The Analysis of Eight Transcriptomes from All Poriferan Classes Reveals Surprising Genetic Complexity in Sponges

Ana Riesgo,*[†,‡,1] Nathan Farrar,[†,2] Pamela J. Windsor,[†,2] Gonzalo Giribet,[1] and Sally P. Leys[2]

[1]Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University

[2]Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada

[‡]Present address: Department of Animal Biology, Universitat de Barcelona, Barcelona, Spain

[†]These authors contributed equally to this work.

***Corresponding author:** E-mail: anariesgogil@gmail.com.

Associate editor: John H. McDonald

Abstract

Sponges (Porifera) are among the earliest evolving metazoans. Their filter-feeding body plan based on choanocyte chambers organized into a complex aquiferous system is so unique among metazoans that it either reflects an early divergence from other animals prior to the evolution of features such as muscles and nerves, or that sponges lost these characters. Analyses of the *Amphimedon* and *Oscarella* genomes support this view of uniqueness—many key metazoan genes are absent in these sponges—but whether this is generally true of other sponges remains unknown. We studied the transcriptomes of eight sponge species in four classes (Hexactinellida, Demospongiae, Homoscleromorpha, and Calcarea) specifically seeking genes and pathways considered to be involved in animal complexity. For reference, we also sought these genes in transcriptomes and genomes of three unicellular opisthokonts, two sponges (*A. queenslandica and O. carmela*), and two bilaterian taxa. Our analyses showed that all sponge classes share an unexpectedly large complement of genes with other metazoans. Interestingly, hexactinellid, calcareous, and homoscleromorph sponges share more genes with bilaterians than with nonbilaterian metazoans. We were surprised to find representatives of most molecules involved in cell–cell communication, signaling, complex epithelia, immune recognition, and germ-lineage/sex, with only a few, but potentially key, absences. A noteworthy finding was that some important genes were absent from all demosponges (transcriptomes and the *Amphimedon* genome), which might reflect divergence from main-stem lineages including hexactinellids, calcareous sponges, and homoscleromorphs. Our results suggest that genetic complexity arose early in evolution as shown by the presence of these genes in most of the animal lineages, which suggests sponges either possess cryptic physiological and morphological complexity and/or have lost ancestral cell types or physiological processes.

Key words: **RNAseq, signaling, epithelia, adhesion, coordination, sex-determination, innate immunity.**

Article

Introduction

Despite a plethora of genomic data now available current metazoan phylogeny is still in flux, especially with respect to the basalmost branching phyla Porifera, Ctenophora, Cnidaria, and Placozoa. Their branching order is however, fundamental for understanding the early evolution of animal features such as tissues and epithelia, nerves and coordination, immune recognition, and propagation of the germ lineage. Traditional markers such as 18S rRNA tend to place sponges as the sister lineage to the rest of metazoans (Zrzavý et al. 1998; Medina et al. 2001), and while recent hypotheses using transcriptomic and genomic data from ctenophores have challenged this view (Dunn et al. 2008; Hejnol et al. 2009; Ryan et al. 2013), the outcome depends on the type of model and parameters used in analysis of these data sets. Other analyses of the same data either confirm that sponges are the sister lineage to the rest of animals (Pick et al. 2010; Philippe et al. 2011) or fail to resolve this dichotomy (Nosenko et al. 2013).

It is commonly considered that morphological complexity in animals is acquired over evolutionary time (McShea 1996). Sponges are morphologically simple in comparison with ctenophores, which possess complex structures such as gonads, nerves, and muscles, structures that are not known at all in sponges. The absence of certain homeodomains in both ctenophores and sponges led Ryan et al. (2010) to suggest an early branching of ctenophores and sponges prior to placozoans, cnidarians, and bilaterians. In this way, perhaps complex structures may have appeared in two branches of early evolving animals almost simultaneously, one in ctenophores and the other giving rise to the rest of animals. Alternatively, sponges might be derived, and having specialized for a filter-feeding lifestyle, become morphologically simplified by losing ancestral cell types. Therefore, the molecular basis to create complex structures might be still present in sponges even though in structure they appear simple.

Although sponges might appear morphologically "simple," analysis of the genomes and transcriptomes of *Amphimedon queenslandica* (Class Demospongiae) and *Oscarella carmela* (Class Homoscleromorpha) revealed a remarkable molecular complexity (Nichols et al. 2006; Srivastava et al. 2010; Conaco et al. 2012; Nichols et al. 2012). The morphology of *A.*

queenslandica corresponds perfectly well to the textbook view of a sponge: a massive body with branching aquiferous canals lined by a single cell layer—the pinacoderm—and enclosing very few cell types many of which are pluripotent (Simpson 1984). There are however, over 8,500 species of sponge currently recognized with an estimated 12–18,000 more to be described (Appeltans et al. 2012). These live in diverse habitats—from the abyssal deep sea to freshwater lakes and rivers—and have contended with changes caused by uptake of symbionts, infection, changing temperature, salinity, and food abundance over at least 600 My (Conway Morris 1993; Jackson et al. 2007). Some are carnivorous (Vacelet and Boury-Esnault 1995) and many different ways of forming a skeleton, from calcium carbonate, aragonite, to spongin and silica, exist (Simpson 1984; Jackson et al. 2007). Within the context of this extraordinary sponge diversity, the genomes and now transcriptomes of *Amphimedon* and *Oscarella* are informative, but by no means conclusive in terms of providing the absolute gene complement of sponges.

To provide a wider framework for understanding the molecular complexity of sponges, we sequenced the transcriptomes eight sponge species covering all four currently recognized poriferan classes (fig. 1): *Aphrocallistes vastus* (Class Hexactinellida); *Chondrilla nucula*, *Ircinia fasciculata*, *Petrosia ficiformis*, *Spongilla lacustris*, and *Pseudospongosorites suberitoides* (Class Demospongiae); *Sycon coactum* (Class Calcarea); and *Corticium candelabrum* (Class Homoscleromorpha). These species represent for the most part sponges that have been well studied in other contexts (cell biology, ecology, physiology), and for which quality starting material could be obtained. We analyzed protein families and Gene Ontologies, and specifically screened each transcriptome for the presence of genes involved in signaling, neuronal and ionic conduction, epithelia, immunity, and reproduction.

Results

Sequence Assembly and Annotation

The cDNA libraries rendered between 38,866,233 reads for *Pseudospongosorites* and 234,585,429 for *Spongilla* of which, between 67% in *Spongilla* and 87% in *Aphrocallistes* of the reads after the thinning process were used for the assemblies (supplementary table S1, Supplementary Material online). The assembly of each species produced between 10 and 65 Mb of assembled contigs in all species (supplementary table S1, Supplementary Material online).

The average length of the contigs was close to 500 bp in all data sets (supplementary table S1, Supplementary Material online), with the transcriptomes of *Aphrocallistes*, *Sycon*, *Petrosia*, and *Chondrilla* showing the greatest N50 values (N50 is a weighted median statistic such that 50% of the entire assembly is contained in contigs equal to or larger than this value in base pairs). The *Corticium* transcriptome had a large number of short contigs that resulted in low N50 values (supplementary table S1, Supplementary Material online). The average coverage per contig was 190 reads for the transcriptomes of all species in the study (supplementary table S1, Supplementary Material online).

Sponges host a great number of symbionts (mainly bacteria) within their tissues that are impossible to remove prior to cDNA construction. To assess the percentage of sequences that can be assigned to metazoans bacteria and/or protozoans, we compared the results of each independent Blast analysis against separate databases containing metazoan, bacterial, and protozoan proteins (supplementary fig. S1A, Supplementary Material online), and found that most contigs in all data sets returned hits from Metazoa. For the Blast against the Bacteria database, *Corticium* had the highest number of contigs with hits not found in Metazoa and Protozoa (supplementary fig. S1A, Supplementary Material online). Only *Sycon* showed unique hits against Protozoa,

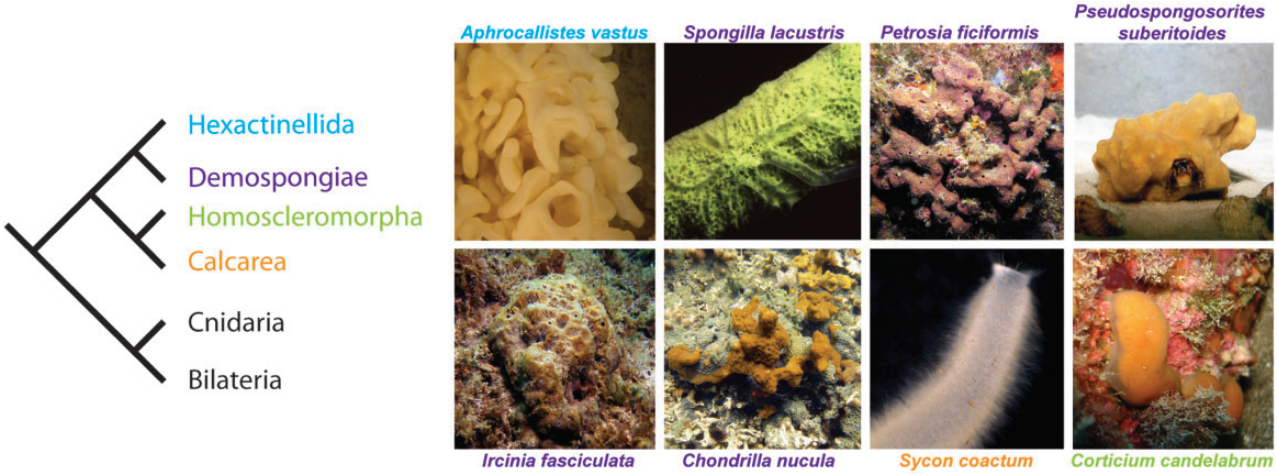

Fig. 1. (*A*) Phylogenetic relationships in the phylum Porifera. (*B*) Images of each sponge used in this study: *Aphrocallistes vastus*, *Spongilla lacustris*, *Petrosia ficiformis*, *Pseudospongosorites suberitoides*, *Ircinia fasciculata*, *Chondrilla nucula*, *Sycon coactum*, and *Corticium candelabrum* (the authors who provided the images are mentioned in the acknowledgements). Colored names indicate their phylogenetic affiliation in (*A*). The species *Amphimedon queenslandica* and *Oscarella carmela*, used for gene comparisons, belong to Demopospongiae and Homoscleromorpha, respectively.

while the other species produced both protozoan and metazoan hits (supplementary fig. S1A, Supplementary Material online).

Among those sequences that blasted to metazoans (more than 60% of the contigs assembled for *Aphrocallistes*; between 40% and 50% in the demosponges; and around 20–30% in *Sycon* and *Corticium*), most hits were assigned to sponges (mostly *Amphimedon*) for the demosponges, and to bilaterians in the case of *Aphrocallistes*, *Sycon*, and *Corticium* (supplementary fig. S1B, Supplementary Material online).

Regardless of the potential different physiological states of the sponges when collected, the percentage of sequences with assigned gene ontology (GO) terms was similar for the ontology categories "biological function," "molecular function," and "cellular component" for all data sets (supplementary fig. S1C, Supplementary Material online), allowing comparisons at that level. It is important to note however, that the total number of GO terms retrieved for each data set was very different, with 32,604 in *Corticium* and only 6,501 in *Petrosia*. For all data sets, in the GO category "biological process," the primary metabolic process was the most abundant term, in the "molecular function" category, catalytic activity was most common, and in "cellular component," macromolecular complex was most abundant.

Protein Families in Porifera

We analyzed the number of protein families (Pfams) in each sponge data set and found the highest number of Pfams in *Corticium* (50,798) and the lowest in *Pseudospongosorites* (10,137). The number of Pfams in *Corticium* could be either due to the high number of symbionts reported in the transcriptomic data set (supplementary fig. S1D and E, Supplementary Material online) or an enrichment of certain domains in this particular species. We found very similar abundances for all protein families in *Aphrocallistes*, *Chondrilla*, *Ircinia*, *Spongilla*, *Petrosia*, and *Pseudospongosorites*, and a different profile in *Sycon* and *Corticium* (supplementary fig. S1D and E, Supplementary Material online). We obtained the functions for all the most abundant protein families and grouped them under the following categories: signaling, cell adhesion, immune system and metabolism and structural/cytoskeletal, and those which showed more than one main function were grouped under "multiple functions" (supplementary fig. S1D and E, Supplementary Material online). Again, in most cases, the number of protein families was higher in *Corticium* and *Sycon* than in the other species. The families showing the larger differences in the "cell adhesion" category were I-set and Laminin_EGF, where both *Sycon* and *Corticium* had higher values. For example, the family MAM was only in *Sycon* and *Corticium*. For the category metabolism and differentiation, *Corticium* had a much larger complement of protein kinases (Pkinase) and sulfatases.

Targeted Gene Study

We used the transcriptome data sets to search for specific genes in pathways related to the acquisition of morphological complexity in metazoans. Absences of genes in any of the species studied here should be interpreted with caution given that transcriptomes convey expressed transcripts. To confirm the presence/absence of all the genes at the genomic level, we carried out the same surveys of the *A. queenslandica* genome and *O. carmela* draft genome (Nichols et al. 2006; Srivastava et al. 2010; Nichols et al. 2012).

Developmental Toolkit Genes

Developmental signaling pathways have been considered a hallmark of metazoan complexity. While most components of the major metazoan signaling pathways are present in sponges, some key absences have been noted (Nichols et al. 2006). We specifically examined the Hedgehog, Wnt, TGF-β, and Notch-Delta pathways. Most components of the Hedgehog (Hh) signaling pathway were present in all sponge classes (fig. 2A). As Hh proteins are composed of an N-terminal signaling domain and a C-terminal Hint domain, similarity to Hh may have simply reflected the presence of one of these domains. We therefore included one row for sequences containing the hedge domain and a row for Hint domains. We found no instances in which both a hedge and a hint domain were in the same sequence, and no true hedgehog proteins were found in any sponge. Finally, Hedgling (Hling) proteins were identified by the whole or partial presence of the domain architecture identified in Adamska, Matus, et al. (2007), as well as simple blast similarity. Importantly, in *Aphrocallistes*, *Corticium*, and *Sycon* we also found *smoothened*, a component of the Hedgehog pathway that was thought to be absent from sponges (Nichols et al. 2006; Adamska, Matus, et al. 2007; Srivastava et al. 2010; Ingham et al. 2011). Phylogenetic analysis indicates that the sponge sequences lie within the smoothened family, and do not cluster together with closely related *frizzled* genes, and predicted 3D structure supports this finding (fig. 2B and C).

Wnts and other Wnt pathway components identified in *A. queenslandica* (Adamska, Degnan, et al. 2007; Adamska et al. 2010), *Suberites domuncula* (Adell et al. 2003, 2007), *Lubomirskia baikalensis* (Harcet et al. 2010), *Ephydatia muelleri* (Windsor and Leys 2010), and *O. carmela* (Nichols et al. 2006; Lapebie et al. 2009) were also found in our transcriptomes (fig. 2). Of note is the possible absence of *Wnt* in *Aphrocallistes*, despite the presence of other *Wnt* signaling components in that sponge.

Our findings confirm and expand the presence of TGF signaling components in all four sponge classes. We found TGF family ligands as well as TGF family receptors and/or activin receptors in all eight sponge transcriptomes. Whereas the TGF-β ligand antagonist *noggin* and the downstream effectors *SMADs* were found in all species, *nodal* was not identified in any of the transcriptomes (fig. 2).

Homologs of *notch* and *delta* were found in all four sponge classes (fig. 2) as had been reported previously for *O. carmela* and *A. queenslandica* (Nichols et al. 2006; Richards and

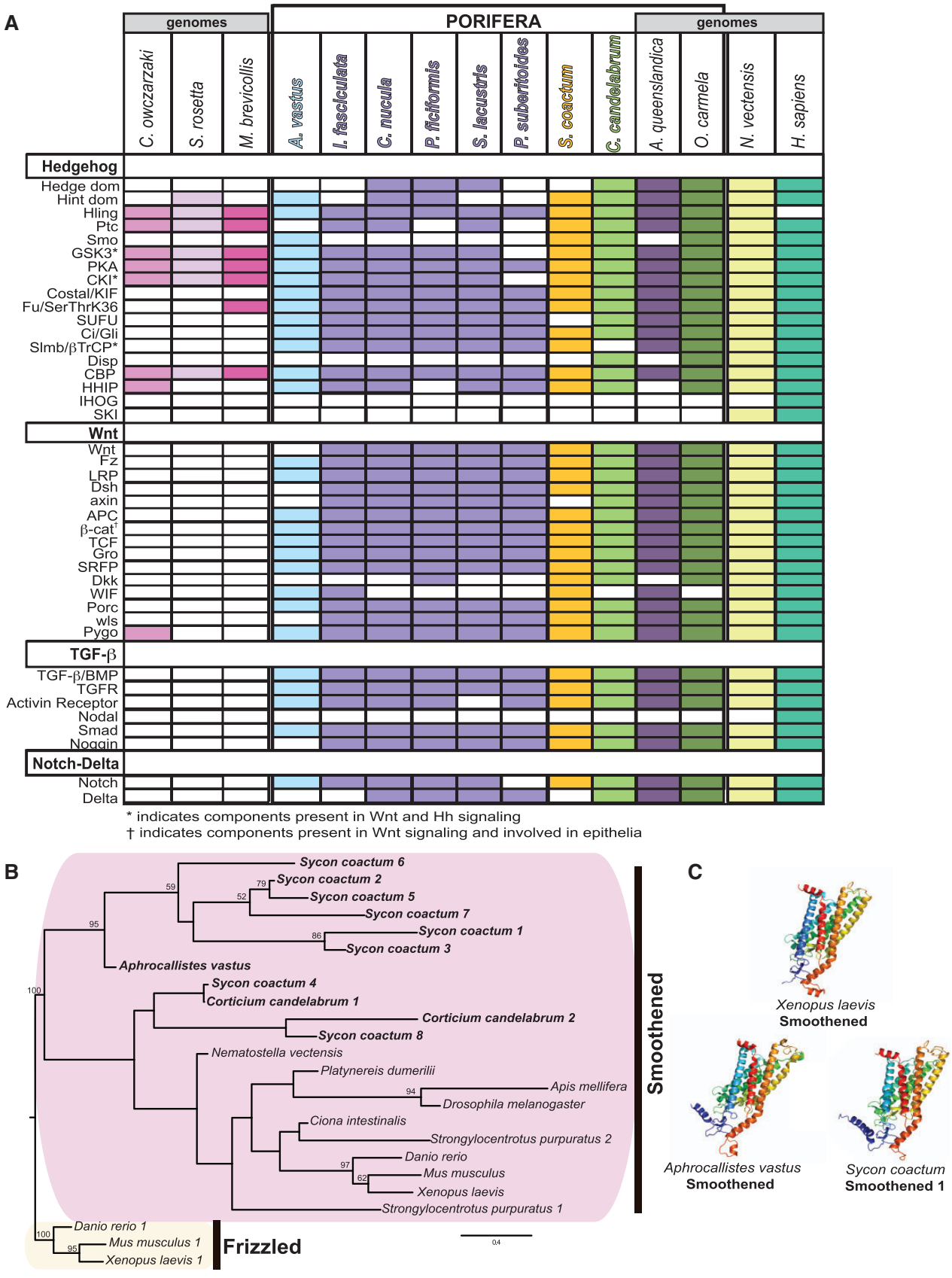

Fig. 2. Signaling molecules identified in Porifera. (*A*) Presence of the genes belonging to the signaling pathways for Hedgehog, Wnt, TGF-β, and Notch-Delta. (*B*) Evolutionary relationships of sponge Smoothened proteins determined with Maximum Likelihood analysis. Bootstrap support is shown for nodes greater than 50% of 500 pseudoreplicates. Accession numbers of sequences used to construct the phylogenetic tree are available in supplementary table S2, Supplementary Material online. (*C*) Three-dimensional reconstructions of the Smoothened proteins in two distant sponges (*A. vastus* and *S. coactum*) and the vertebrate *Xenopus laevis*.

Degnan 2012). Some of the sponge sequences showed characteristics aligning them more closely with *jagged* than *delta*, but phylogenetic analysis confirmed that they cluster with *delta* of other sponges (not shown).

Neuronal Signaling: Postsynaptic Densities and Neurotransmission

Genes associated with postsynaptic densities (PSDs) and signaling via neurotransmitters are shown in figure 3. Perhaps unsurprisingly we found the same general set of genes in all eight transcriptomes as shown previously for *A. queenslandica* (Sakarya et al. 2007). Several genes however, were not found in some sponges, and importantly we were unable to identify some genes previously described from *Amphimedon* in any of the eight transcriptomes. For example, *citron* was not found in *Pseudospongosorites* and *cortactin* was not found in *Petrosia* or *Ircinia*, while *homer* was absent from *Petrosia*. Given the presence of these genes in most of the demosponges we studied, it is likely that the variability reflects the fact that these are transcriptomes (only expressed genes are detected) rather than genomes (where all genes are detected).

Our findings are also broadly consistent with gene absences reported from the *Amphimedon* genome (Sakarya et al. 2007; in supplementary material, Supplementary Material online), although in contrast to Alié and Manuel (2010) we did not find a true *Shaker-type K+* channel in any sponge nor could we find *neurolignin* or *stargazin* in any of the eight transcriptomes (as reported for *Amphimedon* by Sakaraya et al. 2007). We did find a sequence that blasted to a Kv subfamily-A type, which is characterized as shaker-like in *Corticium*, and genes with sequence similarity to shaker-like Kv channel were also found in the three unicellular eukaryote genomes, but it does not necessarily imply Kv channels are present and further characterization is required.

One significant finding that differs from both previous reports on sponge PSD genes was the presence of an *ionotropic glutamate receptor* (iGluR) in *Corticium*, *Sycon*, and *Iricinia*. The iGluRs present in *Corticium* and *Ircinia* appear AMPA-like possessing a Q/R site and all three sponges appear to possess most of the pore motif, SYTANLAAF. Phylogenetic analysis confirmed these channels group with other metazoan iGluRs (fig. 3; supplementary fig. S2, Supplementary Material online).

We attempted to identify core components of the catecholamine signaling pathway (adrenaline, noradrenaline, epinephrine, etc.; fig. 3). Curiously, while we found pieces of the biosynthesis pathway in the transcriptomes, we were unable to show the full pathway in any single sponge species, and some components were missing from all species. For example, dopamine-β-hydroxylase, which catalyzes the reaction of dopamine to norepinephrine, was identified in all transcriptomes, yet *DOPA decarboxylase*, which produces dopamine from L-DOPA could only be identified in *Sycon* and *Corticium*. Furthermore, while gene prediction suggests that the *Amphimedon* genome encodes a *tyrosine hydroxylase* (GI:340369773) an enzyme that catalyzes the reaction of tyrosine to L-DOPA, this gene was not identified in any of the eight sponge transcriptomes studied here, yet *phenylalanine hydroxylase*, the gene that encodes for a protein that catalyzes the synthesis of tyrosine from phenylalanine was identified in all eight transcriptomes. Finally, adrenergic receptors were identified, but subtypes including a *dopamine receptor* were not. Therefore, while components of the catecholamine signaling pathway are present among the four classes of sponges, a more complete picture of this pathway could not be constructed even from data from all of the new transcriptomes. Nevertheless, we did find *glutamate decarboxylase* (which carries out synthesis of GABA from glutamate), which supports previous reports that show both glutamate and GABA are physiologically active in demosponges (Elliott and Leys 2010).

Adhesion and Epithelia

Genes involved in maintenance of epithelial polarity, in adhesion to other cells and to a basal matrix, and genes involved in secretion of a basement membrane have previously been considered indicative of tissue-level differentiation in some sponges and not others. We found this an ideal opportunity to survey the transcriptomes for the same genes studied previously (Nichols et al. 2012; Fahey and Degnan 2010). Our findings are summarized in figure 5. Unsurprisingly, we found most of the polarity genes *Par3*, *Par6*, *Lgl*, *scribble*, and *disks large* (fig. 4). We had difficulty identifying strict homologs of *Patj*, previously identified in *Amphimedon* (Fahey and Degnan 2010), but in contrast to previous work, we found good evidence for *stardust/Pals*, and the ligand of *stardust*, *crumbs*, in all sponge transcriptomes. In terms of adhesion, we found protocadherin in all sponges, and the components typically associated with *cadherin* adhesion (*b-catenin*, *alpha-catenin*, *p120 catenin*, and *vinculin*) were all present. What was unusual in comparison with earlier surveys of the *Amphimedon* genome was the presence of homologs of *claudin* in the three demosponges (*Spongilla*, *Pseudospongosorites*, and *Chondrilla*) as well as in the homoscleromorph *Corticium* and in the calcareous sponge *Sycon* (fig. 4).

Also in contrast to previous work, we found homologs of important components of basement membrane genes, including *type IV collagen* in *Spongilla, Ircinia, Chondrilla, Sycon, and Corticium* (the latter two were shown previously by Leys and Riesgo 2012). *Perlecan* and *nidogen*—molecules that connect the cell membrane to the protein type IV collagen—were also found in all except two of the demosponge transcriptomes (*nidogen* was not identified in *Ircinia*, and *perlecan* was not found in *Pseudospongosorites*). The genes for Laminins, which play a fundamental role in basement membrane assemblage as well as focal adhesion to the extracellular matrix, are composed of three nonidentical chains, alpha, beta, and gamma, whose specific functions depend on the tissue in which they are present (fig. 4). The three chains (alpha, beta, and gamma) were only found in *Corticium* and *Chondrilla* (fig. 4), whereas in the other sponge transcriptomes we found only two of the chains, or just one in the case of *Pseudospongosorites* (fig. 4).

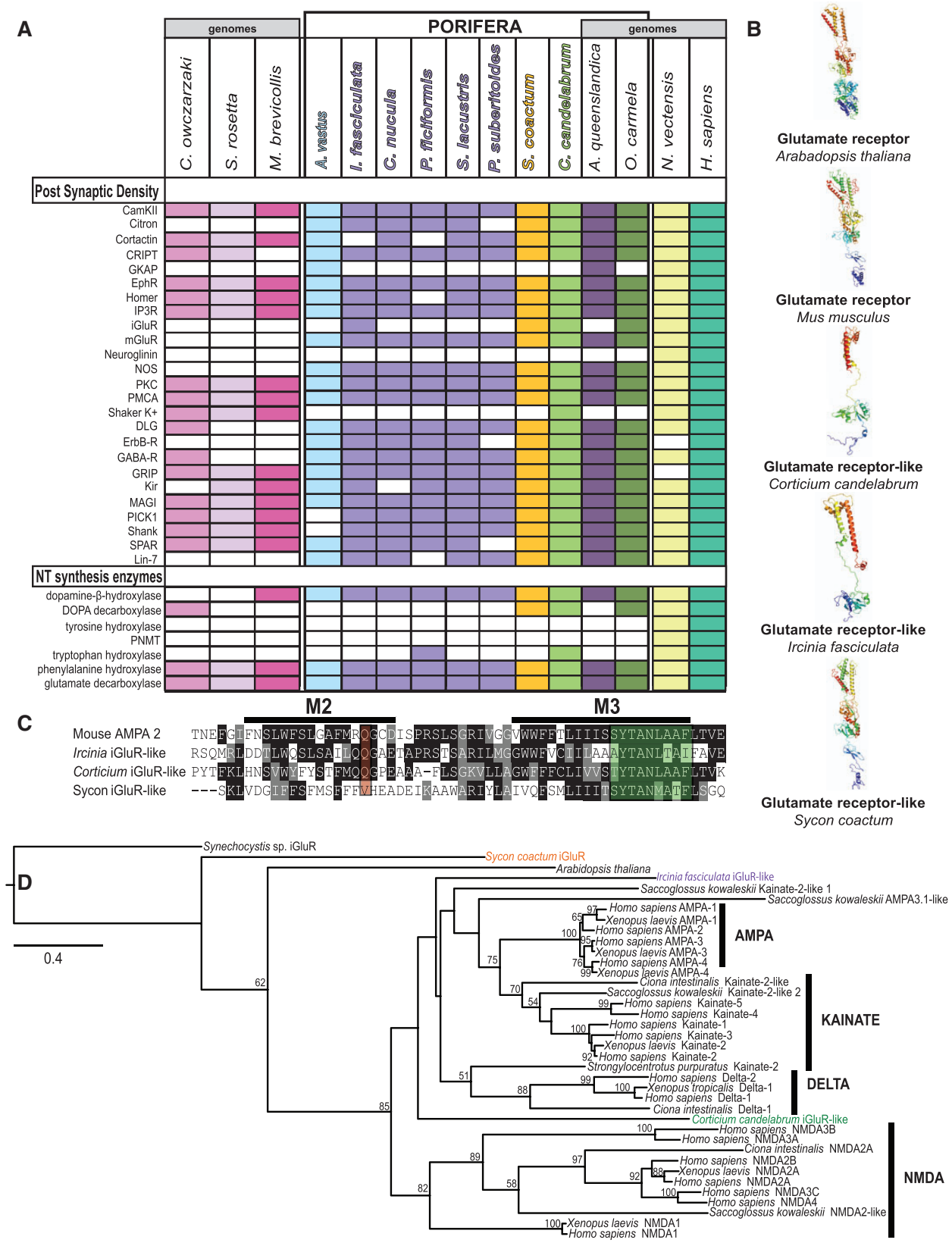

Fig. 3. PSD molecules identified in Porifera. (*A*) Presence of genes comprising the PSD and those involved in synaptic neurotransmission in metazoans. (*B*) Three-dimensional reconstructions of the ionotropic glutamate receptors in the plant *Arabidopsis thaliana*, the vertebrate *Mus musculus*, and three sponges, *I. fasciculata*, *C. candelabrum*, and *S. coactum*. (*C*) Alignment of the ionotropic glutamate receptor proteins from the mouse (*M. musculus*; AMPA type) and the three sponges *I. fasciculata*, *C. candelabrum*, and *S. coactum*. (*D*) Evolutionary relationships between ionotropic glutamate receptors found in bacteria, plants, and metazoans determined with Maximum Likelihood analysis. Bootstrap support is shown for nodes greater than 50% of 500 pseudoreplicates. Accession numbers of sequences used to construct the phylogenetic tree are available in supplementary table S2, Supplementary Material online. Colored sequences are matched to the sponges in figure 1*A*.

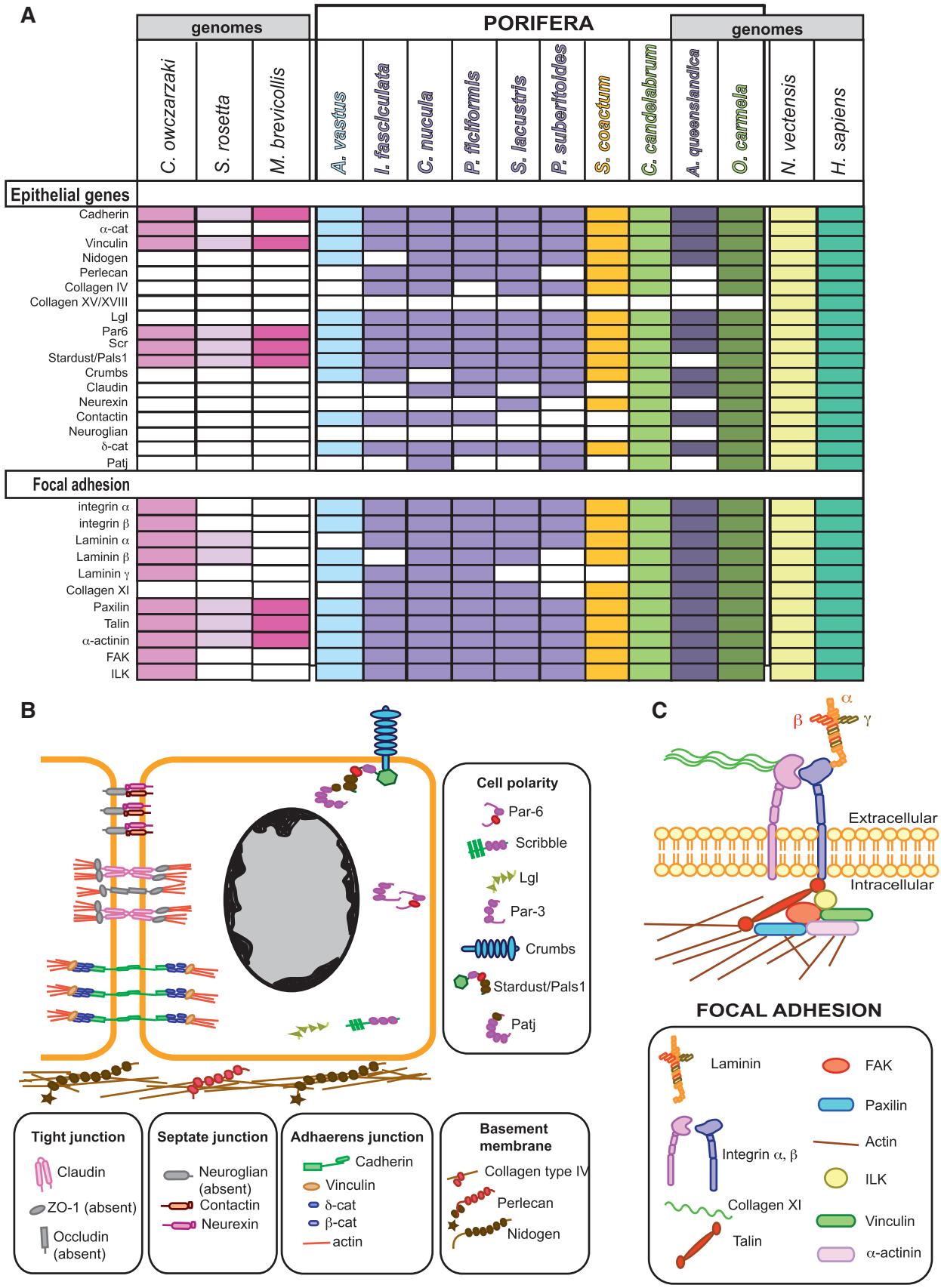

FIG. 4. Cell adhesion, focal adhesion and epithelial development molecules identified in Porifera. (*A*) Presence of the genes involved in cell adhesion, epithelia formation, and focal adhesion. (*B*) Schematic depicting the cell adhesion, polarity signaling, and basement membrane proteins found in metazoans. (*C*) Schematic depicting focal adhesion molecules found in the cell membranes of a metazoan.

Adhesion of cells to the surrounding extracellular matrix together with their stimulation by growth factors are key features that help cells to survive, proliferate, differentiate, or migrate in all animals (Turner 2000; Labouesse and Georges-Labouesse 2003). Cell adhesion is enabled via transmembrane *integrins* and their coupling with extracellular components such as *collagen* and *laminins* as well as their anchoring to *actin* through several protein components such as *focal adhesion kinase*, *paxillin*, *talin*, *integrin-linked kinase*, and *vinculin*. We found all basic components of this mechanism (*focal adhesion kinase, paxilin, talin, integrin alpha and beta, filamin, alpha-actinin*, and *vinculin*) in all eight transcriptomes (fig. 4). In addition, we found the fibrillar *collagen XI*, known for giving support to connective tissues in mammals, in all species except for *Pseudospongosorites* (fig. 4).

Innate Immunity

While in vertebrates the immune system has a two-tier system consisting of either phagocytic activity or the opsonization and direct lysis of pathogens via the complement cascade, basally branching invertebrate phyla typically lack phagocytic activity and only have the ability to detect, contain, and kill pathogens (Miller et al. 2007). The complement cascade has been fully described in three cnidarian species (Miller et al. 2007), but the *Amphimedon* genome has important absences (Srivastava et al. 2010). We focused on the 11 major gene families involved in immunity, as shown in Miller et al. (2007) and Srivastava et al. (2010). We found all of them in almost all eight sponge transcriptomes, with a few exceptions (fig. 5), and significantly, there was only one sequence in the unicellular eukaryote genomes (fig. 5). The *nuclear factor kappa-light-chain-enhancer of activated B cells* (NF kB), the *interleukin receptor-associated kinase 1/4* (IRAK 1/4), *TGF-b activated kinase* (TAK-1), the *TNF receptor-associated factors* (TRAF), and the *interferon regulatory factor* (IRF), were found in all species of sponges and the latter was also found in *Capsaspora* (fig. 5). The *Toll/interleukin 1 receptor 2* (TLR2) was found in *Ircinia, Petrosia,* and *Corticium* (fig. 5). In contrast, the *myeloid differentiation primary response 8* (MyD88) gene was found in all sponges except *Sycon* (fig. 5). α_2-*Macroglobulin* (A2M) is an evolutionarily conserved element of the innate immune system whose best-characterized function is the clearance of active proteases from the tissue in many animals (Armstrong and Quigley 1999); it is thought to be absent in *Amphimedon* (Srivastava et al. 2010), but we found *A2M* in all the transcriptomes. Similarly, the *mannose-binding lectin associated serine protease* (MASP), which is responsible for activation of the *lectin* complement pathway (Iwaki et al. 2011) was not found in *Amphimedon* (Srivastava et al. 2010) but was found in the *Corticium* transcriptome (fig. 5).

Reproductive Machinery: Germ Line, Sex Determination, Pheromones, and Vitellogenesis

The ability of differentiated cells to dedifferentiate into dedicated reproductive cell populations (gametes) is exclusive to multicellular animals. Whether these cells are segregated early in the development of the individual, or are continually transformed from undifferentiated cells varies among animal phyla. In sponges, it appears that the mechanism of gamete determination is triggered by environmental cues and involves somatic cell differentiation into gametes (Riesgo and Maldonado 2008). A well-known germ line machinery exists in metazoans (Ewen-Campen et al. 2010), even though some of the genes may be involved in maintaining totipotency and not specifically in germ line determination (Juliano and Wessel 2010). Knowing what genetic machinery used for germ line (and eventually gamete) specification and sex determination (Miller et al. 2003) exists in sponges can shed light in the evolution of reproduction in metazoans. Of the 20 genes known to be involved in determination of the germ line, we found 11 (with some exceptions) in sponges (fig. 6). The genes *germ cell-less* and *pumilio* were not present in *Aphrocallistes* and *Pseudospongosorites*, and *boule* was not found in *Chondrilla* (fig. 6).

For sex determination, all metazoans investigated use *Dmrt* genes, which work as tissue-specific developmental regulators that integrate information about sex, position, and time to direct narrow populations of cells toward male or female fates (Kopp 2012). The sex determination gene *DMRT1* was found exclusively in *Corticium*, while *FEM-1* (a gene involved in gamete specification that appears broadly in metazoans; e.g., Mckeown and Madigan 1992) was found in all sponge transcriptomes (fig. 6). Another important event genetically and environmentally regulated after gametogenesis is gamete release, which is usually synchronized using pheromones in marine invertebrates (Hardege and Bentley 1997; Painter et al. 1998; Counihan et al. 2001). Even though it has been suggested that pheromones may synchronize gamete release in *Neofibularia nolitangere* (Hoppe and Reichert 1987), it is only very recently that the presence of a pheromone precursor was shown in a demosponge (Novo et al. 2013). In our study, the transcriptomes from all sponge species contained the precursor of the pheromone *attractin* (fig. 6) and there was a high degree of conservation of amino acid sequence in all sponges except *Sycon*.

Vitellogenesis is also a fundamental reproductive process that occurs during gametogenesis not only in sponges but also in all metazoans; it allows embryos and lecithotrophic larvae to survive until they develop feeding structures. The variety of processes converging in the formation of a yolk platelet is remarkable in sponges, as are the various morphologies of yolk (Riesgo and Maldonado 2009), but the genetic regulation of the yolk formation has been investigated only in bilaterians (Wiley and Wallace 1981; Bownes 1986). In our data sets, one or several *vitellogenin* genes were also found in all species except for *Aphrocallistes* (fig. 6), even though the sequences were very divergent among species (not shown).

Discussion

We searched the transcriptomes of eight sponges for genes that have been considered important for metazoan body organization and function. We also checked for the presence of these genes in three well-referenced unicellular eukaryotic genomes, *Capsaspora owzcarzaki*, *Monosiga brevicollis*, and

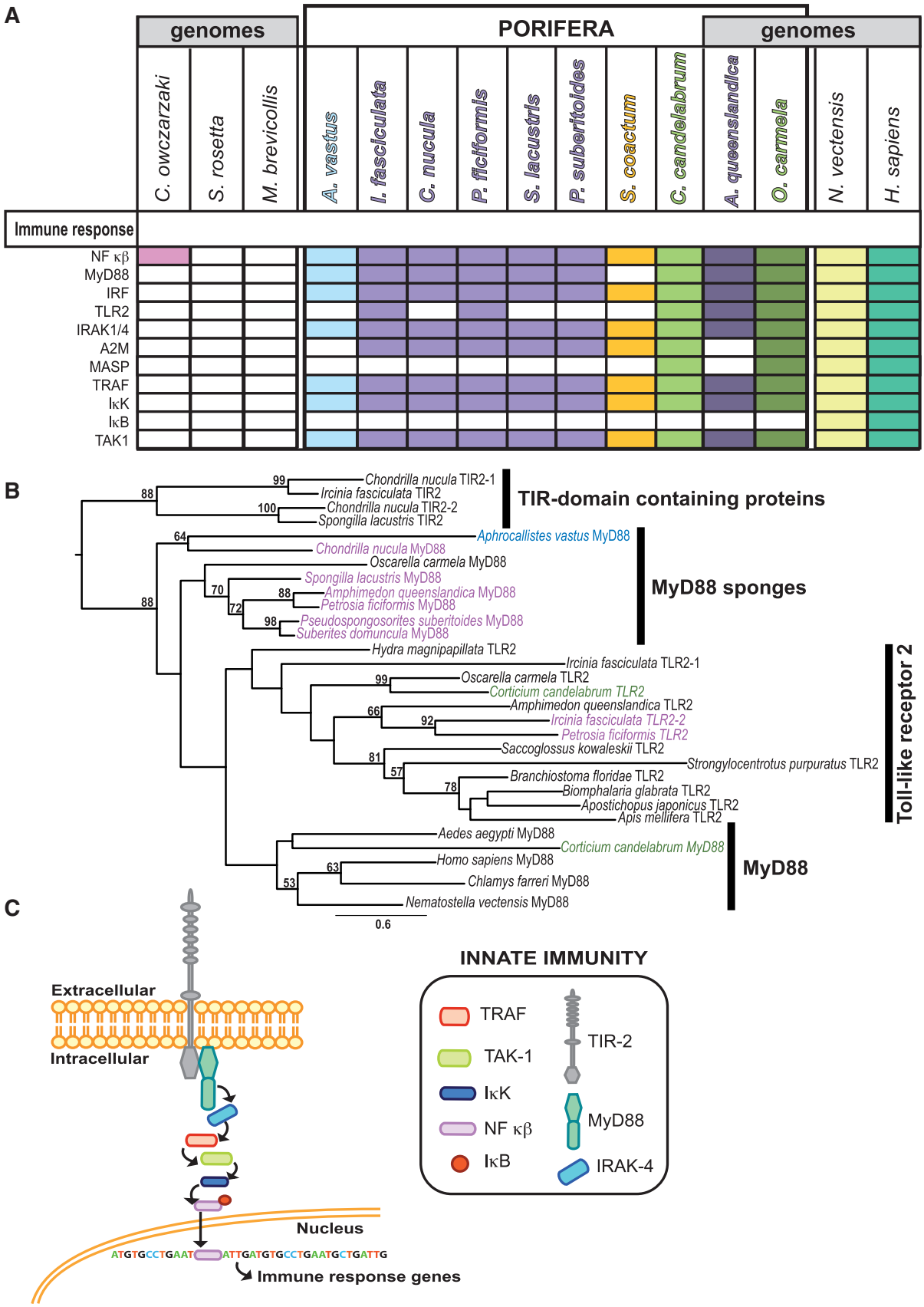

FIG. 5. Innate immunity molecules identified in Porifera. (*A*) Presence of the genes involved in the innate immune response in metazoans. (*B*) Evolutionary relationships of sponge *MyD88* and *Toll-like receptors* determined with maximum likelihood analysis. Bootstrap support is shown for nodes greater than 50% of 500 pseudoreplicates. Colored sequences are matched to the sponges in figure 1*A*. Accession numbers of sequences used to construct the phylogenetic tree are available in supplementary table S2, Supplementary Material online. (*C*) Schematic depicting innate immunity molecules in the cell membrane of a metazoan.

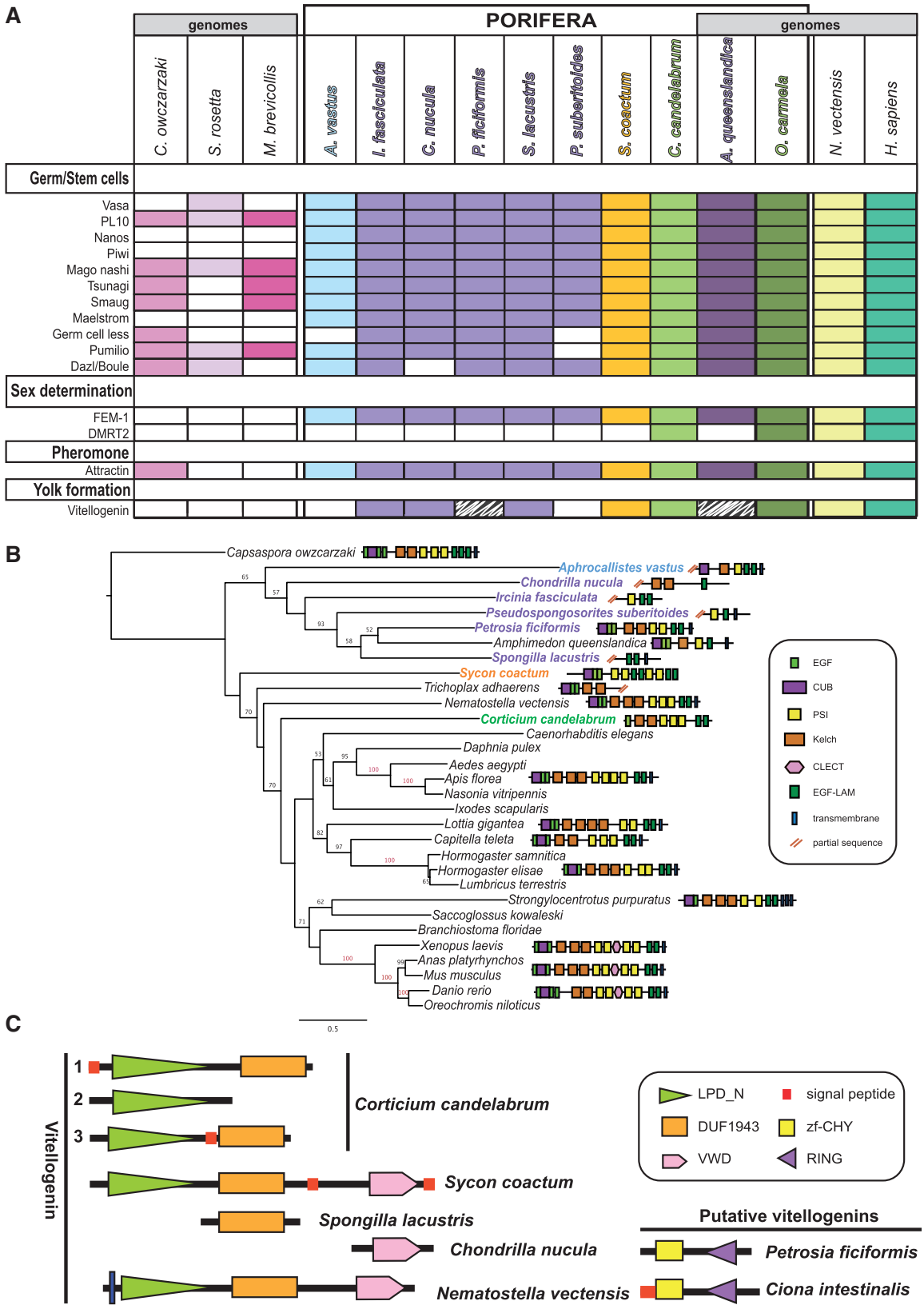

Fig. 6. Reproductive, sex determination, pheromone, and vitellogenesis genes identified in Porifera. (*A*) Presence of the genes involved in germ line and sex determination, pheromone communication, and vitellogenesis in metazoans. Hashmarks denote the presence of a sequence with similarity to vitellogenin-like proteins that are highly divergent from vitellogenin. (*B*) Evolutionary relationships of the sponge pheromone precursor protein attractin as determined with Maximum Likelihood analysis. Bootstrap support is shown for nodes greater than 50% of 500 pseudoreplicates. Colored sequences are matched to the sponges in figure 1*A*. Accession numbers of sequences used to construct the phylogenetic tree are available in supplementary table S2, Supplementary Material online. (C) Protein schematic illustrating the domain structure of vitellogenin and putative-vitellogenin proteins in sponges, cnidarians, and ascidians.

Salpingoeca rosetta, and confirmed their presence in two other metazoan genomes (the cnidarian *Nematostella vectensis* and the vertebrate *Homo sapiens*). We found that few genes in the sponge transcriptomes were sponge-specific—sponges shared between 20% and 50% of genes with other metazoans, between 6.5% and 32% with other eukaryotes (protozoans), and a moderate number were bacterial (between 6% and 24%)—either from bacterial symbionts in the sponge, or in the water they filter; it is unlikely they arise from horizontal gene transfer, which can be the case in premetazoans (Tucker 2013). With respect to Pfam domains, *Corticium* contained the highest number, which is similar to that found in the transcriptomes of annelids and molluscs (45,000–59,000) and higher than arthropods (~35,000) (Riesgo, Andrade, et al. 2012). However, the transcriptome of *Corticium* may also contain high numbers of bacterial Pfams due to the abundance of symbionts in its mesohyl. From the genes shared with metazoans, in most of the sponges studied more than 50% were most similar to bilaterian genes. In fact, *Aphrocallistes*, *Sycon*, and *Corticium* showed less than 25% of similarity between their genes and the sponge genes in the NCBI databases; 75% were more similar to bilaterian genes. The number of annotated genes in the sponge transcriptomes was very similar to that of other nonmodel organisms that also have very few genetic resources (Riesgo, Andrade, et al. 2012; Pérez-Porro et al. 2013). More than 30% of the transcriptomic contigs were not assigned any annotation, which highlights the necessity for a greater effort in sequencing and annotation of sponge genomes and/or transcriptomes.

The complement of genes present in sponges appears far greater than previously understood from the single and now well-studied demosponge species *A. queenslandica* (Srivastava et al. 2010). The publication of the transcriptome of *O. carmela* suggested that homoscleromorph sponges have a far greater complexity than demosponges (Nichols et al. 2012). Yet, the demonstration here of the broad presence of genes in so many functional categories—development, signaling, adhesion, epithelia, immune recognition, and germ-lineage/specification—across Demospongiae, Homoscleromorpha, Calcarea, and Hexactinellida shows that sponges are universally much more complex at the molecular level than previously appreciated.

What then defines metazoan complexity? We found a number of genes which previously had been associated with complex structures of metazoans and thought to be absent in sponges. For example, we found homologs of *smoothened*, *type IV collagen*, and *iontropic glutamate* receptors in several of the sponges. We found also quite clear differences across the sponge classes—almost all metazoan genes were found in the homoscleromorph *Corticium*; *Sycon* shared many of these, but the sequences in *Sycon* though blasting with high e value to the same genes, and folding to the same proposed 3D structure, always had highly divergent amino acid sequences. Interestingly we also found many of these homologs in the glass sponge *Aphrocallistes*. Some of these differences could be due to variation in transcriptome quality (coverage and length of contigs—the most complete transcriptomes were those of *Petrosia*, *Corticium*, *Sycon*, and *Aphrocallistes*), but they may also reflect differences in the length of time that the sponge classes have been separated, and in their level of tissue/functional complexity. The fact that sequences found in *Corticium*, *Sycon*, and *Aphrocallistes* were more similar to bilaterian sequences than to sponge sequences in the databases supports the notion that compared with other sponge groups Demospongiae, broadly speaking, have diverged significantly from other metazoans (i.e., many of these traits are apomorphic).

Gene Searches

We looked for a total of 127 genes involved in development, neuronal and epithelial signaling, immunity and reproduction. Out of those 127, a total of 100 (78% of the genes) were already identified in the *Amphimedon* genome (Srivastava et al. 2010). In our study, we found 119 genes (18 more than in the *Amphimedon* genome) that were thought to be absent in sponges, mainly in *Corticium* and *Sycon* (discussed later), which brings the percentage of genes shared by sponges and other metazoans in the pathways studied here to 93%.

Developmental Toolkit Genes

In unicellular opisthokonts such as *Monosiga* and *Salpingoeca* components of the conserved metazoan toolkit signaling pathways are mostly absent (King et al. 2008; Fairclough et al. 2013); our survey of the genome of *Capsaspora* largely agreed with that finding and confirmed the apparent absence of signaling pathway components outside Metazoa. Notably, however, in *Dictyostelium*, the signaling system involved in forming fruiting bodies contains several elements considered to be critical for Wnt signaling: *Fz receptors*, *GSK3β*, and *β-catenin* homologs (called *GSKA* and *aardvark* in this organism) and a *dkk* (Guder et al. 2006; Harwood 2008). Interactions between some members of the Wnt pathway therefore may be predate the origin of metazoans. We found that sponges possess many metazoan toolkit signaling genes in the Wnt, Hedgehog, TGF-β, and Notch-Delta pathways, and our study of TGF-β and Notch-Delta signaling gave results generally consistent with what is already know from *Amphimedon* (Richards et al. 2008; Richards and Degnan 2009; Srivastava et al. 2010; Richards and Degnan 2012).

The lack of a Hedgehog ligand is unsurprising as it is missing from the genome of *Amphimedon* (Nichols et al. 2006; Adamska, Matus, et al. 2007; Srivastava et al. 2010). It was found in expressed sequence tags (ESTs) from *O. carmela* (Nichols et al. 2006), but the fragment did not contain all diagnostic domains and the characteristic domain structure of a full-length *hedgehog*, therefore these data should be considered cautiously (see Matus et al. 2008 and Ingham et al. 2011 for further discussion). Interestingly, homologs of the *smoothened* gene in *Aphrocallistes*, *Sycon*, and *Corticium* were found and these new findings push the origin of Smoothened genes further back in metazoan evolution. The lack of *smoothened* in demosponge transcriptomes

further supports the idea that the demosponges may have lost certain genes. Furthermore, this trend highlights the need for genomic data from a wider variety of basal branching metazoans—especially sponges—to allow a more complete assessment of the origins of signaling pathways and other characteristic metazoan genes.

Neuronal Signaling

Sponges lack conventional neuronal signaling systems and so it is intriguing that molecules of the protein-rich PSD have been characterized in *Amphimedon* (Sakarya et al. 2007; Alié and Manuel 2010; Srivastava et al. 2010). We found that PSD genes are present in all four classes of poriferans and there is little variation among species, demonstrating that PSD genes were present in the poriferan ancestor. Other genes known to be involved in the development of nervous systems in metazoans (neuralians sensu Nielsen 2012) have been identified in the *Amphimedon* genome, such as *elav- mushashi-like* RNA binding genes, neural transcription factors like *Notch*, *Delta*, and *BHLH* (Richards et al. 2008; Gazave et al. 2009; Richards and Degnan 2012). These genes are also widely expressed in the sponge transcriptomes we studied. Molecules involved in signaling (e.g., *G-coupled receptors* [*GPCRs*]), and neuroendocrine secretion are known in part from the *Amphimedon* genome (Srivastava et al. 2010), but clearly this is where sponges vary in complexity.

We found the first evidence of a rapid, ionic-based receptor in sponges, the *ionotropic glutamate receptors* (*iGluRs*). Both *Corticium* and *Ircinia* transcriptomes have sequences with good similarity to vertebrate *iGluRs*, and we found a similar although more divergent sequence in the calcareous sponge *Sycon*. This pattern of divergence was noted in many genes identified in *Sycon*, an observation that may simply indicate an accelerated rate of evolutionary change in that lineage. In contrast to sponges, in the unicellular eukaryote *Capsaspora* only slightly more than half the PSD genes were present. In general, a small number of structural elements of the PSD that in other metazoans lie deeper within the cell can be found in unicellular eukaryotes, while sponges seem to possess a larger set of PSD genes, notably with an increase in the presence of receptor and signaling molecules. Our results are consistent with the scenario summarized by Ryan and Grant (2009), in which the PSD evolves by adding complexity through the addition of channels and receptors while leaving the underlying scaffolding largely intact. The PSD—as a structure—therefore appears to be a characteristic of neuralians.

In contrast, components of classical neurotransmitter synthesis pathways do not seem to form a coherent group, but rather are scattered throughout the different sponge transcriptomes. The presence of genes encoding for enzymes known to be important components of neurotransmitter synthesis, yet lacking the full synthesis pathway could be viewed as evidence of gene loss. An alternate explanation however, is that some enzymes associated with neurotransmitter synthesis are involved in the production of secondary metabolites in the sponge, rather than classical neurotransmitters to be used for signalling. For example, the gene *DOPA decarboxylase*, which codes for an enzyme involved in the synthesis of both dopamine and serotonin, was found in both *Corticium* and *Sycon* yet neither type of receptor was unambiguously identified in any transcriptome. However, a number of serotonin-derived alkaloids have been identified from the demosponges *Hyrtios erectus* and *H. reticulatus* (Salmoun et al. 2002). This may explain why serotonin has been visualized in sponge tissue and why some synthetic enzymes are present, yet why no clear functional role for the neurotransmitter has been demonstrated (Emson 1966; Lentz 1966; Weyrer et al. 1999; cf. Ellwanger and Nickel 2006); we were also unable to find sequences for a serotonin receptor in the genomes of *Amphimedon* and *Oscarella*. However, the ubiquitous presence of other enzymes such as *glutamate decarboxylase* among all the sponges is consistent with the demonstrated physiological roles for GABA in sponge physiology (Elliott and Leys 2010). Full genomes showing the concrete absence of any molecules, and careful physiological assessments of neurotransmitter effects on sponges, are needed to fully appreciate the roles these enzymes and molecules might play in sponge physiology and behavior.

It is somewhat surprising not to find sequences for voltage gated potassium channels (Kv) in most of the sponges. Until the genomes of these animals can be surveyed it is difficult to draw conclusions about these absences; however, such a consistent absence in most of the sponge transcriptomes could also be an indication of the lack of the need in sponges for rapid changes in membrane potential, typically mediated by Kv channels. The next obvious step will be to experimentally characterize the *shaker-like* sequences we found in the unicellular eukaryotes, and in *Corticium*.

Adhesion and Epithelia

Epithelia are complex and highly versatile structures and are one of the unifying characters of multicellular organisms. Even aggregates of unicells form epithelial-like characters such as adherens junctions, for example, when *Dictyostelium* amoebae congress to form fruiting bodies (Dickinson et al. 2011, 2012). Adherens junctions probably provide support in the raised structure formed of clones of cells. But the full complement of epithelial characters requires proteins that allow adhesion, sealing, polarity, and stability. When these features were assessed for *Amphimedon* (Fahey and Degnan 2010), it was determined that sponges possessed genes allowing polarized epithelia, but lacked the essential conventional molecules typically thought to seal the epithelium from the environment or to stabilize it by attachment to a basement membrane. These conclusions were thought to support the absence of morphological structures for occlusion or a basement membrane in most sponges. In contrast, where these genes were found in *Oscarella*, their presence was justified by ultrastructure showing a typical basement membrane, although it was not considered what this structure would be needed for in a homoscleromorph sponge and not in a demosponge or calcareous sponge.

We found transcripts for genes with homology to *claudin*—involved in sealing the spaces between cells in deuterostomes—and surprisingly *type IV collagens* and other basement membrane genes such as *nidogen* and *perlecan*, which attach *type IV collagen* to the cell's plasma membrane, were present in nearly all sponges. Overall, there was no pattern of presence/absence of these genes across the sponge classes that might provide a hint as to the lineage of evolution. The collagen type IV amino acid sequence from *Spongilla* was the most divergent, which might reflect the recent radiation of sponges into freshwater and the challenges of that environment. Indeed, the changes involved in the marine–freshwater transition would be exciting to revisit with a survey of ion channels and transport molecules in addition to sealing of epithelia.

Innate Immunity

Immune genes evolve at an extraordinarily rapid pace, which makes it difficult to draw up hypotheses about their evolution (Hughes 1997; Hibino et al. 2006). Presumably, the pace of mutation is driven by intense selection in the interplay between host and pathogen. As a consequence, finding immune gene homologs with standard molecular strategies and inferring primitive states is a difficult task. The *Amphimedon* genome encodes several molecules involved in innate immunity including *Nod-like* and *Toll-like receptors*, *IRAK*, *MyD88*, *IRF*, and *IKK* (Srivastava et al. 2010); many immune molecules, however, seem to be largely "eumetazoan" acquisitions (i.e., found in Cnidaria, Ctenophora, and Bilateria). In contrast, we found the most complete molecular machinery involved in sponge immune response to date, finding all genes involved in the innate immune response pathway described in basal invertebrates (Miller et al. 2007). Also, two genes previously described as "eumetazoan" acquisitions, *alpha-2 macroglobulin* and *mannose-binding lectin associated serine protease*, were found in sponges, even though they were absent in *Amphimedon* (Srivastava et al. 2010). While *A2M* was found in all sponges classes, *MASP* was only in *Corticium*. Only one of the components from the selected innate immunity response toolkit was present in one of the three unicellular eukaryotes, similar to that found in the study of Song et al. (2012). Therefore, our results indicate an ancient origin of the innate immune response in metazoan evolution, which predated the separation of sponges and other metazoans.

Reproductive Machinery

In sexually reproducing animals, germ cells are the source of gametes in the adult (Lin 1997). Germ cells carry the hereditary information for the next generation, thus their segregation and protection from a somatic cell fate is essential for animal development and evolution (Buss 1988; Saffman and Lasko 1999; Wylie 1999; Raz 2000). Modern studies identify primordial germ cells more often by the localization of the products of germ-specific genes (Extavour and Akam 2003; Ewen-Campen et al. 2010). Very recently *vasa*, *nanos*, *piwi*, and *PL10* genes were isolated from sponges (Mochizuki et al. 2000, 2001; Funayama et al. 2010; Srivastava et al. 2010), but whether the complete germ line machinery is present in all sponge classes was not investigated. We found that all classes of sponges possess all the genes reported necessary for germ cell determination (note that those specific to *Drosophila* [e.g., *oskar*] were not found). Interestingly, even though sex has not been reported for the three unicellular eukaryotes surveyed here, we found germ line markers in their genomes (*PL10*, *mago nashi*, *smaug*, *pumilio*, *germ cell-less*, and *boule*). The presence of germ line genes in unicellular eukaryotes seems to support the alternative suggestion for the function of these genes: it may be that germ line markers originated in multipotent cells, where they maintain multipotency, and were subsequently co-opted by more specialized embryonic germ cells to determine their germ fate (Juliano and Wessel 2010). Whether these genes are used by sponge cells to maintain multipotency or determine their germ fate remains unknown, but the fact that they possess the complete molecular program for the germ line specification could indicate its potential role in germ line determination, but should be evaluated more closely for their role in determining a germ cell lineage in metazoans.

Sex determination in metazoans involves a wide array of solutions, from splicing-based mechanisms in insects to endocrine regulation in mammals (Kopp 2012). However, the occurrence of the sex determining factors *Dmrt* in all metazoans investigated has emerged as a common theme in sexual dimorphism. The main function of *Dmrt* genes in the gonads of metazoans is to promote male-specific and repress female-specific differentiation (Kopp 2012). Interestingly, we only found an ortholog of *Dmrt1* in *Corticium*, which is a hermaphroditic organism with a remarkable similarity in the gametogenic process with that of other metazoans, for instance, in the continuous and asynchronous production of sperm in the cysts (Riesgo et al. 2007). Likewise, a *Dmrt*1 ortholog is expressed in the hermaphroditic *Acropora millepora* during sexual reproduction (Miller et al. 2003). Therefore, our results could suggest that the sex determination mechanisms involving *Dmrt* genes evolved prior the divergence of Porifera from the rest of metazoans.

Communication via semiochemicals such as pheromones occurs in water by either "sniffing" or by contact chemoreception (Wyatt 2003). The peptide Attractin from *Aplysia* was the first water-borne sex pheromone characterized in invertebrates (Painter et al. 1998); and the full length protein Attractin, which has been found expressed in gonads of mammals (Li et al. 2009) was recently reported in several metazoans including a demosponge (Novo et al. 2013); although the pheromone features were not corroborated. We found that the gene *attractin* was expressed in all classes of sponges, even though it was not originally found in the *Amphimedon* genome. In *Amphimedon*, the gene characterized as *Fanconi anemia group I protein-like* exhibited the highest similarity to *attractin*. Although the potential role of *attractin* on the synchronization of gamete release in sponges remains uninvestigated, it opens the possibility of further research in this novel area.

Vitellogenesis in sponges produces two types of yolk platelets, homogeneous (mainly proteinaceous) and

heterogeneous (lipidic and proteinaceous) (see Simpson 1984; Riesgo and Maldonado 2009 for reviews). The participation of autosynthetic and heterosynthetic (through nurse cells) mechanisms has been described for several species, but the protein precursor has never been characterized in sponges. We found two types of yolk precursors in our sponge transcriptomes, one *vitellogenin* gene in *Ircinia*, *Chondrilla*, *Spongilla*, *Sycon*, and *Corticium*, and a *vitellogenin-like* gene in *Petrosia*. Whether the different genes are involved in the formation of multiple yolk platelets as in *Xenopus* (Wiley and Wallace 1981) needs further study.

Conclusions

This is the first study to survey a wide set of metazoan-specific genes in-depth across all four sponge classes. It provides texture to the question of which molecules might have been present in early animal groups, and more importantly provides the framework for posing new hypotheses on the evolution of multicellularity and animal complexity.

One of the remarkable outcomes of this work is the understanding that most metazoan genes, or the greater complement of genes involved in complex gene pathways, are present in all sponge groups, including genes supposedly absent in the genome of the demosponge *Amphimedon*. Importantly, there are very few key absences (mainly these concern rapid signaling molecules), but overall transcriptomic data sets proved useful to detect complex molecular machineries, even though they are known to contain genes expressed only at a given time.

Determining gene function in sponges is the next challenge. Usually, function of genes in an organism is inferred by comparing to gene function known from other animals. For example, *occludins* are known to seal the epithelium at its apical surface in mice, therefore the presence of *occludins* is taken to imply that sponge epithelia seal. Only one functional study has shown that sponge epithelia are sealed against the outside milieu (Adams et al. 2010), but the species in which this work was carried out, *Spongilla lacustris*, does not appear to possess *occludins* in our transcriptome. In *Spongilla*, therefore, sealing function could be the work of other as yet unspecified molecules.

Here, we show that the great majority of metazoan genes are present in all sponge groups, but sponges do not have conventional structures, behavior, or even mechanisms of development. So either the genes we know from other animals have a different function in sponges, and were co-opted later in the evolution of metazoans for the function we are familiar with, as suggested for adhesion molecules in unicellular eukaryotes (Sebé-Pedrós et al. 2010), or the structure has been lost in sponges. In this sense, the recent publication of the genome of *Mnemiopsis leydi* (Ryan et al. 2013), which places ctenophores as the sister group to all other animals, suggests that structures such as nerves might have been present in the metazoan ancestor and were secondarily lost in placozoans and sponges. Another alternative is that sponge genes might carry out a similar function to that in other animals, but we do not yet recognize that function because we do not distinguish the morphology of such a different structure.

Our comparative transcriptomic analysis strengthens the view that sponge complexity as revealed by their molecular toolkit is poorly reflected in their morphology, as it has also been shown in placozoans (Srivastava et al. 2008). Quite interestingly, our data provide an indication that demosponges have diverged substantially from other classes of Porifera, and highlights the strong similarity of genes in calcareous, homoscleromorph and even hexactinellid sponges with those in other metazoans. Our data also shows that a number of genes are present in calcareous and homoscleromorph sponges but absent in the Silicea (hexactinellids and demosponges). Both the greater number of genes in Calcarea and Homoscleromorpha and the similarity of those sequences in those two species lend support to the suggested sister relationship of these two groups shown in recent phylogenetic studies (Nosenko et al. 2013; Ryan et al. 2013). Taken together the overall view given by patterns shown in gene presence and absence across the four sponge classes supports the idea that sponges are monophyletic with Demospongiae + Hexactinellida and Calcarea + Homoscleromorpha forming sister groups, a hypothesis that is in agreement with the latest phylogenomic analysis using several sponge taxa (Nosenko et al. 2013). One noteworthy observation provided by the data presented by the publication of the ctenophore genome is the remarkable similarity in the gene absences in the PSD and signaling pathways in the ctenophore genome (Ryan et al. 2013) and our study. These absences (and their significance) deserve more attention given that they may support the placement of ctenophores at the base of the metazoan tree.

Materials and Methods

Sample Collection

We collected tissue samples from 8 sponge species, belonging to the four currently recognized classes (fig. 1): *A. vastus* (Hexactinellida), *S. coactum* (Calcarea), *I. fasciculata*, *C. nucula*, *P. ficiformis*, *S. lacustris*, *P. suberitoides* (Demospongiae), and *C. candelabrum* (Homoscleromorpha). Collecting information is provided in supplementary table S1, Supplementary Material online. Hereafter, we refer to each animal by its genus to ease readability.

Sample Preparation

To avoid contamination from epibionts, prior to fixation tissues were cleaned carefully using a stereomicroscope. A piece of sponge tissue was removed with razor blades that were rinsed in RNAseZap (Ambion, TX, USA). All procedures were carried out on ice and quickly to avoid RNA degeneration. Tissues were either flash-frozen in liquid nitrogen and stored at −80 °C or they were immersed in at least ten volumes of RNA*later* at 4 °C for 1 h, incubated overnight at −20 °C, and subsequently stored in the same buffer at −80 °C until RNA was extracted (sometimes samples placed in RNA*later* were transported back to the laboratory at room temperature, where they were stored at −80 °C) (see supplementary

table S1, Supplementary Material online, for details). The amount of tissue used depended on the extent of the spicule skeleton: in most cases 20–80 mg of tissue was used but for *Petrosia* and *Aphrocallistes*, 200 mg was needed due to the large silica skeleton (Riesgo, Pérez-Porro, et al. 2012).

mRNA Extractions

Two different methods of RNA extraction were used: 1) total RNA extraction followed by mRNA purification for *Corticium* and 2) direct mRNA extraction for all other species. Protocols used for both extraction types are available elsewhere (Riesgo, Andrade, et al. 2012; Riesgo, Pérez-Porro, et al. 2012). Total RNA from *Aphrocallistes* was extracted using the Norgen Biotek Animal Tissue RNA Purification Kit (Norgen Biotek, Thorold, ON, Canada). Quantity and quality (purity and integrity) of mRNA were assessed by three different methods, reported in Riesgo, Pérez-Porro, et al. (2012) and shown in supplementary table S1, Supplementary Material online.

Next-Generation Sequencing

For all sponges except *Aphrocallistes*, next-generation sequencing was performed using Illumina GAII and HiSeq2000 (Illumina, Inc., San Diego, CA, USA) platforms at the FAS Center for Systems Biology at Harvard University. mRNA concentrations between 20 and 79.9 ng/µl (supplementary table S1, Supplementary Material online) were used for cDNA synthesis with the TruSeq RNA sample preparation kit (Illumina, Inc.), as described previously (Riesgo, Andrade, et al. 2012; Riesgo, Pérez-Porro, et al. 2012). cDNA was ligated to homemade adapters in one sample of *Petrosia* (5′-ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT GGT T-3′), whereas ds cDNA was ligated to Illumina adapters in the rest of species. Size-selected cDNA fragments of approximately 300 bp excised from a 2% agarose gel were amplified using Illumina PCR Primers for Paired-End reads (Illumina, Inc.) and 18 cycles of the PCR program 98 °C – 30 s, 98 °C – 10 s, 65 °C – 30 s, 72 °C – 30 s, followed by an extension step of 5 min at 72 °C.

The concentration of the cDNA libraries was measured with the QubiT dsDNA High Sensitivity (HS) Assay Kit using the QubiT Fluorometer (Invitrogen, Carlsbad, CA, USA). The quality of the library and of size selection was checked using the "HS DNA assay" in a DNA chip for Agilent Bioanalyzer 2100 (Agilent Technologies, CA, USA). cDNA libraries were considered successful when the final concentration was higher than 1 ng/µl and the Bioanalyzer profile was optimal (Riesgo, Pérez-Porro, et al. 2012). Successful libraries were brought to 10 or 7 nM depending on the initial concentration prior to sequencing. The paired-end reads had lengths of 100 or 150 bp, depending on availability of sequencers (Illumina GAIIx or HiSeq).

The *Aphrocallistes* transcriptome was prepared by LC Sciences (http://www.lcsciences.com/, last accessed February 13, 2014) using 1 µg of total RNA for polyA tail selection of the mRNA (supplementary table S1, Supplementary Material online). Library preparation was performed using also the TruSeq RNA sample preparation kit (Illumina, Inc.) following the manufacturer's instructions and the sequencing of a 9 nM library performed on HiSeq2000 with paired-end 100 nt reads also by LC Sciences (Texas).

Sequence Assembly

Thinning and trimming for the raw reads was done with CLC Genomics Workbench 5.1 (CLC bio, Aarhus, Denmark). Thinning refers to discarding of nucleotides and/or entire reads based on quality parameters. It was performed using either 0.05 or 0.005 as the limit, based on *Phred* quality scores, and resulting quality of the thinned reads was visualized in FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/, last accessed February 13, 2014). After thinning, only those terminal bases with a *Phred* quality score under 30 were trimmed (where a *Phred* score of 30 corresponds to a probability of 10^{-3} of incorrect base calling), producing sequences of unequal size. Reads were rescreened to check for presence of adapter or primer sequences using FastQC, and if present, adapters or primers were removed using with CLC Genomics Workbench 5.1.

De novo assemblies with all data sets thinned and/or trimmed were performed with CLC Genomics Workbench 5.1 (CLC bio, Aarhus, Denmark), or Trinity (http://trinityrnaseq.sourceforge.net/, last accessed February 13, 2014). Global alignments for the de novo assemblies were used with the following parameters: mismatch cost = 2; insertion cost = 3; deletion cost = 3; length fraction = 0.5; similarity = 0.8; and randomly assigning the nonspecific matches. Best *k*-mer length was estimated by the software.

Sequence Annotation

For each species, two methods were used for annotation: a more global assignment of gene ontology using Blast and a more specific assignment of domain by HMMer. For Blast, contigs shorter than 300 bp were removed, as very few of these short contigs retrieved results for Gene Ontology assignments (Riesgo, Andrade, et al. 2012). The remaining contigs were independently mapped against three different selections of the nonredundant (*nr*) NCBI database (all Metazoan proteins in *nr*, all Bacterial proteins in *nr*, and all Protozoan proteins in *nr*) using the BlastX program of the Blast suite. All searches were conducted with Blast+ (Altschul et al. 1997; Camacho et al. 2009) using an e-value cutoff of 1e−5. We used the output file from the blast against Metazoa which contained the best hits and Blast2GO v2.5.0 (Conesa et al. 2005) to retrieve the Gene Ontology (GO) terms and their parents associated with the top Blast hit for each sequence. Searches for specific genes were carried out using HMMer hidden Markov models using Interproscan tools (http://www.ebi.ac.uk/Tools/pfa/iprscan/, last accessed February 13, 2014) and either HMM profiles present in the PFAM Protein families database or HMMerbuilds generated specifically using sequences downloaded from NCBI.

To estimate the complexity of the complements of genes involved in different pathways, independently from the general Blast results, we selected gene targets from conserved developmental signaling pathways, and genes associated

with postsynaptic signaling, germ lineage and reproduction, adhesion, and innate immune regulation. We retrieved at least three different sequences of the selected protein targets from GenBank (http://www.ncbi.nlm.nih.gov/genbank/, last accessed February 13, 2014) from a range of metazoan groups to use them for searches in our transcriptome data sets. These sequences were aligned using T-COFFEE (Notredame et al. 2000), MAFFT (Katoh et al. 2005), or MUSCLE (Edgar 2004) depending on the level of conservation of the protein, and the alignments used to create HMM profiles for each protein of interest. HMMER searches were performed against all eight transcriptomes, translated into all six reading frames. We selected only the hits with the maximum similarity (cutoff of 1e−5; which varied greatly between groups), and checked each open reading frame with ORF finder (http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi, last accessed February 13, 2014). A similar approach was performed using the software CLC Genomics Workbench 5.0, selecting three protein sequences from other metazoans and blasting them using the local Blast suite plug-in with each contig list as the targeted database. Each predicted protein sequence was reverse blasted against the database *nr* in NCBI using the BlastP and DELTA-Blast programs (http://blast.ncbi.nlm.nih.gov/, last accessed February 13, 2014) and the domain structure checked with SMART (http://smart.embl-heidelberg.de/, last accessed February 13, 2014) using HMMER, PFAM domain, and internal repeats searching. To avoid the bias in the detection of genes derived from the use of cnidarian or bilaterian protein queries, each time we found a target gene in the transcriptomes, we added the sequence to the list of protein queries to improve the searches.

We also confirmed the presence/absence of the same set of conserved developmental signaling pathways, and neuronal signaling, germ lineage and reproduction, adhesion, and innate immune regulation in three different unicellular eukaryote species (*C. owczarzaki*, *M. brevicollis*, and *S. rosetta*) to determine whether the appearance of the genes in sponges was a novel acquisition or a feature shared with these unicellular organisms. We used the same gene targets and searched in the genomes with the BlastP engine implemented in the Broad institute website using default settings (http://www.broadinstitute.org/annotation/genome/multicellularity_project/MultiHome.html, last accessed February 13, 2014). We also used searched the *A. queenslandica* genome and *O. carmela* draft genome (Nichols et al. 2006; Srivastava et al. 2010; Nichols et al. 2012) to confirm the presence/absence of each genes at the genomic level. The presence of sequences from the cnidarian and bilaterian taxa were verified by either Blast or in the literature. We performed 3D reconstructions of the translated sequences of the targeted genes using PHYRE2 for protein fold recognition (Kelley and Sternberg 2009).

Sequences obtained in this study are available in supplementary file S1, Supplementary Material online, and also deposited under the Bioproject accession numbers in GenBank: *I. fasciculata* (PRJNA225586), *C. nucula* (PRJNA225590), *P. ficiformis* (PRJNA162901), *S. lacustris* (PRJNA225591), *P. suberitoides* (PRJNA225580), *S. coactum* (PRJNA162899), and *C. candelabrum* (PRJNA162903) and *A. vastus* at the ERA archive of the University of Alberta http://hdl.handle.net/10402/era.38025 (last accessed February 13, 2014).

Phylogenetic Analyses

For each of the selected genes, independent protein alignments were built using MUSCLE implemented in SEAVIEW 4.3.0 (Gouy et al. 2010) and MAFFT (Katoh et al. 2005) with default parameters. For the maximum likelihood phylogenetic analysis of the protein sequences, we used RAxML (Stamatakis 2006) with the LG model and an estimated gamma shape parameter and 500 independent searches. Nodal support was estimated via the rapid bootstrap algorithm (500 replicates) using the WAG-CAT model (Stamatakis et al. 2008). Bootstrap resampling frequencies were then mapped onto the optimal tree from the independent searches.

Supplementary Material

Supplementary figures S1 and S2, tables S1 and S2, and file S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank J. Wulff, S. López-Legentil, and J. Chu for images of *Pseudospongosorites*, *Ircinia*, and *Sycon*, and S. Taboada, J. Mora, and M. Maldonado for providing specimens of Mediterranean sponges (S.T. and J.M. provided samples of *P. ficiformis* and M.M. of *C. candelabrum*). W. Gallin and J. Paps gave invaluable advice on molecular analysis. This work was supported by a Marie Curie IOF grant #237219 and a Juan de la Cierva contract of the Spanish Ministry of Economy and Innovation to A.R.; Alberta Innovates Scholarship and NSERC PGSD scholarships to N.F. and P.W.; US National Science Foundation grant DEB 0844881 and by internal funds from the Museum of Comparative Zoology to G.G.; and a NSERC Discovery Grant 222863 to S.P.L.

References

Adams EDM, Goss GG, Leys SP. 2010. Freshwater sponges have functional, sealing epithelia with high transepithelial resistance and negative transepithelial potential. *PLoS One* 5:e15040.

Adamska M, Degnan SM, Green KM, Adamski M, Craigie A, Larroux C, Degnan BM. 2007. Wnt and TGF-β expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning. *PLoS One* 2:e1031.

Adamska M, Larroux C, Adamski M, Green K, Lovas E, Koop D, Richards GS, Zwafink C, Degnan BM. 2010. Structure and expression of conserved Wnt pathway components of the demosponge *Amphimedon queenslandica*. *Evol Dev.* 12:494–518.

Adamska M, Matus DQ, Adamski M, Green K, Rokhsar DS, Martindale MQ, Degnan BM. 2007. The evolutionary origin of hedgehog proteins. *Curr Biol.* 17:R836–R837.

Adell T, Nefkens I, Müller WEG. 2003. Polarity factor "Frizzled" in the demosponge *Suberites domuncula*: identification, expression and localization of the receptor in the epithelium/pinacoderm. *FEBS Lett.* 554:363–368.

Adell T, Thakur AN, Müller WEG. 2007. Isolation and characterization of Wnt pathway-related genes from Porifera. *Cell Biol Int.* 31:939–949.

Alié A, Manuel M. 2010. The backbone of the post-synaptic density originated in a unicellular ancestor of choanoflagellates and metazoans. *BMC Evol Biol.* 10:34.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389–3402.

Appeltans W, Ahyong ST, Anderson G, Angel MV, Artois T, Bailly N, Bamber R, Barber A, Bartsch I, Berta A, et al. 2012. The magnitude of global marine species diversity. *Curr Biol.* 22:2189–2202.

Armstrong PB, Quigley JP. 1999. α2-macroglobulin: an evolutionarily conserved arm of the innate immune system. *Dev Comp Immunol.* 23:375–390.

Bownes M. 1986. Expression of the genes coding for vitellogenin (yolk protein). *Ann Rev Entomol.* 31:507–531.

Buss LW. 1988. Diversification and germ-line determination. *Paleobiology* 14:313–321.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.

Conaco C, Neveu P, Xhou H, Arcila ML, Degnan SM, Degnan BM, Kosik KS. 2012. Transcriptome profiling of the demosponge *Amphimedon queenslandica* reveals genome-wide events that accompany major life cycle transitions. *BMC Genomics* 13:209.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674–3676.

Conway Morris S. 1993. The fossil record and the early evolution of the Metazoa. *Nature* 361:219–225.

Counihan RT, McNamara DC, Souter DC, Jebreen EJ, Preston NP, Johnson CR, Degnan BM. 2001. Pattern, synchrony and predictability of spawning of the tropical abalone *Haliotis asinina* form Heron Reef, Australia. *Mar Ecol Prog Ser.* 213:193–202.

Dickinson DJ, Nelson WJ, Weis WI. 2011. A polarized epithelium organized by β- and α-catenin predates cadherin and metazoan origins. *Science* 331:1336–1339.

Dickinson DJ, Nelson WJ, Wiens M. 2012. An epithelial tissue in *Dictyostelium* challenges the traditional origin of metazoan multicellularity. *BioEssays* 34:833–840.

Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD, et al. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452:745–749.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acid Res.* 32:1792–1797.

Elliott GRD, Leys SP. 2010. Evidence for glutamate, GABA and NO in coordinating behaviour in the sponge, *Ephydatia muelleri* (Demospongiae, Spongillidae). *J Exp Biol.* 213:2310–2321.

Ellwanger K, Nickel M. 2006. Neuroactive substances specifically modulate rhythmic body contractions in the nerveless metazoan *Tethya wilhelma* (Demospongiae, Porifera). *Front Zool.* 3:7.

Emson RH. 1966. The reactions of the sponge *Cliona celata* to applied stimuli. *Comp Biochem Physiol.* 18:805–827.

Ewen-Campen B, Schwager EE, Extavour CGM. 2010. The molecular machinery of germ line specification. *Mol Reprod Dev.* 77:3–18.

Extavour CGM, Akam M. 2003. Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* 130:5869–5884.

Fahey B, Degnan BM. 2010. Origin of animal epithelia: insights from the sponge genome. *Evol Dev.* 12:601–617.

Fairclough SR, Chen Z, Kramer E, Zeng Q, Young S, Robertson HM, Begovic E, Richter DJ, Russ C, Westbrook MJ, et al. 2013. Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate *Salpingoeca rosetta*. *Genome Biol.* 15:R15.

Funayama N, Nakatsukasa M, Mohri K, Agata K. 2010. *Piwi* expression in archeocytes and choanocytes in demosponges: insights into the stem cell system in demosponges. *Evol Dev.* 12:275–287.

Gazave E, Lapébie P, Richards GS, Brunet F, Ereskovsky AV, Degnan BM, Borchiellini C, Vervoort M, Renard E. 2009. Origin and evolution of the Notch signalling pathway: an overview from eukaryotic genomes. *BMC Evol Biol.* 9:249.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 27:221–224.

Guder C, Philipp I, Lengfeld T, Watanabe H, Hobmayer B, Holstein TW. 2006. The Wnt code: cnidarians signal the way. *Oncogene* 25: 7450–7460.

Harcet M, Roller M, Cetkovic H, Perina D, Wiens M, Müller WEG, Vlahovicek K. 2010. Demosponge EST sequencing reveals a complex genetic toolkit of the simplest metazoans. *Mol Biol Evol.* 27: 2747–2756.

Hardege JD, Bentley MG. 1997. Spawning synchrony in *Arenicola marina*: evidence for sex pheromonal control. *Proc Biol Sci.* 264: 1041–1047.

Harwood AJ. 2008. *Dictyostelium* development: a prototypic Wnt pathway?. In: Vincan E, editor. Wnt signalling. Vol. II: Pathway models. New York: Springer/Business Media. p. 21–32.

Hejnol A, Obst M, Stamatakis A, Ott M, Rouse GW, Edgecombe GD, Martinez P, Baguñà J, Bailly X, Jondelius U, et al. 2009. Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc Biol Sci.* 276:4261–4270.

Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH, Terwilliger DP, Buckley KM, Brockton V, Nair SV, Berney K, et al. 2006. The immune gene repertoire encoded in the purple sea urchin genome. *Dev Biol.* 300:349–365.

Hoppe WF, Reichert MJM. 1987. Predictable annual mass release of gametes by the coral reef sponge *Neofibularia nolitangere* (Porifera: Demospongiae). *Mar Biol.* 94:277–285.

Hughes AL. 1997. Rapid evolution of immunoglobulin superfamily C2 domains expressed in immune system cells. *Mol Biol Evol.* 14: 1–5.

Ingham PW, Nakano Y, Seger C. 2011. Mechanisms and functions of Hedgehog signalling across the Metazoa. *Nat Rev Genet.* 12: 393–406.

Iwaki D, Kanno K, Takahashi M, Endo Y, Matsushita M, Fujita T. 2011. The role of Mannose-binding-lectin-associated Serine protease-3 in activation of the alternative complement pathway. *J Immunol.* 187: 3751–3758.

Jackson DJ, Macis L, Reitner J, Degnan BM, Wörheide G. 2007. Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis. *Science* 316:1893–1895.

Juliano C, Wessel G. 2010. Versatile germline genes: when are germline cells segregated during animal development? *Science* 329: 640–641.

Katoh K, Kuma K-I, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33:511–518.

Kelley LA, Sternberg MJE. 2009. Protein structure prediction on the web: a case study using the Phyre server. *Nat Protoc.* 4:363–371.

King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I, et al. 2008. The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature* 451:783–788.

Kopp A. 2012. *Dmrt* genes in the development and evolution of sexual dimorphism. *Trends Genet.* 28:175–184.

Labouesse M, Georges-Labouesse E. 2003. Cell adhesion: parallels between vertebrate and invertebrate focal adhesions. *Curr Biol.* 13: R528–R530.

Lapébie P, Gazave E, Ereskovsky A, Derelle R, Bézac C, Renard E, Houliston E, Borchiellini C TL. 2009. WNT/β-catenin signalling and epithelial patterning in the homoscleromorph sponge. *Oscarella. PLoS One* 4:e5823.

Lentz TL. 1966. Histochemical localization of neurohumors in a sponges. *J Exp Zool.* 162:171–179.

Leys SP, Riesgo A. 2012. Epithelia, an evolutionary novelty of metazoans. *J Exp Zool B Mol Dev Evol.* 318:438–447.

Li J, Wang S, Huang S, Cheng D, Shen S, Xiong C. 2009. *Attractin* gene deficiency contributes to testis vacuolization and sperm dysfunction in male mice. *J Huazhong Univ Sci Technol.* 29:750–754.
Lin H. 1997. The tao of stem cells in the germline. *Ann Rev Genet.* 31: 455–491.
Matus DQ, Magie CR, Pang K, Martindale MQ, Thomsen GH. 2008. The Hedgehog gene family of the cnidarian *Nematostella vectensis* and implications for understanding metazoan Hedgehog pathway evolution. *Dev Biol.* 313:501–518.
Medina M, Collins AG, Silberman JD, Sogin ML. 2001. Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. *Proc Nat Acad Sci U S A.* 98: 9707–9712.
Mckeown M, Madigan SJ. 1992. Sex determination and differentiation in invertebrates: *Drosophila* and *Caenorhabditis elegans*. *Curr Opin Cell Biol.* 4:948–954.
McShea DW. 1996. Perspective: Metazoan complexity and evolution: is there a trend? *Evolution* 50:477–492.
Miller DJ, Hemmrich G, Ball EE, Hayward DC, Khalturin K, Funayama N, Agata K, Bosch TCG. 2007. The innate immune repertoire in Cnidaria—ancestral complexity and stochastic gene loss. *Genome Biol.* 8:R59.
Miller SW, Hayward DC, Bunch TA, Miller DJ, Ball EE, Bardwell VJ, Zarkower D, Brower DL. 2003. A DM domain protein from a coral, *Acropora millepora*, homologous to proteins important for sex determination. *Evol Dev.* 5:251–258.
Mochizuki K, Nishimiya-Fujisawa C, Fujisawa T. 2001. Universal occurrence of the *vasa*-related genes among metazoans and their germline expression in *Hydra*. *Dev Genes Evol.* 211: 299–308.
Mochizuki K, Sano H, Kobayashi S, Nishimiya-Fujisawa C, Fujisawa T. 2000. Expression and evolutionary conservation of *nanos*-related genes in *Hydra*. *Dev Genes Evol.* 210:591–602.
Nichols SA, Dirks W, Pearse JS, King N. 2006. Early evolution of animal cell signalling and adhesion genes. *Proc Nat Acad Sci U S A.* 103: 12451–12456.
Nichols SA, Roberts BW, Richter DJ, Fairclough SR, King N. 2012. Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/β-catenin complex. *Proc Nat Acad Sci U S A.* 109: 13046–13051.
Nielsen C. 2012. Animal evolution: interrelationships of the living phyla. Oxford: Oxford University Press.
Nosenko T, Schreiber F, Adamska M, Adamski M, Eitel M, Hammel J, Maldonado M, Müller WE, Nickel M, Schierwater B, et al. 2013. Deep metazoan phylogeny: when different genes tell different stories. *Mol Phylogenet Evol.* 67:223–233.
Notredame C, Higgins DG, Heringa H. 2000. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol.* 302: 205–217.
Novo M, Riesgo A, Fernández-Guerra A, Giribet G. 2013. Pheromone evolution, reproductive genes, and comparative transcriptomics in Mediterranean earthworms (Annelida, Oligochaeta, Hormogastridae). *Mol Biol Evol.* 30:1614–1629.
Painter SD, Clough B, Garden RW, Sweedler JV, Nagle GT. 1998. Characterization of *Aplysia* Attractin, the first water-borne peptide pheromone in invertebrates. *Biol Bull.* 194:120–131.
Pérez-Porro AR, Navarro-Gómez D, Uriz MJ, Giribet G. 2013. A NGS approach to the encrusting Mediterranean sponge *Crella elegans* (Porifera, Demospongiae, Poecilosclerida): transcriptome sequencing, characterization and overview of the gene expression along three life cycle stages. *Mol Ecol Res.* 13:494–509.
Philippe H, Brinkmann H, Lavrov DV, Littlewood DTJ, Manuel M, Wörheide G, Baurain D. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biol.* 9: e1000602.
Pick KS, Philippe H, Schreiber F, Erpenbeck D, Jackson DJ, Wrede P, Wiens M, Alié A, Morgenstern B, Manuel M, et al. 2010. Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. *Mol Biol Evol.* 27:1983–1987.
Raz E. 2000. The function and regulation of *vasa*-like genes in germ-cell development. *Genome Biol.* 1:1017.1011–1017.1016.
Richards GS, Degnan BM. 2009. The dawn of developmental signalling in the Metazoa. *Cold Spring Harbor Symp Quant Biol.* 74:81–90.
Richards GS, Degnan BM. 2012. The expression of Delta ligands in the sponge *Amphimedon queenslandica* suggests an ancient role for Notch signalling in metazoan development. *Evo Devo.* 3:15.
Richards GS, Simionato E, Peron M, Adamska M, Vervoort M, Degnan BM. 2008. Sponge genes provide new insight into the evolutionary origin of the neurogenic circuit. *Curr Biol.* 18:1156–1161.
Riesgo A, Andrade SCS, Sharma PP, Novo M, Pérez-Porro AR, Vahtera V, González VL, Kawauchi GY, Giribet G. 2012. Comparative description of ten transcriptomes of newly sequenced invertebrates and efficiency estimation of genomic sampling in non-model taxa. *Frontiers Zool.* 9:33.
Riesgo A, Maldonado M. 2008. Differences in reproductive timing among sponges sharing habitat and thermal regime. *Invertebr Biol.* 127:357–367.
Riesgo A, Maldonado M. 2009. Sexual reproduction of demosponges. Berlin: Verlag.
Riesgo A, Maldonado M, Durfort M. 2007. Dynamics of gametogenesis, embryogenesis, and larval release in a Mediterranean homosclerophorid demosponge. *Mar Freshw Res.* 58:398–417.
Riesgo A, Pérez-Porro AR, Carmona S, Leys SP, Giribet G. 2012. Optimization of preservation and storage time of sponge tissues to obtain quality mRNA for next-generation sequencing. *Mol Ecol Res.* 12:312–322.
Ryan JF, Pang K, Program NCS, Mullikin JC, Martindale MQ, Baxevanis AD. 2010. The homeodomain complement of the ctenophore *Mnemiopsis leidyi* suggests that Ctenophora and Porifera diverged prior to ParaHoxozoa. *Evo Devo.* 1:9.
Ryan JF, Pang K, Schnitzler CE, Nguyen A-D, Moreland RT, Simmons DK, Koch BJ, Francis WR, Havlak P, Comparative Sequencing Program NISC, et al. 2013. The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science* 342: 1242592.
Ryan TJ, Grant SGN. 2009. The origin and evolution of synapses. *Nat Rev Neurosci.* 10:701–712.
Saffman EE, Lasko P. 1999. Germline development in vertebrates and invertebrates. *Cell Mol Life Sci.* 55:1141–1163.
Sakarya O, Armstrong KA, Adamska M, Adamski M, Wang I-F, Tidor B, Degnan BM, Kosik KS. 2007. A post-synaptic scaffold at the origin of the animal kingdom. *PLoS One* 2:e506.
Salmoun M, Devijer C, Daloze D, Braekman J-C, van Soest RWM. 2002. 5-Hydroxytryptamine-derived alkaloids from two marine sponges of the genus *Hyrtiois*. *J Nat Prod.* 65:1173–1176.
Sebé-Pedrós A, Roger AJ, Lang FB, King N, Ruiz-Trillo I. 2010. Ancient origin of the integrin-mediated adhesion and signaling machinery. *Proc Nat Acad Sci U S A.* 107:10142–10147.
Simpson TL. 1984. The cell biology of sponges. Berlin: Springer Verlag.
Song X, Ping J, Sheng Q, Liming C, Fei M. 2012. The evolution and origin of animal Toll-like receptor signaling pathway revealed by network-level molecular evolutionary analyses. *PLoS One* 7: e51657.
Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U, Kawashima T, Kuo A, Mitros T, Salamov A, Carpenter ML, et al. 2008. The *Trichoplax* genome and the nature of placozoans. *Nature* 454: 955–960.
Srivastava M, Simakov O, Chapman J, Fahey B, Gauthier MEA, Mitros T, Richards GS, Conaco C, Dacre M, Hellsten U, et al. 2010. The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature* 466:720–726.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol.* 57:758–771.
Tucker RP. 2013. Horizontal gene transfer in choanoflagellates. *J Exp Zool B Mol Dev Evol.* 320:1–9.

Turner CE. 2000. Paxillin and focal adhesion signalling. *Nat Cell Biol.* 2: E231–E236.
Vacalet J, Boury-Esnault N. 1995. Carnivorous sponges. *Nature* 373: 26.
Weyrer S, Rützler K, Rieger R. 1999. Serotonin in Porifera? From developing *Tedania ignis*, the Caribbean fire sponge (Demospongiae). *Mem Queensland Mus.* 44:659–665.
Wiley HS, Wallace RA. 1981. The structure of Vitellogenin. *J Biol Chem.* 256:8626–8634.
Windsor PJ, Leys SP. 2010. Wnt signalling and induction in the sponges aquiferous system: evidence for an ancient origin of the organizer. *Evol Dev.* 12:484–493.
Wyatt TD. 2003. Pheromones and animal behaviour: communication by smell and taste. Cambridge: Cambridge University Press.
Wylie C. 1999. Germ cells. *Cell* 96:165–174.
Zrzavý J, Mihulka S, Kepka P, Bezděk A, Tietz D. 1998. Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. *Cladistics* 14:249–285.