growth factor receptor; Igfr, insulin-like growth factor receptor; Fgfr, fibroblast growth factor receptor; Epha, Eph receptor A; Ret, Ret protooncogene; Musk, skeletal muscle tyrosine kinase receptor; Ddr, epithelial discoidin domain receptor; Jak, Janus kinase; Stam, signal transducing adaptor molecule; Stat, signal transducer and activator of transcription; Pias, protein inhibitor of activated STAT;* Amb, Amoebozoa; Hs, *Homo sapiens*; Cf, *Canis familiaris*; Mm, *Mus musculus*; Gg, *Gallus gallus*; Sp, *Strongylocentrotus purpuratus*; Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*; Sp, *Schizosaccharomyces pombe*; Sc, *Saccharomyces cerevisiae*; Dd, *Dictyostelium discoideum*.

*Filled circles (●) indicate the presence of gene homologs, not necessarily orthologs, in select taxa. Open circles (○) indicate their absence.

†See supporting information for a full gene list and supplemental data.

phasize the role of proteoglycan aggregation factors in species-specific reaggregation of dissociated cells (e.g., refs. 38 and 39), are incomplete. Indeed, sponges entered the historic period of morphological radiation equipped with the adhesion machinery used by eumetazoans to selectively sort differentiated cell populations and to form epithelial barriers between differentiating body compartments.

The observation that ancestral sponges and eumetazoans had an approximately equivalent arsenal of cell-adhesion genes necessary for morphogenetic cell sorting and epithelium formation presents an enigma: Why did eumetazoans diversify while the sponge form remained static? One possible explanation for the disparate patterns of morphological evolution between sponges and eumetazoans is that preexisting gene families developed new interactions that conferred new functions to the ancestors of eumetazoans. The phylogenetic patterns of interactions between cell-adhesion receptors and their cytoskeletal linker proteins offer an example of how the evolutionary sequence of molecular interactions can be reconstructed. Integrins are linked to the actin cytoskeleton by their interactions with talin, vinculin, α-actinin, filamin, and paxillin, all of which are present in *O. carmela*. In contrast, classical cadherins are linked to the actin cytoskeleton through interactions with β-catenin, which also features prominently in Wnt signaling (37). Although

EVOLUTION

Table 2. Eumetazoan cell-adhesion machinery in the sponge *O. carmela*

	Phylogenetic distribution of cell adhesion gene families*								
	Bilateria						Fungi		Amb
Adhesion genes homologs†	Hs	Mm	Gg	Sp	Dm	Ce	Sp	Sc	Dd
Cell-contact and -adhesion proteins									
Contactin	●	●	●	●	●	●	○	○	○
β-integrin	●	●	●	●	●	●	○	○	○
α-integrin	●	●	●	●	●	●	○	○	○
ADAM	●	●	●	●	●	●	○	○	○
Protocadherin	●	●	●	●	●	●	○	○	○
Fat	●	●	●	●	●	●	○	○	○
NCAM	●	●	●	●	●	●	○	○	○
Selectin	●	●	●	●	●	●	○	○	○
Tetraspanin	●	●	●	●	●	●	○	○	○
Plexin	●	●	●	●	●	●	○	○	○
Heparanase	●	●	●	●	●	○	○	○	○
Neurexin	●	●	●	●	●	●	○	○	○
Crumbs	●	●	●	●	●	●	○	○	○
ECM molecules									
Agrin	●	●	●	●	●	●	○	○	○
Cthrc	●	●	●	●	●	●	○	○	○
Col11a2	●	●	●	●	●	●	○	○	○
Col4	●	●	●	●	●	●	○	○	○
Fibulin	●	●	●	●	●	●	○	○	○
Fras1	●	●	●	●	○	○	○	○	○
α-laminin	●	●	●	●	●	●	○	○	○
β-laminin	●	●	●	●	●	●	○	○	○
Netrin	●	●	●	●	●	●	○	○	○
Perlecan	●	●	●	●	●	●	○	○	○
Tenascin	●	●	●	●	●	●	○	○	○
Thrombospondin	●	●	●	●	●	●	○	○	○
Spondin	●	●	●	●	●	●	○	○	○
Fibrillin	●	●	●	●	●	●	○	○	○
40S ribosomal protein SA	●	●	●	●	●	●	●	●	●
Nardilysin	●	●	●	●	●	○	○	○	○
Adhesion-related cytoskeletal linkers									
Ankyrin	●	●	●	●	●	●	○	○	○
α-actinin	●	●	●	●	●	●	●	●	●
Paxillin	●	●	●	●	●	●	○	●	●
β-catenin	●	●	●	●	●	●	○	●	●
p130^{cas}	●	●	●	○	●	●	○	○	○
Rho1 GTPase	●	●	●	●	●	●	○	○	○
Spectrin	●	●	●	●	●	●	○	○	○
Talin	●	●	●	●	●	●	○	○	●
Vinculin	●	●	●	●	●	●	○	○	○
Fascin	●	●	●	●	●	●	○	○	○
Filamin	●	●	●	●	●	●	○	○	●
Parvin	●	●	●	●	●	●	○	○	○
LASP-1	●	●	●	○	●	●	○	○	○

LASP-1, *LIM and SH3 domain protein 1*; *ADAM*, *a disintegrin and metalloprotease*; *Fat*, *Fat protocadherin tumor suppressor*; *NCAM*, *neural cell adhesion molecule*; ECM, extracellular matrix; *Cthrc*, *collagen triple helix repeat containing*; *Col11a2*, *procollagen, type XI, α 2*; *Col4*, *Type IV collagen*; *Fras1*, *Fraser Syndrome 1*; Amb, Amoebozoa; Hs, *H. sapiens*; Cf, *C. familiaris*; Mm, *M. musculus*; Gg, *G. gallus*; Sp, *S. purpuratus*; Dm, *D. melanogaster*; Ce, *C. elegans*; Sp, *Sc. pombe*; Sc, *Sa. cerevisiae*; Dd, *D. discoideum*.

*Filled circles (●) indicate the presence of gene homologs, not necessarily orthologs, in select taxa. Open circles (○) indicate their absence.

†See supporting information for a full gene list and supplemental data.

we find evidence of *β-catenin* and other Wnt pathway components in *O. carmela*, classical cadherins have not been discovered outside of Bilateria (40). However, because β-catenin localizes to adhesive structures in the amoebozoan *D. discoideum* (41), it is likely to have had ancestral adhesive roles that were decoupled from both cadherin-mediated adhesion and Wnt signaling. From the available evidence, we hypothesize that β-catenin had an ancestral role in adhesion, but its role in Wnt signaling may predate its role in cadherin-mediated adhesion.

From the combined efforts of paleontology, systematics, genomics, and "evo-devo," a model of early animal evolution is emerging. Current evidence suggests that most major animal lineages (including bilaterians) (7, 42–44) extend into the pre-Cambrian where they arose from a sponge-like ancestor (9). In response to environmental and ecological changes during the Cambrian, the ancestors of modern eumetazoans underwent rampant morphological radiations, whereas independent ancestral lineages of sponges did not. We now know that the ancestors of both homoscleromorph sponges

and eumetazoans had roughly comparable sets of developmentally important genes. Specifically, *O. carmela* expresses many homologs of genes known from the major eumetazoan signaling pathways and cell-adhesion and transcription factor (36) gene families. What remains unclear is how these genes function in sponges, how they functioned in eumetazoan ancestors, and whether they were coopted to new functions or linkages in the lineage leading to modern eumetazoans. Further studies of expressed genes and protein function in *O. carmela* may provide insights into the intrinsic factors underlying key events in the early evolution of animal phyla.

Materials and Methods

RNA Isolation and mRNA Extraction. An individual specimen of *O. carmela* bearing densely brooded embryos and larvae of all stages was removed from the glass surface of an open-seawater aquarium at the Joseph M. Long Marine Laboratory, University of California (Santa Cruz). Tissue was transported live in seawater to the University of California (Berkeley), where it was washed in sterile seawater, examined microscopically for the presence of contaminating organisms, and frozen in liquid nitrogen. Frozen tissue was immediately homogenized by using an RNase-free mortar and pestle, and total RNA was extracted by using Trizol reagent according to the manufacturer's instructions. The total RNA recovered was quantitated by using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE) and analyzed on a formaldehyde–agarose gel to confirm that it was intact. mRNA was isolated from ≈1 mg of total RNA by using the Oligotex mRNA Mini kit (Qiagen, Valencia, CA) spin column protocol and was quantified by using a NanoDrop spectrophotometer.

cDNA Library Construction. Five micrograms of mRNA was used to create a directionally cloned cDNA library in the vector pSport1 by using the Superscript Plasmid system with Gateway Technology and ElectroMAX DH10B-competent cells (Invitrogen, Carlsbad, CA). mRNA was radiolabeled with 1 μCi (1 Ci = 37 GBq) [a-^{32}P]dCTP and size-fractionated according to the manufacturer's specifications. All precipitation steps were omitted, and we instead used Microcon YM-100 columns (Millipore, Billerica, MA) to wash and concentrate products. This method obviated the need to precipitate nucleic acids by using a nucleic acid carrier and aided efficient product recovery. The resulting cDNA library was analyzed for its quality by PCR screening and sequencing >100 clones. The mean insert size was found to be ≈1,250 bp.

EST Sequencing and Editing. Primary library aliquots were sent to the Broad Institute (Cambridge, MA), where 5,760 colonies were picked and sequenced with M13 primers from both directions. Raw sequences were trimmed for quality and vector content by using the DNA contig management software program, Sequencher v.4.5 (Gene Codes, Ann Arbor, MI). The original EST data are available online at http://cigbrowser.berkeley.edu/cgi-bin/oscarella/nph-blast.pl.

EST Annotation. By using a combination of automated and manual techniques, EST sequences were clustered, assembled into contigs, and annotated. Individual reads were originally clustered by using stringent criteria (e.g., 90% similarity) in Sequencher. The Metazome (www.metazome.com), NCBI nr, and GO (www.geneontology.org) databases were used to give a preliminary annotation of the *O. carmela* EST clusters. For the Metazome analysis, the EST sequences were compared for their homology to the jawed vertebrate centroid (consensus) sequences by using an expected (E)-value cutoff of 0.001. For each *O. carmela*-Metazome BLAST hit, the largest summed score of nonoverlapping high-scoring segment pairs (HSPs) was found. The top fifteen such hits from each *O. carmela* EST were then examined manually. Similar analyses were done by using the NCBI nr and GO databases in place of Metazome. Contigs with high similarity to elements of major signaling and adhesion gene families were identified, and their original, unedited chromatograms were subjected to a second round of editing, clustering, and annotation in which contigs were manually checked for their validity, and artificially clustered contigs were ''dissolved.'' The consensus sequences resulting from the second round of annotation were again analyzed for their homology to sequences in the Metazome, NCBI nr, and GO databases. Furthermore, the sequences excluded in the first round of annotation were reexamined to confirm that all relevant contigs were extracted. All unique contigs of potential interest were pooled and assigned candidate homologies based on the consensus between the highest-scoring hits by using all three annotation tools. The E-values listed in the supporting information, which is published on the PNAS web site, represent the most significant E-values obtained by blasting against the NCBI nr or GO databases (scores range from e-05 to e-161).

We thank E. Begovic, A. Collins, M. Levine, K. McDonald, J. Nelson, N. Patel, V. Pearse, A. Rokas, J. Valentine, N. Weitzel, and the members of the King laboratory for advice and valuable comments on the manuscript. This work was supported, in part, by funding from the Pew Scholars Program and the Gordon and Betty Moore Foundation.

1. Kaesler, R. L. (2003) in *Treatise on Invertebrate Paleontology* (Geol. Soc. Am., Boulder, CO, and Univ. of Kansas, Lawrence, KS), Vol. 2, p. 349.
2. Li, C., Chen, J. & Hua, T. (1998) *Science* **279,** 879.
3. Valentine, J. W. (2004) *On the Origin of Phyla* (Univ. of Chicago Press, Chicago).
4. Halanych, K. M. (2004) *Ann. Rev. Ecol. Syst.* **35,** 229–256.
5. Borchiellini, C., Manuel, M., Alivon, E., Boury-Esnault, N., Vacelet, J. & Le Parco, Y. (2001) *J. Evol. Biol.* **14,** 171–179.
6. Medina, M., Collins, A. G., Silberman, J. D. & Sogin, M. L. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 9707–9712.
7. Peterson, K. J. & Butterfield, N. J. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 9547–9552.
8. Peterson, K. J. & Davidson, E. H. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 4430–4433.
9. Peterson, K. J., McPeek, M. A. & Evans, D. A. D. (2005) *Paleobiology* **31,** 36–55.
10. Borchiellini, C., Chombard, C., Manuel, M., Alivon, E., Vacelet, J. & Boury-Esnault, N. (2004) *Mol. Phylogenet. Evol.* **32,** 823–837.
11. Nichols, S. A. (2005) *Mol. Phylogenet. Evol.* **34,** 81–96.
12. Boury-Esnault, N., Ereskovsky, A., Bezac, C. & Tokina, D. (2003) *Invertebr. Biol.* **122,** 187–202.
13. Boute, N., Exposito, J. Y., Boury-Esnault, N., Vacelet, J., Noro, N., Miyazaki, K., Yoshizato, K. & Garrone, R. (1996) *Biol. Cell* **88,** 37–44.
14. Gumbiner, B. M. (1996) *Cell* **84,** 345–357.
15. Pires-daSilva, A. & Sommer, R. J. (2002) *Nat. Rev. Genet.* **4,** 39–49.
16. Tepass, U., Godt, D. & Winklbauer, R. (2002) *Curr. Opin. Genet. Dev.* **12,** 572–582.
17. Technau, U., Rudd, S., Maxwell, P., Gordon, P. M. K., Saina, M., Grasso, L. C., Hayward, D. C., Sensen, C. W., Saint, R., Holstein, T. W., Ball, E. E. & Miller, D. J. (2005) *Trends Genet.* **21,** 633–639.
18. Müller, W. E., Skorokhod, A. & Müller, I. M. (1999) *Acta Biol. Hung.* **50,** 395–411.
19. Skorokhod, A., Gamulin, V., Gundacker, D., Kavsan, V., Müller, I. M. & Müller, W. E. (1999) *Biol. Bull.* **197,** 198–206.
20. Kruse, M., Müller, I. M. & Müller, W. E. (1997) *Mol. Biol. Evol.* **14,** 1326–1334.
21. Schacke, H., Schroder, H. C., Gamulin, V., Rinkevich, B., Müller, I. M. & Müller, W. E. (1994) *Mol. Membr. Biol.* **11,** 101–107.
22. Cetkovic, H., Müller, W. E. & Gamulin, V. (2004) *Genomics* **83,** 743–745.
23. Cetkovic, H., Grebenjuk, V. A., Müller, W. E. & Gamulin, V. (2004) *Gene* **342,** 251–261.
24. Adell, T., Nefkens, I. & Müller, W. E. (2003) *FEBS Lett.* **554,** 363–368.
25. Davidson, E. H. & Erwin, D. H. (2006) *Science* **311,** 796–800.
26. Tyler, S. (2003) *Integr. Comp. Biol.* **43,** 55–63.
27. Harwood, A. & Coates, J. (2004) *Curr. Opin. Cell Biol.* **16,** 470–476.
28. Varner, J. A. (1999) in *Guidebook to the Extracellular Matrix, Anchor, and Adhesion Proteins*, eds. Kreis, T. & Vale, R. (Oxford Univ. Press, Oxford), pp. 484–485.
29. Fedders, H., Augustin, R. & Bosch, T. C. (2004) *Dev. Genes Evol.* **214,** 72–80.
30. Salie, R., Niederkofler, V. & Arber, S. (2005) *Neuron* **45,** 189–192.

31. Mackem, S. (2006) in *Cell Signaling and Growth Factors in Development*, eds. Unsicker, K. & Krieglstein, K. (Wiley–VCH, Weinheim, Germany), Vol. 2, pp. 523–578.
32. Kintner, C. (2003) in *Handbook of Cell Signaling*, eds. Bradshaw, R. A. & Dennis, E. A. (Academic, San Diego), Vol. 2, pp. 813–826.
33. Leys, S. P. & Ereskovsky, A. (2006) *Can. J. Zool.* **84,** 262–287.
34. Maldonado, D. R. & Bergquist, P. L. (2002) in *Atlas of Marine Invertebrate Larvae*, ed. Young, C. M. (Academic, San Diego), pp. 21–50.
35. Kusserow, A., Pang, K., Sturm, C., Hrouda, M., Lentfer, J., Schmidt, H. A., Technau, U., von Haesleler, A., Hobmayer, B., Martindale, M. Q. & Holstein, T. W. (2005) *Nature* **433,** 156–160.
36. Larroux, C., Fahey, B., Liubicich, D., Hinman, V. F., Gauthier, M., Gongora, M., Green, K., Worheide, G., Leys, S. P. & Degnan, B. M. (2006) *Evol. Dev.* **8,** 150–173.
37. Kreis, T. & Vale, R. (1999) (Oxford Univ. Press, Oxford), pp. 568.
38. Fernandez-Busquets, X. & Burger, M. M. (1997) *J. Biol. Chem.* **272,** 27839–27847.
39. Fernandez-Busquets, X. & Burger, M. M. (2003) *Cell. Mol. Life Sci.* **60,** 88–112.
40. Oda, H., Tgawa, K. & Akiyama-Oda, Y. (2005) *Evol. Dev.* **7,** 376–389.
41. Grimson, M. J., Coates, J. C., Reynolds, J. P., Shipman, M., Blanton, R. L. & Harwood, A. J. (2000) *Nature* **408,** 727–731.
42. Chen, J. Y., Oliveri, P., Gao, F., Dornbos, S. Q., Li, C. W., Bottjer, D. J. & Davidson, E. H. (2002) *Dev. Biol.* **248,** 182–196.
43. Erwin, D. H. & Davidson, E. H. (2002) *Development (Cambridge, U.K.)* **129,** 3021–3032.
44. Valentine, J. W., Jablonski, D. & Erwin, D. H. (1999) *Development (Cambridge, U.K.)* **126,** 851–859.