

Minireview

The gene complement of the ancestral bilaterian - was Urbilateria a monster?

David J Miller* and Eldon E Ball†

Addresses: *ARC Centre of Excellence for Coral Reef Studies and Comparative Genomics Centre, James Cook University, Townsville, Queensland 4811, Australia. †Centre for Molecular Genetics of Development and School of Biology, Australian National University, Canberra, ACT 2601, Australia.

Correspondence: David Miller. Email: david.miller@jcu.edu.au

Abstract

Expressed sequence tag analyses of the annelid *Pomatoceros lamarckii*, recently published in *BMC Evolutionary Biology*, are consistent with less extensive gene loss in the Lophotrochozoa than in the Ecdysozoa, but it would be premature to generalize about patterns of gene loss on the basis of the limited data available.

See research article <http://www.biomedcentral.com/1471-2148/9/240>.

The pioneers of pyrosequencing have a lot to answer for. The availability of 'next generation' DNA sequencers has provided zoologists with unforeseen opportunities to address many basic evolutionary issues and, for those of us whose interests lie beyond the model organisms, these are indeed interesting times. Not so long ago the costs of large-scale expressed sequence tag (EST) analyses were prohibitive, but the recent development of the fast and (relatively) cheap 454, Illumina and SOLiD technologies is enabling large-scale transcriptome analysis, and potentially whole-genome analysis, to be applied to a wide range of animals, providing insights into evolutionary issues that were once considered essentially intractable.

One important and controversial issue that can now be addressed is the gene complement of Urbilateria. It is clear that this ancestor of all bilateral animals had a genome resembling that of a modern vertebrate, but which also contained some genes lost from modern vertebrates, raising the issue of just how many genes were present in the ancestor. In this respect, EST studies on lophotrochozoans, such as that reported in a recent paper in *BMC Evolutionary Biology* [1], are proving particularly informative. Of the three major divisions within Bilateria (the Ecdysozoa, the Lophotrochozoa and the Deuterostomia; Figure 1), Lophotrochozoa, which contains the annelids (worms) and mollusks (including snails) and various minor phyla, is still only poorly represented in terms of whole genome data.

One implication of work on the annelid *Platynereis dumerilii* is that lophotrochozoans may be less derived (are more representative of ancestral character states) than

members of Ecdysozoa [2], but it is not yet clear how representative *Platynereis* is. Takahashi *et al.* [1] have now analyzed a set of ESTs from a second and only distantly related annelid, the serpulid *Pomatoceros lamarckii*, which differs from *Platynereis* both morphologically and in lifestyle. *Platynereis* is a free-living predator, whereas *Pomatoceros* lives within a tube that it constructs and captures food from the surrounding waters using a crown of feeding tentacles through which water is filtered. Nevertheless, data from the two species lead to the same conclusion, that annelids (and perhaps lophotrochozoans in general) are less derived than the insects and nematodes investigated so far. One focus of the paper [1] was patterns of gene sharing and gene loss between *Pomatoceros* and the other major groups of organisms. These figures are summarized in Figure 1: *Pomatoceros* shares a significant number of genes (158; 7% of the total) only with deuterostomes and other lophotrochozoans, but a much smaller number (23 genes; 1% of the total) only with ecdysozoans and other lophotrochozoans. In addition, 11 genes shared only with non-bilaterians were identified, illustrating the ubiquity of gene loss.

The significance of cnidarians for understanding the urbilaterian gene complement

Although the genome of a choanoflagellate (thought to be the closest living relatives of the animals) [3] showed us that some 'animal-specific' genes came earlier in evolution, many other genes really are unique to metazoans, and these include components of several signaling pathways (such as the Wnt, transforming growth factor β and nuclear hormone receptor pathways). When the genomes of bilaterians are compared with those (admittedly, as yet few) available for 'lower' (non-bilaterian) animals, one fact that clearly stands out is that a quantitative leap in terms of signaling molecule complexity preceded the emergence of the Cnidaria (the phylum that includes hydras, sea anemones, corals and jellyfish). Whereas most or all of the animal-specific signaling systems seem to be present in the genomes of Porifera (sponges) and Placozoa (placozoans; very simple animals with only three or four distinct cell

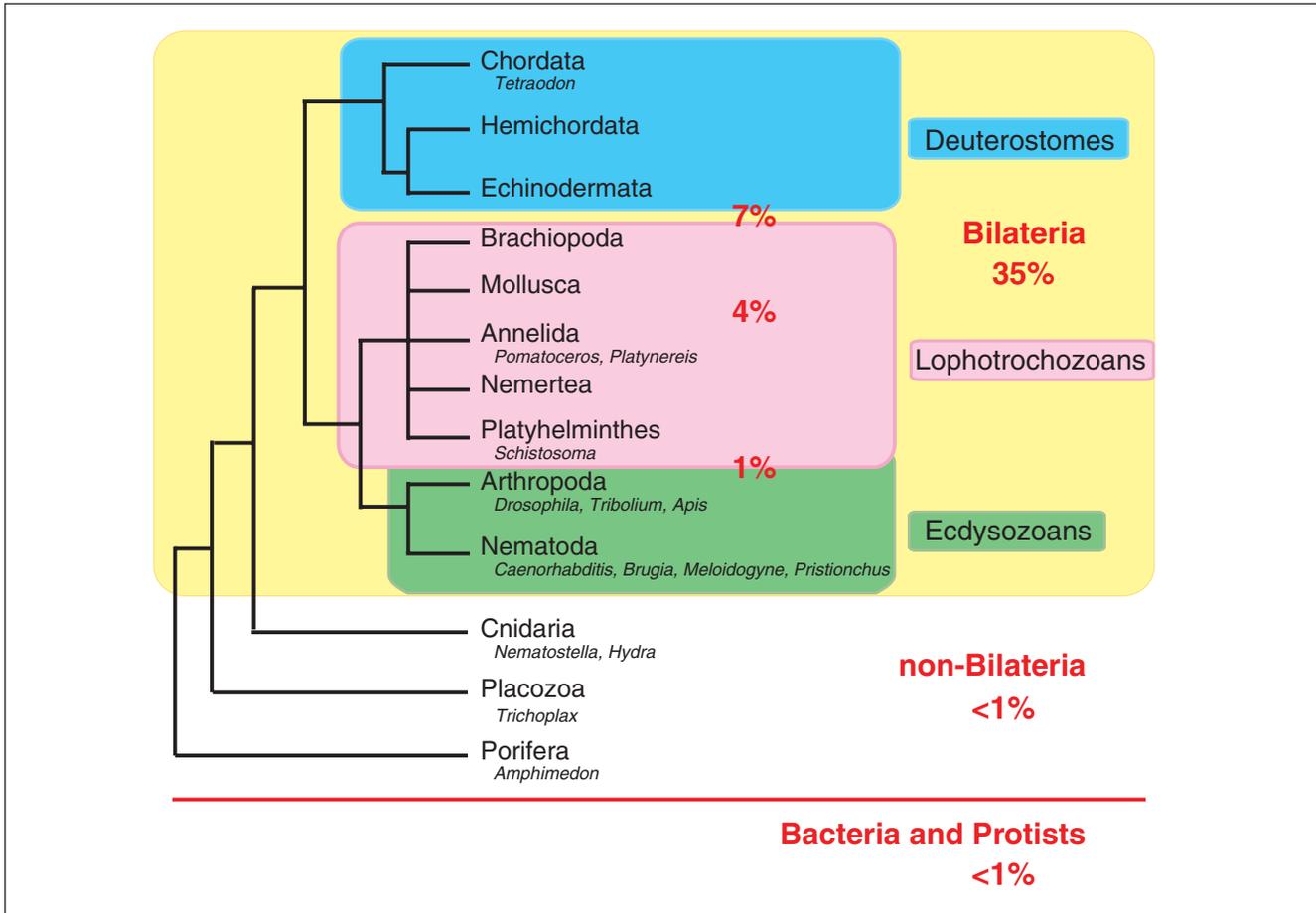


Figure 1

A simplified view of animal phylogeny, showing the taxonomic position of groups and organisms mentioned in the text (genera are in italics). Taxa above the red line are animals. Relationships among the non-bilaterian phyla remain controversial, but the topology shown reflects the current consensus. Numbers in red are percentages of the total number (2,308) of *Pomatoceros* ESTs with matches against specific taxonomic groups. Numbers on the boundaries between taxonomic groups are shared exclusively between *Pomatoceros* and those groups, whereas in the cases of non-bilaterians and bacteria/protists the numbers reflect ESTs shared between *Pomatoceros*, lophotrochozoans and those taxonomic groups. For example, 7% of *Pomatoceros* ESTs are shared only with lophotrochozoans and deuterostomes, and less than 1% are shared only with lophotrochozoans and non-bilateral animals or with lophotrochozoans and bacteria or protists. Over half (1,205; 52%) of the total number of *Pomatoceros* ESTs had no matches in the databases.

types), they are much less highly elaborated than in cnidarians or bilaterians. For example, whereas the sea anemone *Nematostella* has 12 Wnts, most of which are recognizable as homologs of specific Wnt types known from bilaterians [4], the sponge *Amphimedon* and the placozoan *Trichoplax* each have only three Wnts that are not easily assignable. In addition, these lower animals seem to have much less well developed arsenals of signaling molecule antagonists. The situation with respect to transcription factors is a little less clear cut but, for example, the homeobox gene complement of *Nematostella* is much more bilaterian-like than are those of sponges and placozoans [5]. In summary, cnidarians seem to be particularly important in terms of understanding the urbilaterian gene repertoire.

Have lophotrochozoans lost fewer genes?

There is a widespread perception that ecdysozoans have lost more of the ancestral gene set than have deuterostomes or lophotrochozoans. This notion has its roots in early comparisons (for example, [6]) between cnidarians, vertebrates and the model ecdysozoans (fly and nematode worm), which clearly demonstrated that gene loss was much more extensive in *Drosophila* and *Caenorhabditis* than in vertebrates. Since that time, genome data have become available for a broader range of species, so to what extent does this generalization still hold? Is it possible (or meaningful) to generalize - have ecdysozoans in general lost more genes than lophotrochozoans or deuterostomes, or do we still have too few whole genome sequences to be able to say?

There are now whole-genome data for over 20 insect species (this includes 12 *Drosophila* species) and a handful (5) of nematodes. Comparisons between insects and vertebrates (for example, [7]) indicate that gene loss is largely a function of rates of evolution and divergence times, and does not discriminate between vertebrates or insects. Among insects, *Drosophila* has a particularly high rate of evolution, whereas the beetle *Tribolium* and the honeybee *Apis* have lower rates of evolution and have lost fewer of the ancient genes present in Urbilateria. Although vertebrates in general have lost fewer ancient genes, the chicken is a clear outlier, having lost more genes in the 'universal single-copy orthologs' and 'universal multi-copy orthologs' categories than any of the five insects included in the Wyder *et al.* analysis [7]. Moreover, whereas the two *Caenorhabditis* species and the parasitic species *Brugia malayi* and *Meloidogyne hapla* all show the 'typical' ecdysozoan pattern of extensive gene loss, a fifth nematode species, *Pristionchus pacificus*, is not so reduced [8].

What about lophotrochozoans? Whole genome sequences have been determined for seven genera (eight species; two *Schistosoma* species), but so far very few large-scale analyses have been published. The idea that the genomes of lophotrochozoans are less derived than those of ecdysozoans comes largely from work on the annelid *P. dumerilii*, which shows that this organism is closer to vertebrates than to ecdysozoans in intron structure and retention, and in protein coding sequence similarity (see, for example, [2]). The only lophotrochozoans with sequenced and analyzed genomes, *Schistosoma mansoni* and *S. japonicum*, are both parasitic plathyhelminths, which, consistent with other parasites, have undergone extensive gene loss and divergence. Thus, they can hardly be considered representative of phyla consisting mostly of free-living forms. There are bound to be derived lophotrochozoans, just as there are derived ecdysozoans and derived deuterostomes.

So, although this limited sample of two annelids [1] is consistent with greater gene loss in Ecdysozoa than in Lophotrochozoa, it is still very early days, and it would be premature to yet draw general conclusions. We await with interest the analysis and publication of more lophotrochozoan genomes, particularly those of free-living flatworms, mollusks, and some of the smaller phyla, such as bryozoans, nemertines and brachiopods.

Any gene can go

One surprising implication of comparative genomics is that no gene is indispensable; every animal seems to have lost hundreds of what one might have assumed were 'core requirement' genes. For example, Wyder *et al.* [7] report that 40% of ancient orthologous genes were lost in a least one of the ten animals included in their analysis (five insects and five vertebrates). One example of loss of a core gene is the case of the Toll receptor in *Hydra magnipapillata*.

Whereas *Nematostella* and other members of the basal cnidarian class Anthozoa have a canonical Toll receptor, *Hydra* (which is a member of the more derived class Hydrozoa) has lost this gene [9]. *Hydra* seems to have undergone non-orthologous gene replacement, Toll receptor function being fulfilled by two unrelated proteins [10]. Evolution sometimes dispenses with whole pathways, for instance the entire DNA methylation system in the case of dipterans (flies and mosquitoes).

Inferring the ancestral complement - was Urbilateria a monster?

All animals have lost genes, but it does not follow that Urbilateria was a monster in terms of gene content. Bilaterian animals typically have around 20,000 genes (range 11,500 to 28,000; *Meloidogyne* to *Tetraodon*), but a substantial fraction of these are taxonomically restricted at some level. Many of these taxonomically restricted genes are paralogs or highly diverged members of large gene families, generated by duplication events that have occurred at all levels. On the basis of the currently available data, the core bilaterian gene set probably contained fewer than 10,000 genes, the caveat being that the available data are rather limited. Taxonomic gaps need to be plugged, and more data for non-bilaterians in particular will be critical in revealing the genomic makings of Urbilateria. Far too few whole-genome sequences are yet available for firm estimates to be made, but it is clear that there is no need to invoke monsters - either hopeful or hopeless.

References

1. Takahashi T, McDougall C, Troscianko J, Chen W-C, Jayaraman-Nagarajan A, Shimeld SM, Ferrier DEK: **An EST screen from the annelid *Pomatoceros lamarckii* reveals patterns of gene loss and gain in animals.** *BMC Evol Biol* 2009, **9**:240.
2. Raible F, Tessmar-Raible K, Osoegawa K, Wincker P, Jubin C, Balavoine G, Ferrier D, Benes V, de Jong P, Weissenbach J, Bork P, Arendt D: **Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*.** *Science* 2005, **310**: 1325-1326.
3. King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I, Marr M, Pincus D, Putnam N, Rokas A, Wright KJ, Zuzow R, Dirks W, Good M, Goodstein D, Lemons D, Li W, Lyons JB, Morris A, Nichols S, Richter DJ, Salamov A, JGI Sequencing, Bork P, Lim WA, Manning G, *et al.*: **The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans.** *Nature* 2008, **451**:783-788.
4. Kusserow A, Pang K, Sturm C, Hroudá M, Lentfer J, Schmidt HA, Technau U, von Haeseler A, Hobmayer B, Martindale MQ, Holstein TW: **Unexpected complexity of the Wnt gene family in a sea anemone.** *Nature* 2005, **433**:156-160.
5. Larroux C, Luke GN, Koopman P, Rokhsar DS, Shimeld SM, Degnan BM: **Genesis and expansion of metazoan transcription factor gene classes.** *Mol Biol Evol* 2008, **25**:980-996.
6. Kortschak RD, Samuel G, Saint R, Miller DJ: **EST analysis of the cnidarian *Acropora millepora* reveals extensive gene loss and rapid sequence divergence in the model invertebrates.** *Curr Biol* 2003, **13**:2190-2195.
7. Wyder S, Kriventseva EV, Schröder R, Kadowaki T, Zdobnov EM: **Quantification of ortholog losses in insects and vertebrates.** *Genome Biol* 2007, **8**:R242.

8. Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, Dinkelacker I, Fulton L, Fulton R, Godfrey J, Minx P, Mitreva M, Roeseler W, Tian H, Witte H, Yang SP, Wilson RK, Sommer RJ: **The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism.** *Nat Genet* 2008, **40**:1193-1198.
9. Miller DJ, Hemmrich G, Ball EE, Hayward DC, Khalturin K, Funayama N, Agata K, Bosch TC: **The innate immune repertoire in Cnidaria - ancestral complexity and stochastic gene loss.** *Genome Biol* 2007, **8**:R59.
10. Bosch TC, Augustin R, Anton-Erxleben F, Fraune S, Hemmrich G, Zill H, Rosenstiel P, Jacobs G, Schreiber S, Leippe M, Stanisak M, Grötzinger J, Jung S, Podschun R, Bartels J, Harder J, Schröder JM: **Uncovering the evolutionary history of innate immunity: the simple metazoan *Hydra* uses epithelial cells for host defence.** *Dev Comp Immunol* 2009, **33**: 559-569.

Published: 19 November 2009

doi:10.1186/jbiol192

© 2009 BioMed Central Ltd