Phylogenomics Revives Traditional Views on Deep Animal Relationships

Hervé Philippe,1,11 Romain Derelle,2,11 Philippe Lopez,2 Kerstin Pick,3,5 Carole Borchiellini,6 Nicole Boury-Esnault,6 Jean Vacel,6 Emmanuelle Renard,6 Evelyn Houlist,7 Eric Quénèc,5 Corinne Da Silva,6 Patrick Wincker,6 Hervé Le Guyader,2 Sally Leys,3 Daniel J. Jackson,3,10 Fabian Schreiber,4 Dirk Erpenbeck,5 Burkhard Morgenstern,3,4 Gert Wörheide,5,6 and Michaël Manuel2,*

1Centre Robert-Cedergren
Département de Biochimie
Université de Montréal
Succursale Centre-Ville
Montréal, Québec H3C3J7
Canada
2UPMC, Univ Paris 06
UMR 7138 Systématique, Adaptation, Evolution
CNRS UPMC MNHN IRD, Case 05
Université Pierre et Marie Curie
7 quai St Bernard
75005 Paris
France
3Courant Research Center Geobiology
Georg-August-Universität Göttingen Goldschmidtstr. 3
4Abteilung Bioinformatik
Institut für Mikrobiologie und Genetik
Goldschmidtstr. 1
37077 Göttingen
Germany
5Department of Earth- and Environmental Sciences &
GeoBioCenterLMU
Ludwig-Maximilians-Universität München
Richard-Wagner-Str. 10
80333 München
Germany
6Aix-Marseille Université
CNRS UMR 8540 DIMAR
Centre d’Océanologie de Marseille
Station Marine d’Endoume
rue de la Batterie des Lions
13 007 Marseille
France
7UPMC, Univ Paris 06
CNRS UMR 7009 Biologie du Développement
Observatoire Océanologique
06230 Villefranche-sur-Mer
France
8Genoscope and CNRS UMR 8030
2 rue Gaston Crémieux
91057 Evry
France
9Department of Biological Sciences
CW 405
University of Alberta

Canada
10School of Integrative Biology
The University of Queensland
Brisbane 4072
Australia

Summary

The origin of many of the defining features of animal body plans, such as symmetry, nervous system, and the mesoderm, remains shrouded in mystery because of major uncertainty regarding the emergence order of the early branching taxa: the sponge groups, ctenophores, placozoans, cnidarians, and bilaterians. The “phylogenomic” approach [1] has recently provided a robust picture for intrabilaterian relationships [2, 3] but not yet for more early branching metazoan clades. We have assembled a comprehensive 128 gene data set including newly generated sequence data from ctenophores, cnidarians, and all four main sponge groups. The resulting phylogeny yields two significant conclusions reviving old views that have been challenged in the molecular era: (1) that the sponges (Porifera) are monophyletic and not paraphyletic as repeatedly proposed [4–9], thus undermining the idea that ancestral metazoans had a sponge-like body plan; (2) that the most likely position for the ctenophores is together with the cnidarians in a “coelenterate” clade. The Porifera and the Placozoa branch basally with respect to a moderately supported “eumetazoan” clade containing the three taxa with nervous system function. This new phylogeny provides a stimulating framework for exploring the important changes that shaped the body plans of the early diverging phyla.

Results and Discussion

A Comprehensive Phylogenomic Data Set to Address Basal Metazoan Evolution

Previous studies of basal metazoan relationships by molecular phylogeny techniques (e.g., [3–8, 10, 11]) have proposed contradictory and often poorly supported trees, leaving major issues such as the phylogenetic status (monophyly or paraphyly) of sponges and the position of ctenophores and placozoans unsettled. These inconsistencies may reflect insufficient molecular sampling and/or inadequate taxon sampling of the diversity of extant nonbilaterian metazoan lineages [1, 11–13]. We have adopted a phylogenomic approach specifically aimed at clarifying the basal metazoan relationships, involving more comprehensive sampling of all the major early branching animal lineages. By using newly generated cDNA sequences in addition to publicly available sequences, we have assembled a metazoan data set enriched in species representing the early diverging phyla (see Experimental Procedures and Supplementary Data available online). The data set comprises 128 different protein-coding genes (30,257 unambiguously aligned
positions) for 11 outgroup species and 44 metazoans, including 9 sponge species, 3 ctenophores, 9 cnidarians, the placozoan *Trichoplax*, and a representative sampling of bilaterian species. Among the 55 terminal taxa, 24 are complete or nearly complete (≤5% of missing data), and only 27% of positions in the final alignment are absent (see Table S2).

This is the first phylogenomic data set to include all four main sponge lineages: Demospongiae, by far the most species-rich sponge group, is represented by four species, chosen to maximize morphological and phylogenetic diversity; Hexactinellida and Calcispongia are each represented by two species; and Homoscleromorpha is represented by a chimerical operational taxonomic unit created from two species of the genus *Oscarella*.

The Sponges Restored as a Monophyletic Group
Our data set was analyzed by Bayesian inference analysis, via the CAT model of sequence evolution [14], conceived to reduce artifacts resulting from mutational saturation and unequal rates of substitution, which are major problems when analyzing ancient events [13, 15]. To explore the effect of outgroup taxa on the metazoan interrelationships obtained, we performed three analyses with different taxon samplings (Figure 1): rooted analysis with a paraphyletic outgroup comprised of the fungi, ichthyosporeans, *Capsaspora*, and choanoflagellates ("outgroup 1;" tree shown in Figure 1 and Figure S1; with bootstrap supports [BS] in pink in Figure 1); analysis rooted with just choanoflagellates, the metazoan sister group [16] ("outgroup 2;" BS in blue in Figure 1, tree in...
Figure 2. Characters of the Sponge Body Plan and Their Evolution
(A) Schematic section of an adult sponge (bottom) and SEM picture showing a choanocyte, the sponge collar cell (top, choanocyte from Chelonaplysilla noevus, Demospongiae). The arrows indicate the direction of circulation of water in the aquiferous system of the sponge. Abbreviations: atr, atrial cavity; cb, cell body; cc, choanocyte chamber; col, collar of microvilli; ex, exhalant canal; fl, flagellum; in, inhalant canal; mes, mesohyl; osc, osculum (or exhalant orifice); ost, ostium (or inhalant orifice); pin, pinacoderm (thin epithelial layer, limiting the sponge body on its external surface and within the canals); sp, spicule.
(B) Most parsimonious scenario for the evolution of sponge body plan characters, imposed on a scheme of sponge paraphyly.
(C) Most parsimonious scenario assuming sponge monophyly.
In (B) and (C), the gray branches indicate the presence of sponge body plan characters (aquiferous system, internalized choanocyte chambers, pinacoderm) and the black branches indicate the absence of these characters. The gray horizontal line indicates character acquisition; the hollow horizontal line indicates character loss. “Sponges 1, 2, and 3” correspond to the major lineages (silicisponges, homoscleromorphs, and calcisponges), of which exact branching order varies among published studies recovering sponge paraphyly.

Figure S2; and unrooted analysis (BS in black in Figure 1, tree in Figure S3). The topology resulting from the rooted analyses (trees shown in Figure 1 and Figures S1 and S2) was statistically well supported at most nodes, and its general features were in line with previous studies [2, 3]: choanoflagellates positioned as the sister group to the Metazoa, with Bilateria, Protostomia, Lophotrochozoa, and Ecdysozoa each forming well-supported monophyletic groups. These rooted trees provide strong evidence that the sponge species all belong together in a monophyletic group (Porifera) (bootstrap support = 90% and 96% with outgroup 1 and outgroup 2, respectively). The branch leading to the Porifera is short (Figure 1), accounting for the difficulty in recovering sponge monophyly in previous molecular analyses. This presumably reflects closely spaced splitting events during the Proterozoic era when the sponge lineages emerged.

Extant sponges are a diverse group sharing a number of common body plan features, notably a system of internal canals and choanocyte chambers through which water flows, and a thin epithelial covering called the pinacoderm (Figure 2A). Although morphological character analyses firmly support the hypothesis that the sponges form a monophyletic group [5, 17], rRNA analyses have repeatedly indicated that they are paraphyletic, with the calcisponges and/or the homoscleromorphs positioned closer to eumetazoans than to the other sponges [4–8]. It is worth noting, however, that sponge monophyly could not be ruled out unequivocally in many of these studies because of poor statistical support [6, 7, 10]. The previously proposed hypothesis of sponge paraphyly had significant implications for understanding the origin of multicellular animals, because it would imply that characters shared by all sponge lineages are ancestral for the Metazoa and that eumetazoans are derived from animals with a sponge-like body plan [4, 5, 8, 9] (Figure 2B).

The significant support for sponge monophyly in the present study allows us to return to the idea that a sponge body plan (notably featuring an aquiferous system with internalized choanocyte chambers and the pinacoderm) evolved in the stem line of the Porifera (Figures 2C and 3). The specialized collar apparatus of sponge choanocytes has often been assumed to be an ancient feature shared with choanoflagellates, based on phenotypic similarity [16]. However, many ultrastructural details of choanoflagellate and choanocyte cells are different, such as the length and spacing of the microvilli and the organization of the microtubule cytoskeleton. Their functional properties also differ, with the microvilli of choanoflagellates but not of choanocytes being contractile. Their similarity might thus represent convergence, with choanocytes being a synapomorphy (shared derived character) of Porifera. It is clear in any case that, rather than reflecting the ancestral animal form, adult sponges are better considered as highly specialized organisms, possibly having acquired a sedentary life style from a hypothetical pelagic ancestor. Notably, the absence of obvious symmetry in many adult sponges fuelled the popular idea that the last common metazoan ancestor lacked defined axial organization [18, 19]. In fact the adult bodies of hexactinellids, calcisponges, homoscleromorphs, and nonbilaterian eumetazoans are characterized by axial symmetry, as is the larval organization of sponges [20], ctenophores, and cnidarians. This suggests that the common ancestor of all animals may have showed symmetry around a single polarity axis [21], and thus that the asymmetry of the adult body in most demosponges and in Trichoplax is likely to be derived rather than ancestral (Figure 3).

Lessons from Relationships within the Porifera
In line with some previously published phylogenies (e.g., [6, 7, 11]), our analysis placed hexactinellids and demosponges together to form the Silicea Gray, 1867 [22] sensu stricto (with maximal bootstrap support in all analyses) characterized by siliceous spicules organized around a well-defined protein axial filament [23] and by a particular class of membrane phospholipids known as demospongic acids [24]. Concerning the enigmatic Homoscleromorpha, our analyses clearly excluded...
them from the demosponges and favored a sister group relationship to the Calcispongiae (with highest support of 91% [BS] obtained in the analysis with outgroup 2), in line with results from 18S rRNA analyses [25, 26] but in conflict with traditional classification schemes (see [27]). The siliceous spicules without defined axial filament found in some Homoscleromorpha [23] thus might have evolved independently from those of hexactinellids and demosponges (Figure 3). In addition, homology of siliceous spicules between the latter two taxa is uncertain because they are absent in the Dictyoceratida, represented here by Carteriospongia foliascens, the earliest-branching Demospongiae taxon in our phylogeny (Figure 1) (see [25, 28]). Whether the thick basi-epithelial basement membrane of homoscleromorph larvae and adults, which shares homologous biochemical components with eumetazoan basement membranes [29, 30], was inherited from a common metazoan ancestor and subsequently reduced or lost in most sponges and in Trichoplax, or acquired independently in homoscleromorphs and eumetazoans, cannot be decided from our analyses (Figure 3).

The Coelenterata Clade Revived
A recent phylogenomic analysis suggested that the cnidarians, a phylum of marine, mostly planktonic and gelatinous animals, diverged earlier than sponges [3]. This highly unorthodox hypothesis would see the dismantling of the clade Eumetazoa (ctenophores, cnidarians, and bilaterians), despite their sharing of many key characteristics such as nerve and muscle cells and a differentiated digestive system (absent in sponges and in Trichoplax). Polyphyly of eumetazoans would thus imply several independent acquisitions of these features, or their secondary loss in sponges and/or placozoans [31]. Our rooted analyses are not consistent with the basal position of ctenophores, but rather suggest the existence of a Coelenterata [32] (Ctenophora + Cnidaria) clade, placed within a monophyletic Eumetazoa (Figure 1). A recent study [11] also obtained the coelenterate grouping, but with low bootstrap support, and within a heterodox scheme of eumetazoan polyphyly. Historically, the coelenterate grouping [32] was based on certain anatomical resemblances between ctenophores and the cnidarian medusae (e.g., gelatinous body, tentacles, and “radial” symmetry) that were later considered convergences [33]. In fact, the complex body plan of ctenophores (with eight longitudinal rows of ciliated “comb rows,” a ramified endodermal gastro-vascular system, a complex sensory apparatus located at the aboral pole, and a prevalence of biradial symmetry [19]) differs markedly from that of the cnidarians. Apart from some common embryological features (central yolk and similar unipolar cleavages; animal pole corresponding to adult mouth), there are no clear-cut morpho-anatomical synapomorphies supporting the Coelenterata. The very long branch leading to the ctenophores (see Figure 1) makes their position prone to perturbation by the long-branch attraction (LBA) artifact [34]. The basal position of ctenophores suggested by Dunn et al. [3] might thus have resulted from attraction of the ctenophores by the distant outgroup taxa used to root the tree. This problem was alleviated in the present study by more comprehensive species sampling.
and by the use of the CAT model. That ctenophores are indeed attracted by distant outgroups is empirically demonstrated in our analyses by the observed increase in branch support for Coelenterata and Eumetazoa after partial or total removal of outgroup taxa (Figure 1). Thus, when distant outgroups (notably fungi) were used (as in [3]) (outgroup 1), the Coelenterata were moderately supported (BS = 78%) and the Eumetazoa were poorly supported (BS = 62%). With choanoflagellates as the only outgroup (outgroup 2), support for Coelenterata and Eumetazoa increased remarkably (BS = 93% and BS = 90%, respectively). Even higher support for the coelenterates was obtained by unrooted analysis (BS = 99%). We further checked that the position of ctenophores was not due to artificial attraction by the long branch leading to medusozooan cnidarians (Hydrozoa + Scyphozoza) (see Figure 1), by an analysis excluding these species (Figure S4): ctenophores still grouped with anthozoan cnidarians (a short branch), with high support (BS = 91%).

Our results not only suggest that ctenophores are the sister group to cnidarians but also that eumetazoans are monophyletic, implying single acquisition during animal evolution of nerve and muscle cells and/or the digestive system, in line with conventional ideas. These findings are at odds with the schemes of eumetazoan polyphyly proposed in two other recent phylogenetic studies [3, 11], both of which used more limited taxonomic sampling of nonbilateral metazoans and more phylogenetically distant outgroups. It is clearly premature to make a final conclusion on basal metazoan relationships, because not all our analyses yielded significant statistical support values, and the influence of outgroup taxonomy on tree topology might indicate that there is conflict in the data. As additional data from more nonbilateral species becomes available, the remaining doubts should finally be resolved. It should be noted that the position of the placozoan Trichoplax with respect to sponges and eumetazoans remains poorly supported in our analyses (Figure 1) and that recent investigations focused on placozoan relationships [11, 35] provided contradictory results, leaving this question unresolved.

Body Plan Evolution among the Eumetazoans

The proposed restoration of the Coelenterata implies that cnidarians and ctenophores are phylogenetically equally related to the bilaterians and has implications with respect to the origin of mesoderm and of bilateral symmetry. These body plan features have been classically thought to be evolutionary novelties of the Bilateria, but their origin has been suggested to date back to the common cnidarian-bilaterian ancestor from recent developmental gene evidence [36–38]. The mesoderm-like muscle cell lineage of ctenophores [37] might be homologous with the mesoderm of the Bilateria and with mesoderm-like derivatives previously identified in cnidarians [37, 39]. Concerning symmetry, parsimony optimization favors an independent evolution of anatomical bilaterality in the bilaterians and in anthozoan cnidarians (Figure 3), but the significance of the biradial anatomy of the ctenophores [21] remains to be evaluated, for instance through the study of the developmental regulatory genes unilaterally expressed in cnidarians and in the bilaterians [38].

Our new proposal of basal metazoan relationships provides a stimulating framework for furthering our understanding of early metazoan evolution. It suggests that several key features of metazoan body plans were affected by events of convergence or reversion (Figure 3), contrasting with the traditional conception of metazoan evolution dominated by a gradual increase in morphological complexity. It should motivate detailed exploration of many aspects of character transformations during evolution, development, and metamorphosis, as well as the relationships of larval to adult traits.

Experimental Procedures

EST Sequencing

Fresh samples of Sycon raphanus, Oscarella lobularis, and Oopscas minuta were collected in the Mediterranean near Marseille (France). Ephydatia muelleri gemmules from Belgium were incubated in the lab until production of young adult sponges. Samples of Heterochone calyx were collected in British Columbia (Canada) and re-aggregated tissue was used as starting material. Carteriopsis fongalia was collected at Lizard Island (Great Barrier Reef, Australia) and Leucetta chagosensis at North Stradbroke Island (Australia). Pleurobrachia pileus adults were collected in Villefranche-sur-Mer (France). For Clybia hemisphaerica, the starting material was a strain cultured at the Marine Station in Villefranche-sur-Mer. Frozen samples, RNA Later (QIAGEN)-preserved, or extracted total RNA (depending on the species) were sent to Genome Express (O. minuta), RZPD (S. raphanus, O. lobularis, E. muelleri), Express Genomics (P. pileus and C. hemisphaerica), and the Max Planck Institute for Molecular Genetics in Berlin (Germany) (H. calyx, C. fongalia, L. chagosensis) for cDNA library construction. ESTs were sequenced at the Max Planck Institute for Molecular Genetics (Berlin, Germany) (H. calyx, C. fongalia, L. chagosensis, C. hemisphaerica) or at the Genoscope (Evry, France) (all other species). Numbers of sequenced ESTs were approximately 2,000 (O. minuta, E. muelleri, S. raphanus, O. lobularis), 4,000 (H. calyx, C. fongalia, L. chagosensis), 30,000 (P. pileus), and 90,000 (C. hemisphaerica). All these newly sequenced EST collections are publicly available in dbEST/GenBank (http://www.ncbi.nlm.nih.gov/dbEST/). The alignment used for phylogenetic analyses is provided as Supplemental Data.

Data Assembly

We built upon phylogenomic data sets previously assembled [13, 40]. These alignments were updated, via the protocol described in [41], with the addition of newly generated sequences, and of sequences publicly available from the Trace Archive (http://www.ncbi.nlm.nih.gov/Traces/) and the EST Database (http://www.ncbi.nlm.nih.gov/dbEST/) of GenBank at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). In addition, 23 new genes sampled for at least two main poriferan clades were added. All these genes are likely to be orthologs because they are in single-copy in most of the opisthokonts, few recent duplications being observed mostly in vertebrates and Drosophila. To further evaluate the possibility of hidden paralogy, we inferred single-gene phylogenies and looked for any strongly supported conflict with the super-matrix tree according to protocol described in [42]. At a bootstrap threshold of 70%, conflicts were observed for only 6.5% of the testable bipartitions, less than the expected error rate. None of these conflicts could be easily explained by hidden paralogy (see details on these analyses in the Supplemental Experimental Procedures).

As previously demonstrated [13], taxon sampling has a major impact in phylogenomic studies. In addition to the nine sponges, nine cnidarians, three ctenophores, and one placozoan available, we therefore selected 22 slowly evolving representative taxa among available Bilateria (based on previous branch length comparison). To reduce the potential impact of long-branch attraction (LBA) [34], we also incorporated all available ichthyospores and choanoflagellates (taxa hypothesized to be the closest unicellular relatives of Metazoa) to break the long-branch leading to the distantly related fungal outgroup (for which only the slow-evolving chytridiomycetes and zygomycetes were used). Ambiguously aligned regions were removed with Gblocks [43]. Sequence selection and concatenation were performed with SCaFoS [44]. To reduce the amount of missing data in the final alignment, we discarded under-sampled genes. Only genes sampled for at least two-thirds of the species (36 out of 55) were retained. The resulting gene selection (128 genes) yielded an alignment of 30,257 unambiguously aligned positions. For all but two genes, the four major diplolast lineages (Porifera, Cnidaria, Ctenophora, and Placozoa) were represented by at least one species; at least three of the main poriferan clades (Demospongiae, Hexactinellida, Homoscleromorpha, Calcispongia) were represented for 65% of the genes.
Phylogenetic Analyses

To analyze our phylogenomic data set, we used the CAT model [14] with PhyloBayes version 2.3, which has been shown in some contexts to be less prone to LBA artifacts than other models [13, 15]. We performed statistical comparisons of the CAT model with the GTR model (of which other matrix-based models—WAG, JTT, or LG—are special cases) by using cross-validation tests as described in [40], based on the topology of Figure 1. Ten replicates were run: 9/10 for the learning set and 1/10 for the test set. MCMC were run for 3000 (1500) cycles for the CAT (GTR) model, 1500 (100) being discarded as burn-in. The CAT model was found to have a much better statistical fit than did GTR (a likelihood score of 3033.36 ± 123.824 in favor of CAT). For the plain posterior estimation, two independent chains were run for a total number of 15,000 cycles (corresponding to ~1,200,000 generations) saving every ten cycles and discarding the first 5,000 cycles (burn-in). Each cycle consists in a complicated series of updates of all components of the parameter vector, including an average of 20 topological updates. The posterior consensus tree was obtained by pooling the tree lists of four independent runs. For each node, we compared the posterior probabilities inferred from two independent chains. The maximum difference we observed was 0.11 and 0.2 for the trees with 55 and 44 species, respectively; the average difference being 0.0017 and 0.0011. The differences corresponded to nodes poorly supported by bootstrap values. Posterior probabilities were always identical for nodes having a bootstrap support >90%. Bootstrap percentages were obtained by running 100 independent pseudoreplicates, for 10,000 cycles (900,000 generations) each (because of CPU limitations). Bibliographic references for phylogeny softwares are provided in the Supplemental Experimental Procedures.

Supplemental Data

Supplemental Data include Supplemental Experimental Procedures, four figures, three tables, and the alignment and can be found with this article online at http://www.current-biology.com/supplemental/S0960-9822(09)00805-7.

Acknowledgments

We thank Muriel Jager for technical help and Ronald Jenner for advice and discussion. We are grateful to Evelyn Richelle-Maurer for providing the Ephydatia gemmules, to Thierry Pérez for collecting Oopscacas minuta and Oscarella lobularis specimens, and to Jean-Yves Carval for collecting Pleurobrachia pileus specimens. H.P. acknowledges Génome Québec, the Canadian Research Chair, and the Université de Montréal for financial support and the Réseau Québécois de Calcul de Haute Performance for computational resources. M.M. acknowledges the French Ministry of Research ("ACI jeunes chercheurs" and ANR NT_NV_52 Genocnidaire), the Consortium National de Recherche en Génomique, the Genoscope, and the GIS Institut de la Génomique Marine for financial support. S.L. acknowledges funding from NSERC (for the collection and processing of Heterochone). G.W. and B.M. acknowledge financial support through the Canadian Research Chair, and the Université de Montréal for critical resources that supported cdNA library construction of Leucetta chagosensis.

Received: November 13, 2008
Revised: February 22, 2009
Accepted: February 25, 2009
Published online: April 2, 2009

References