The Genome of the Ctenophore *Mnemiopsis leidyi* and Its Implications for Cell Type Evolution


**Introduction:** An understanding of ctenophore biology is critical for reconstructing events that occurred early in animal evolution. The phylogenetic relationship of ctenophores (comb jellies) to other animals has been a source of long-standing debate. Until recently, it was thought that Porifera (sponges) was the earliest diverging animal lineage, but recent reports have instead suggested Ctenophora as the earliest diverging animal lineage. Because ctenophores share some of the same complex cell types with bilaterians (such as neural and mesodermal cells), the phylogenetic position of ctenophores affects how we think about the early evolution of these cell types.

**Methods:** We have sequenced, annotated, and analyzed the 150-megabase genome of the ctenophore *Mnemiopsis leidyi*. We have performed detailed phylogenetic analyses on these new data using both sequence matrices and information on gene content. We conducted extensive genomic inventories on signaling pathway components and genes known to be critical to neural and mesodermal cell types, among others.

**Results:** Our phylogenetic analyses suggest that ctenophores are the sister group to the rest of the extant animals. We find that the sets of neural components present in the genomes of *Mnemiopsis* and the sponge *Amphimedon queenslandica* are quite similar, suggesting that sponges have the necessary genetic machinery for a functioning nervous system but may have lost these cell types. We also find that, although *Mnemiopsis* has most of the genes coding for structural components of mesodermal cells, they lack many of the genes involved in bilaterian mesodermal specification and, therefore, may have independently evolved these cell types.

**Discussion:** These results present a newly supported view of early animal evolution that accounts for major losses and/or gains of sophisticated cell types, including nerve and muscle cells. This evolutionary framework, along with the comprehensive genomic resources made available through this study, will yield myriad discoveries about our most distant animal relatives, many of which will shed light not only on the biology of these extant organisms but also on the evolutionary history of all animal species, including our own.

---

**FIGURES IN THE FULL ARTICLE**

Fig. 1. *M. leidyi* life history and anatomy.

Fig. 2. Previously proposed relationships of the five deep clades of animals.

Fig. 3. Tree produced by maximum-likelihood analysis of the EST set.

Fig. 4. Tree produced by maximum-likelihood analysis of gene content.

Fig. 5. The origin of postsynaptic genes.

Fig. 6. Inventory of myogenic components in *M. leidyi*.

**SUPPLEMENTARY MATERIALS**

Materials and Methods

Figs. S1 to S10

Tables S1 to S31

References

The phylogenetic position of the ctenophore *Mnemiopsis leidyi* and its implications regarding the origin of mesodermal cell types. (A) Adult *M. leidyi*. (B) Summary of the relationships of the five main branches of animals and the outgroup Choanoflagellata. (C) Inventory of myogenic specification genes in *Mnemiopsis*. Components present in the *Mnemiopsis* genome are in blue, and names are underlined. Absent components are in red. The lack of many of these factors in *Mnemiopsis* indicates that ctenophore mesodermal cell types are specified differently than in bilaterians, suggesting that they perhaps evolved independently in these two lineages.

*Corresponding author. E-mail: andy@mail.nih.gov*
The Genome of the Ctenophore *Mnemiopsis leidyi* and Its Implications for Cell Type Evolution

Joseph F. Ryan,1,2 Kevin Pang,2 Christine E. Schnitzler,2 Anh-Dao Nguyen,3 R. Travis Moreland,4 David K. Simmons,3 Bernard J. Koch,1 Warren R. Francis,4 Paul Haviak, NISC Comparative Sequencing Program,5 Stephen A. Smith,6,8 Nicholas H. Putnam,7 Steven H. D. Haddock,4 Casey W. Dunn,2 Tyra G. Wolfsberg,1 James C. Mullikin,1,6 Mark Q. Martindale,3 Andreas D. Baxevanis2

An understanding of ctenophore biology is critical for reconstructing events that occurred early in animal evolution. Toward this goal, we have sequenced, assembled, and annotated the genome of the ctenophore *Mnemiopsis leidyi*. Our phylogenomic analyses of both amino acid positions and gene content suggest that ctenophores rather than sponges are the sister lineage to all other animals. *Mnemiopsis* lacks many of the genes found in bilaterian mesodermal cell types, suggesting that these cell types evolved independently. The set of neural genes in *Mnemiopsis* is similar to that of sponges, indicating that sponges may have lost a nervous system. These results present a newly supported view of early animal evolution that accounts for major losses and/or gains of sophisticated cell types, including nerve and muscle cells.

The phylogenetic position of ctenophores presents a challenge to our understanding of early animal evolution, especially as it relates to complex features such as cell types. The stark difference between the body plans of ctenophores and that of all other animals makes comparisons inherently difficult. Genomic sequencing of animals (1–4) and their closest relatives (5) provides invaluable insight into the molecular innovations contributing to the morphological diversity exhibited among modern-day animals. The vast majority of sequenced animal genomes are from Bilateria, the clade that includes most animal species (including humans and traditional model systems). Three of the four nonbilaterian metazoan lineages—Porifera (sponges), Placozoa, and Cnidaria (for example, sea anemones, corals, hydroids, and jellyfish)—have at least one species with a sequenced genome. The absence of a complete genome sequence from the fourth nonbilaterian metazoan lineage, Ctenophora (or comb jellies), has made it difficult to resolve the earliest evolutionary events in the animal tree of life and reconstruct the likely gene inventory of the most recent common ancestor of animals.

Ctenophores are gelatinous marine animals characterized by eight longitudinal rows of ciliated comb plates that run along their oral-aboral axis (Fig. 1, A to C). Their bodies consist of an inner gastrodermal layer and an outer epidermal layer separated by a mesoglea. The muscular system, deployed in discrete regions of the body (for example, in the body wall, pharynx, and tentacles), is composed almost exclusively of smooth muscle cells; however, sarcomeric muscles have been reported in a single ctenophoran species (6). The ctenophore nervous system includes the apical sensory organ, a peripheral subepithelial nerve net, neurons that run through the mesoglea, and nerves associated with the tentacles. Most ctenophores, unlike all other animals, have specialized adhesive cells called colloblasts that are involved in prey capture. Most species are hermaphrodite and capable of self-fertilization. Fertilized eggs undergo a highly stereotyped ctenophore-specific cleavage program (Fig. 1, D to M), with embryogenesis in most species leading to a free-swimming cydippid stage that displays most of the features of the adult body plan (that is, development is direct).

*Mnemiopsis leidyi* is a lobate ctenophore native to the coastal waters of the western Atlantic Ocean. This species has recently invaded the Black, Caspian, and North Seas, causing major economic and ecological impact to native species in those areas. *M. leidyi* has been used effectively to study regeneration (7), axial patterning (8, 9), and bioluminescence (10–12). In addition, a cell lineage fate map (13–15), as well as resources for collecting and spawning, has been established (16), promoting *M. leidyi* as a leading model for evolutionary and developmental studies.

The phylogenetic relationship of ctenophores to other animals has been a source of long-standing debate. The group lacks a reliable fossil record, and, on the basis of morphological features, ctenophores have been assigned various positions in animal phylogeny, including as sister to cnidarians in a clade called Coelenterata (sometimes called Radiata) (Fig. 2A) and as sister to Bilateria (Fig. 2B). Phylogenetic analyses of ribosomal RNA show little or no support uniting ctenophores with cnidarians or bilaterians and have tended to place ctenophores sister to a clade that includes all animals besides Porifera (Fig. 2C). Phylogenetic studies have also produced conflicting results, with a series of multigene analyses placing ctenophores sister to all other metazoans (Fig. 2D) (17, 18), and another, based primarily on ribosomal proteins, supporting the Coelenterata hypothesis (Fig. 2A) (19). Yet another study, also based primarily on ribosomal characters but with expanded taxon sampling, upheld the relationship of ctenophores as sister to all metazoans except Porifera (similar to Fig. 2C) (20). On the basis of its simple morphology, it has been suggested that Placozoa is the sister group to all animals (Fig. 2E) (21). Ctenophores have also been placed in a clade of nonbilaterian animals called “Diploblastica,” on the basis of a curated set of nuclear and mitochondrial proteins and a small morphological matrix (Fig. 2F) (22). The most recent analyses of the placement of sponges and ctenophores indicated that supermatrix analyses of the publicly available data are sensitive to gene selection, taxon sampling, model selection, and other factors (23). The inconsistency of reports about the phylogenetic position of ctenophores (table S1) has made it difficult to evaluate morphological, developmental, and experimental data involving these animals in an evolutionary context, complicating efforts to understand the early evolution of animals.

**Genome Sequencing and Assembly**

Genomic DNA was isolated from the embryos of two self-fertilized adult *M. leidyi* collected in Woods Hole, Massachusetts. DNA from one embryo pool was used to construct a library for Roche 454 sequencing. We generated 7.3 million raw reads, which yielded 2.5 Gb of sequence. Using the Phusion assembler (24), we assembled these data into 24,884 contigs, constituting 150 Mb of sequence and providing roughly 12-fold coverage of the genome. DNA from the other embryo pool was used to create two mate-pair libraries for Illumina GA-II sequencing, one with a 3-kb insert and the other with a 4-kb insert. After duplicate read-pairs were removed, 4.2 million and 2.6 million pairs remained for the 3- and 4-kb insert libraries, respectively. These reads were used to construct scaffolds of the original set of Roche 454 contigs. The final assembly consists of 5100 scaffolds, resulting in 160-fold physical coverage and an N50 of 187 kb (supplementary materials). To test the accuracy and completeness of our assembly, we aligned 99.4% of 15,752 public expressed sequence tags (ESTs) to our assembly. The average coverage of each alignable EST,
as determined by baa.pl (25), was 98.2%. In 94.8% of cases, a single EST mapped completely to a single scaffold. These numbers suggest that the assembly is both complete and accurately assembled.

Characteristics of the *M. leidyi* Genome

The *M. leidyi* genome is among the smallest 7% of genomes when compared with those cataloged in the Animal Genome Size Database (26) and is densely packed with gene sequences. It encodes 16,548 predicted protein-coding loci, which make up 58% of the genome, and we conservatively assign 44% of these gene predictions into homology groups with non-ctenophores. The average length of an unspliced *M. leidyi* transcript is 5.8 kb. Eight percent of predicted genes are embedded within other genes. This number of nested intronic genes is high compared to other genomes (table S2), but may be inflated owing to a subset of these being alternatively expressed exons. The level of repetitive sequence in the *M. leidyi* genome is low to moderate, as compared to other metazoans (tables S3 and S4); this has made it possible to produce a high-quality genome assembly based on paired-end and mate-pair sequencing alone. Additional characteristics of this genome are presented in tables S5 to S10.

Phylogenetic Position of *M. leidyi*

The availability of the complete genome of *M. leidyi* has allowed us to improve on the ctenophore sampling used in previous phylogenomic analyses of gene sequence evolution. We assessed two data matrices that differ in breadth of taxon sampling and fraction of missing data: a "Genome Set" that includes only data from complete genomes (13 animals, 19.6% missing data) and an "EST Set" that includes partial genomic data from many taxa (58 animals, 64.9% missing data). We analyzed both matrices by using maximum-likelihood [with the GTR+Γ model as implemented in RAxML (27)] and Bayesian [with the CAT model as implemented in PhyloBayes (28)] methods. To understand the effect of outgroup selection on our ingroup topology, we included four different sets of non-metazoan outgroups (table S11) in each combination of method and matrix. This multifactorial strategy yielded a total of 16 analyses (Table 1).

We found no support in any of these analyses for Coelenterata (Cn,Ct), Diploblastica (Bi,), or Placozoa being the sister lineage to the rest of Animals (Tr,). We recovered strong support for a sister relationship between Cnidaria and Bilateria (Cn,Bi) and for a clade of Placozoa, Cnidaria, and Bilateria (Tr,Cn,Bi). Maximum-likelihood analyses support the placement of Ctenophora as sister group to all other
Despite an average run time of 205 days per run, none of the Bayesian analyses on the EST data set are concordant with all possible topologies and (Ct,Po) is the same as (Bi,Cn,Tr). Hypotheses, and most hypotheses are represented as trees in Fig. 2. In the absence of nonanimal Choanata, choanoflagellates; and Animalia, no outgroups. Columns represent support for tested hypotheses. Most hypotheses are represented as trees in Fig. 2. In the absence of nonanimal outgroups, (Ct) and (Po) are concordant with all possible topologies and **(Ct,Po) is the same as (Bi,Cn,Tr). Despite an average run time of 205 days per run, none of the Bayesian analyses on the EST data set converged; convergence was monitored by using the maxdiff statistic generated by the bcmap program within PhyloBayes (>0.3).

**Table 1. Support for various hypotheses across 16 phylogenetic analyses.** Two amino acid matrices (Genome Set and EST Set) were analyzed with two different method/model combinations (ML indicates maximum-likelihood with the GTR+I model using RAxML (27) and Bayes is Bayesian with the CAT model using PhyloBayes (28)), using four different sets of nonmetazoan outgroups for each analysis (Opisthokonta are fungi, amoeboids, and choanoflagellates; Holozoa, amoeboids and choanoflagellates; Choanata, choanoflagellates; and Animalia, no outgroups). Columns represent support for tested hypotheses, and most hypotheses are represented as trees in Fig. 2. In the absence of nonanimal outgroups, (Ct) and (Po) are concordant with all possible topologies and **(Ct,Po) is the same as (Bi,Cn,Tr). Despite an average run time of 205 days per run, none of the Bayesian analyses on the EST data set converged; convergence was monitored by using the maxdiff statistic generated by the bcmap program within PhyloBayes (>0.3).

<table>
<thead>
<tr>
<th>Genome Set</th>
<th>ML</th>
<th>Bayes</th>
</tr>
</thead>
<tbody>
<tr>
<td>104,840 cols</td>
<td>13 animals</td>
<td>80.4% occupied</td>
</tr>
<tr>
<td>Opisthokonta</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Holozoa</td>
<td>96</td>
<td>0</td>
</tr>
<tr>
<td>Choanata</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Animalia</td>
<td>96</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EST Set</th>
<th>ML</th>
<th>Bayes***</th>
</tr>
</thead>
<tbody>
<tr>
<td>88,384 cols</td>
<td>58 animals</td>
<td>35.1% occupied</td>
</tr>
<tr>
<td>Opisthokonta</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Holozoa</td>
<td>96</td>
<td>0</td>
</tr>
<tr>
<td>Choanata</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Animalia</td>
<td>96</td>
<td>0</td>
</tr>
</tbody>
</table>

Metazoa (Ct) regardless of data matrix used (Fig. 3). The Bayesian analysis of the genome set strongly supports a clade of Ctenophora and Porifera (Ct,Po) as the sister group to all other Metazoa. This relationship also receives some support in our maximum likelihood trees, and we suspect that the result is due to poor taxon sampling in the Genome Set. However, until there are more complete genomes available to test this hypothesis, this relationship cannot be completely dismissed. Despite an average run time of 205 days per run, none of the Bayesian analyses on the EST data set converged (maxdiff > 0.3). The lack of convergence in these analyses suggests that the application of this method to this data set is insufficient to resolve this relationship.

The analyses run without nonmetazoan outgroups show strong support for a monophyletic clade of Cnidaria and Bilateria (Table 1). This evidence contradicts the idea that long-branch attraction between ctenophores and the outgroup is the sister group to all other Metazoa, but there is substantial support for ctenophore as the sister group to the rest of animals (Table 1). Furthermore, our results strongly show that Placozoa, Cnidaria, and Bilateria (that is, ParaHoxozoa) are monophyletic. Given the sensitivity of the molecular sequence evolution analyses to taxon sampling and inference method, consistent with other recent analyses (23), we also examined the evolution of gene content.

We clustered genes by using default parameters in OrthoMCL (31) and used these clusters to construct a gene presence/absence matrix. Using RAxML with a GTR+I model, we conducted a weighted likelihood-based analysis on this matrix. We then calibrated sites on the basis of the congruence of columns to known bilaterian relationships with the “–F” parameter in RAxML. The result of this analysis was a tree supporting Ctenophora as the sister group to all other animals (Ct) (Fig. 4) and the rejection of all other alternative topologies (in Fig. 2) at the 5% confidence level by likelihood-based statistical hypothesis testing (table S14). The pattern of presence and absence of gene families and signaling pathway components seen in previous studies is consistent with these results (32–36). Our reanalysis of an expanded set of near intron pairs (37) was also consistent with these results (fig. S2).

**Cell Signaling Components in M. leidyi**

Across Bilateria, there are seven major cell signaling pathways that play important roles during embryological development: Wnt, transforming growth factor–β (TGF-β), receptor tyrosine kinase (RTK), Notch, nuclear receptor, Hedgehog, and Janus kinase (JAK)/signal transducers and activators of transcription (STAT) (38). Comparisons of nonbilaterian (2–4) and nonmetazoan genomes (5, 39) show that some of these signaling pathways evolved before the evolution of animal multicellularity, others are specific to metazoan evolution, and some were lineage-specific innovations. The cell signaling components present in the M. leidyi genome include the RTK family, which predates the origin of Metazoa (40); the TGF-β signaling pathway (33), thought to have evolved in the metazoan common ancestor (39); and the canonical Wnt signaling pathway (34). Notably absent from both the TGF-β and Wnt pathways are the major bilaterian antagonists; members of the Wnt/PCP (planar cell polarity) pathway, such as Flamingo and Strabismus, are not present. Relatively few components of the Notch pathway (tables S1S and S16) are present, and many of those lack key diagnostic domains. M. leidyi also lacks most of the major genes necessary for Hedgehog signaling [for example, the Hedgehog ligand, the smoothened receptor, and SUFU (suppressor of fused)]. Last, the JAK/STAT pathway is most likely a bilaterian innovation because there are no true JAK orthologs in M. leidyi or any other nonbilaterians reported to date.

**Neural Components in M. leidyi**

Ctenophores have a nervous system consisting of a nerve net, mesogleal fibers, and tentacular nerves (41). In contrast to the cnidarian nervous system, which contains an ectodermal and endodermal nerve net, the nerve nets of ctenophores consist of polygonal nerve cords spread under the ectodermal epithelium; these nerve nets show high levels of regional specialization and concentrations associated with the apical sensory organ/ polar fields and tentacle bulbs, structures without clear homologs in any other animal groups (42). Unlike in cnidarians and bilaterians, immunological investigations have failed to detect the presence of serotonin in ctenophores (43). Ctenophore nervous systems are also unique in their abundance.
Fig. 3. Tree produced by maximum-likelihood analysis of the EST Set. The tree was produced from a matrix consisting of 242 genes and 104,840 amino acid characters. Circles on nodes indicate 100% bootstrap support. Support placing ctenophores as sister to the rest of Metazoa is 96% of 100 bootstrap replicates.
of synaptic connections and their presynaptic morphology (44).

Many of the genes known to be critical to the nervous system of bilaterians and cnidarians are present in the sponge *Amphimedon queenslandica*, an animal without a nervous system. It has been hypothesized that the origin of the nervous system in nonsponges coincided with the origin of a few neural components that are absent from *A. queenslandica* (4, 45), but our phylogenetic results and the absence of these same components in *M. leidyi* challenge this hypothesis. Both *A. queenslandica* and *M. leidyi* contain orthologs of transcription factors involved in bilaterian and cnidarian neural development, including *lhx* (46), bHLH (basic helix-loop-helix), six, gli, and sox (classes B, C, E, and F) genes. The neural differentiation RNA binding genes ELAV and Musashi, as well as the axon guidance genes neuropxin, semaphorin, plexin, and an ephrin receptor, are all present in both *A. queenslandica* and *M. leidyi*. However, netrin, slit, and unc-5, involved in axon guidance, are absent from both genomes.

Many of the genes involved in the formation of bilaterian synapses and neural differentiation are present in both *A. queenslandica* and *M. leidyi*—but again, sponges and ctenophores lack a similar set of synaptic scaffolding genes (tables S17 and S18), all of which are present in cnidarians and bilaterians (Fig. 5). The pattern of presence and absence of these scaffolding genes is consistent with these genes being primitively absent in sponges and ctenophores. Almost all of the enzymes involved in the biosynthesis of dopamine and other catecholamine neurotransmitters are also absent in both *A. queenslandica* and *M. leidyi* (table S19). An exception to this shared pattern with sponges is the presence of two definitive opsin genes in *M. leidyi*, but not *A. queenslandica*, that are expressed in photoreceptors (light-producing cells), as well as in putative photosensory cells in the apical sense organ (12).

**Mesoderm Components in *M. leidyi***

Ctenophores have several cell types (such as distinct muscle cells and mesenchymal cells) that, in bilaterians, are characteristically derived from mesodermal tissues. Cell lineage studies (14) have indicated that these cells are derived from a true endomesoderm because mesodermal cells are generated from precursors that also give rise to the endodermal portions of the gut; this is similar to the
endomesodermal origins of mesoderm in virtually all bilaterians. However, screening the *M. leidyi* genome reveals a surprising result in that almost none of the genes involved in bilaterian mesoderm development can be found (Fig. 6 and tables S20 and S21). Functional components of the fibroblast growth factor, Notch, Hedgehog, and the nodal (TGFB superfamily) pathways, all of which are important in the segregation of mesoderm in different bilaterian forms, are also not observed. Other genes known to be involved in bilaterian mesoderm development, such as gli/glis genes, are expressed in neural (but not mesodermal) cells in *M. leidyi* (47).

**Mesoderm and Neural Components Also Absent from Other Ctenophores**

To test whether these absences from the *M. leidyi* genome were true for other ctenophores, we searched the deeply sequenced transcriptomes of seven other ctenophore species (*Bathyctena chuni*, *Beroe forskalii*, *Charistephane fugiens*, *Euplokamis dunlapae*, *Hormiphora californensis*, *Lampea lactea*, and *Thalassocalyce inconstans*) for FGF (fibroblast growth factor), Hedgehog, nodal, twist, snail, Lbx, NK4, NK3, NK2, Myf5, Noggin, Mrf4, Myogenin, Eomesoderm, GATA, MyoD, and troponin. We were able to identify putative snail genes in *T. inconstans* and *E. dunlapae* and putative GATA genes in five of the seven species. We were unable to identify the other 15 missing genes in any of these ctenophore transcriptomes (tables S22 and S23). A phylogenetic analysis of ionotropic glutamate receptor sequences from *M. leidyi* and these ctenophore transcriptomes suggests that the ctenophore receptors form a sister clade to the bilaterian glutamate receptors.

---

**Fig. 5. Origin of postsynaptic genes.** A possible configuration for postsynaptic genes. Genes are colored by their node of origin (inset). Names of genes present in the *M. leidyi* genome are underlined. Accession numbers of *M. leidyi* genes are given in table S16.

**Fig. 6. Inventory of myogenic components in *M. leidyi*.** Components present in the *M. leidyi* genome are in blue, and names are underlined. Absent components are in red. (A) The main structural components of smooth muscle are present in the *M. leidyi* genome. All structural components are present except for troponin (in red). (B) The majority of signaling molecules and transcription factors involved in specifying and differentiating the mesoderm of bilaterian animals are absent from the genome of *M. leidyi*. The asterisks next to Snail and GATA indicate that these components have been identified in the transcriptomes of other ctenophores.
are not specific to the
discussion here. Within ctenophores, the majority of absences
from the rest of animals. These results indicate
that, like glutamate receptors after ctenophores diverged
within the lineage of radiate cnidarians and bilaterians, which might explain some
of the most distant animal relatives, many of which will
shed new light not only on the biology of these
together with alternative hypotheses. With a
taken as a whole, we show that gene
content data support Ctenophora as the sister
group to all other animals, a conclusion supported
by phylogenetic analysis of amino acid matrices
from concatenated protein sequences. However,
these analyses are sensitive to taxon sampling and phylogenetic methods and, therefore, provide
some support for alternative hypotheses. With a
catenophore genome in hand, we show that gene
content data support Ctenophora as the sister
group to all other animals and statistically reject
competing hypotheses. It will be important to test
this result once more genomic data are available from other ctenophores, sponges, and other relevant
groups. Nevertheless, this result is congruent with the
structure and inventory of a variety of genes
families and signaling pathways, as well as genes
essential to neural and mesodermal cell types.
It appears that much of the genetic machinery
necessary for a nervous system was present in the
ancestor of all extant animals. This pattern suggests
that a less elaborate nervous system was present in the
tetazonan ancestor and was secondarily re-
duced in placozoans and sponges. The alternative
is that neural cell types arose independently in both
the ctenophore lineage and the lineage that led to
cnidarians and bilaterians, which might explain some
of the unique aspects of the ctenophore nervous
system. Resolving these alternative hypotheses will
require functionally characterizing the nervous
system-related genes in ctenophores and sponges.
Like the nervous system, the mesoderm
appears to have had a complex evolutionary history.
Our results are consistent with several alternative
hypotheses. One possibility is that the mesoderm
was present in the most recent common ancestor
of ctenophores and bilaterians but was lost in
sponges, placozoans, and cnidarians. However,
given the absence of the majority of genes involved
in the specification and differentiation of the
bilaterian mesoderm from the M. leidyi genome,
it appears more likely that ctenophores indepen-
dently evolved mesodermal cell types after they
diverged from the rest of animals. This interpre-
tation is compatible with a recent report that striated
musculature evolved independently in bilaterians,
cnidarians, and in the ctenophore E. dunalpae (49).
The implications of these findings go well be-
dy the rearrangement of the branches of the meta-
zoan tree of life, arguing for a new way of thinking
regarding the emergence and/or conservation of
what heretofore were considered to be unique
and indispensable biological features. Likewise,
theories on the evolution of animal multicellularity
have to be reevaluated. This evolutionary frame-
work, along with the comprehensive genomic
resources made available through this study, will
undoubtedly yield myriad new discoveries about
our most distant animal relatives, many of which will
shed new light not only on the biology of these
extant organisms but also on the evolutionary
history of all animal species, including our own.

Discussion and Conclusion
The sequence of the M. leidyi genome has given rise
to multiple categories of evidence that support
the placement of ctenophores as the sister
group to all other animals, a conclusion supported
by phylogenetic analysis of amino acid matrices
from concatenated protein sequences. However,
these analyses are sensitive to taxon sampling and phylogenetic methods and, therefore, provide
some support for alternative hypotheses. With a
catenophore genome in hand, we show that gene
content data support Ctenophora as the sister
group to all other animals and statistically reject
competing hypotheses. It will be important to test
this result once more genomic data are available from other ctenophores, sponges, and other relevant
groups. Nevertheless, this result is congruent with the
structure and inventory of a variety of genes
families and signaling pathways, as well as genes
essential to neural and mesodermal cell types.

Methods

Phylogenetic Analysis of Concatenated
Gene Matrices
We analyzed two matrices constructed from con-
catenated protein sequences. One consisted of
M. leidyi amino acids added to a genome-based
data matrix that was reported in the A. queenslandica
genome paper (4). The second used a phenetic
sequence clustering method as described previously
(18). We generated maximum-likelihood
trees with the GTR+T model using RAxML (27)
and Bayesian trees with the CAT model using
PhyloBayes (29). All alignments and trees are avail-
able at http://research.nhgri.nih.gov/manuscripts/
Baxevanis/science2013_supplement/

Phylogenetic Analysis of Gene Content
We assembled a presence/absence matrix of gene
clusters and analyzed these data with RAxML
under the GTR-gamma model of rate heteroge-
neity. We used known bilaterian relationships to
generate a weight matrix in RAxML. We used
per-site log likelihoods generated in RAxML as
input to CONSEL (53) to generate P values for
alternative hypotheses.

References and Notes
1. J. A. Chapman et al., The dynamic genome of Hydra.
pmid:20228792
2. N. H. Putnam et al., Sea anemone genome reveals
ancestral eumetazoan gene repertoire and genomic
3. www.nhgri.nih.gov/monkeys/mnhg_home.htm;
pmid:17635350
4. M. Srivastava et al., The Trichoplax genome and the
nature of placozoans.
pmid:18715981
5. M. Srivastava et al., The Amphimedon queenslandica
genome and the evolution of animal complexity.
pmid:20686567
6. N. King et al., The genome of the chaenoflagellate
Monosiga brevicollis and the origin of metazoa.
pmid:18273011
7. G. O. Mackie, C. E. Mills, C. L. Singla, Structure and
function of the prehensile tentilla of Euplokomis
(Ctenophora, Cydippida). Zoomorphology 107, 319–337
8. J. Q. Henry, M. Q. Martindale, Regulation and
regeneration in the ctenophore Mnemiopsis leidyi.
pmid:11071786
and the evolutionary origins of the bilaterian body plan.
pmid:12220977
10. K. Pang, M. Q. Martindale, Developmental expression
of homeobox genes in the ctenophore Mnemiopsis
doi:10.1007/s00222-008-0257-2; pmid:18504608
11. M. Anctil, Ultrastructure of the luminescent system of
the ctenophore Mnemiopsis leidyi. Cell Tissue Res.
12. G. Freeman, G. T. Reynolds, The development of
bioluminescence in the ctenophore Mnemiopsis
pmid:4150750
13. E. Schnitzler et al., Genetic specification, evolution,
and expression of photoprotein and opsin genes in
Mnemiopsis leidyi: A new view of ctenophore photocytes.
pmid:22325493
in the Ctenophora: The inductive role of β1 microtubules
in organizing cten row formation in the “mosaic
embryo, Mnemiopsis leidyi. Development 124,
15. M. Q. Martindale, J. Q. Henry, Intracellular fate mapping
in a basal metazoan, the ctenophore Mnemiopsis
leidyi, reveals the origins of mesoderm and the existence
of indeterminate cell lineages. Dev. Biol. 214, 243–257
16. G. Reverberi, G. Ortolani, On the origin of the
ciliated plates and mesoderm in the Ctenophore.
17. K. Pang, M. Q. Martindale, Mnemiopsis leidyi spawning
and embryo collection. CSH Protoc. 2008, pdb.prot5085
18. A. Dunn et al., Broad phylogenomic sampling
improves resolution of the animal tree of life. Nature
pmid:18322464
19. D. Hejrai et al., Assessing the root of bilaterian animals
pmid:19759036
20. H. Philippe et al., Phylogenomics revises traditional views

Downloaded from http://science.sciencemag.org on May 28, 2019


The Base of the Animal Tree?

The identity of the most basal lineages of the animal kingdom evolutionary tree has long been contested. Ryan et al. (p. 10.1126/science.1242592; see the Perspective by Rokas) sequenced the genome of the ctenophore the warty comb jelly or sea walnut, *Mnemiopsis leidyi*, and conclude that ctenophores alone, not sponges or the clade consisting of both ctenophores and cnidarians, are the most basal extant animals. The results suggest a specific evolutionary process that likely occurred—including repeated gains and loss of mesoderm, expansion of genes associated with the cell cycle, growth signaling, apoptosis, and epithelial and neural cell types. Furthermore, previous hypotheses regarding the evolution of animals may require re-evaluation.