Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida

Sarah J. Bourlat1, Thorhildur Juliusdottir2, Christopher J. Lowe3, Robert Freeman4, Jochanan Aronowicz3, Mark Kirschner5, Eric S. Lanter6,7, Michael Thorndyke7, Hiroaki Nakano7, Andrea B. Kohn8, Andreas Heyland8, Leonid L. Moroz9, Richard R. Copley7 & Maximilian J. Telford1

Deuterostomes comprise vertebrates, the related invertebrate chordates (tunicates and cephalochordates) and three other invertebrate taxa: hemichordates, echinoderms and Xenoturbellida. The relationships between invertebrate and vertebrate deuterostomes are clearly important for understanding our own distant origins. Recent phylogenetic studies of chordate classes and a sea urchin have indicated that urochordates might be the closest invertebrate sister group of vertebrates, rather than cephalochordates, as traditionally believed8–10. More remarkable is the suggestion that cephalochordates are closer to echinoderms than to vertebrates and urochordates, meaning that chordates are paraphyletic. To study the relationships among all deuterostome groups, we have assembled an alignment of more than 35,000 homologous amino acids, including new data from a hemichordate, starfish and Xenoturbella. We have also sequenced the mitochondrial genome of Xenoturbella. We support the clades Olfactores (urochordates and vertebrates) and Ambulacaria (hemichordates and echinoderms)6. Analyses using our new data, however, do not support a cephalochordate and echinoderm grouping and we conclude that chordates are monophyletic. Finally, nuclear and mitochondrial data place Xenoturbella as the sister group of the two ambulacrarian phyla1. As such, Xenoturbella is shown to be an independent phylum, Xenoturbellida, bringing the number of living deuterostomes to four.

Although molecular phylogenetic studies have prompted changes in the membership of the deuterostome superphylum (two groups—chaetognaths11 and lophophorates12)—have been reclassified as protostomes and one species, Xenoturbella bocki, added13 there remains an unchanging core of three phyla: echinoderms, hemichordates and chordates. Three main groups are recognized within the chordates themselves—urochordates, including the ascidians and larvacens; cephalochordates or lancelets; and vertebrates, including fish and tetrapods.

Molecular studies have also changed our view of the relationships between deuterostome taxa. Hemichordates, long thought of as the sister group of the chordates, are now considered to be closest to echinoderms in a supraphyletic clade named the Ambulacaria1. Most recently, analyses of hundreds of genes from expressed sequence tag (EST) and whole-genome sequencing projects have gathered data from the two groups previously missing: a hemichordate, starfish and Xenoturbella. We added to the urochordate plus vertebrate clade10) seem to be credible components of the deuterostomes, two further aspects of the phylogeny of the group remain contentious.

The first stems from a remarkable result in a phylogenomic analysis in which cephalochordates were associated with echinoderms rather than with other chordates1. This result, if correct, has profound implications for our understanding of early deuterostome evolution and the origins of the chordate body plan, because it suggests that a number of conserved chordate features were present in the common ancestor of all deuterostomes and were secondarily lost in echinoderms and hemichordates. Although this is clearly a controversial finding, it deserves serious consideration. The initially surprising relocation of the hemichordates from chordate sister group to the Ambulacaria has already shown that certain characteristics that were thought to be specific to the chordate lineage are likely to derive from the deuterostome common ancestor; the homology of gill slits and endostyle in hemichordates and chordates is now supported by studies both of morphology and of shared gene-expression patterns6,11,12.

The second open question concerns the position of the worm Xenoturbella. Consideration of its small subunit ribosomal RNA gene (SSU) indicated that Xenoturbella might be a deuterostome, related to the Ambulacaria but not included in either the echinoderms or the hemichordates1. This result is supported by the lack in Xenoturbella of a complex molecular character that unites the two ambulacrarian phyla—a change in their mitochondrial genetic code whereby the codon ATA codes for the amino acid isoleucine rather than for methionine as in most other animals, including Xenoturbella1. At the same time, there are several reasons to question the position of Xenoturbella. There is low statistical support for its exclusion from the hemichordates in the SSU study1, the change in genetic code is known to be homoplastic because it has occurred independently in the rhabditophoran playthelminths13; separate analyses of mitochondrial Cytochrome oxidase I sequences placed Xenoturbella with the hemichordates1, and there are detailed similarities between the structure of the epidermis of hemichordates and Xenoturbellal4 that have led to suggestions that Xenoturbella is a neotenous hemichordate15.

To test the robustness of previous work and to extend the phylogenomic approach to all of the main deuterostome groups, we have gathered data from the two groups previously missing: a hemichordate, Saccoglossus kowalevski (206,715 ESTs including those from Lowe et al.16) and Xenoturbella bocki (1,372 ESTs). We gathered data from the two groups previously missing: a hemichordate, Saccoglossus kowalevski (206,715 ESTs including those from Lowe et al.16) and Xenoturbella bocki (1,372 ESTs). We...
the sampling of the echinoderm clade with data from the star Solaster stimpsoni (5,614 ESTs). We searched these ESTs for orthologues of the 146 genes used by Philippe et al.22. We then supplemented these with 35 genes having a representative orthologue in our Xenoturbella ESTs, as this taxon had the fewest matches to the data set. Xenoturbella is the sister group of the Ambulacraria by performing phylogenetic analyses using maximum-likelihood and bayesian approaches recapitulated previous results (Fig. 1a). All nodes received maximum support from bayesian analysis (bayesian posterior probability (BPP) = 100%). Maximum-likelihood non-parametric bootstrap support (MLBP) was lower for some nodes (see Supplementary Information). The overall topology of this tree conforms to the ‘new animal phylogeny’18 with monophyletic protostomes, deuterostomes, lophotrochozoans (represented here by molluscs and an annelid), and ecdysozoans (represented here by arthropods and a nematode; see Supplementary Information). Within the deuterostomes, we find support for monophyletic groups of vertebrates, cyclostomes (hagfish and lamprey) and Olfactores (vertebrates plus urochordates). Importantly, we recovered the relationship linking cephalochordates not to other chordates but to echinoderms: BPP = 100%, Treefinder17 LRSH support (TF) = 75% and MLBP = 81%. This relationship was unexpected and we wanted to test the possibility that it was caused by systematic error. Two powerful ways of addressing systematic errors are to include additional taxa, thereby enabling better detection of misleading multiple substitutions, and to remove problematic taxa (for example, long branches); we have used both approaches29.

By adding sequences from three taxa that are likely to be related to the echinoids (asteroid, hemichordate and Xenoturbella) we could test the possibility that the cephalochordate/echinoderm relationship might result from a systematic error such as long-branch attraction21. The effect of including these three taxa in our analysis is clear; support for a cephalochordate plus echinoderm clade disappears and support for the canonical monophyletic chordates increases (Fig. 1b). The monophyletic chordate clade in our new analysis has BPP = 100% and TF= 53%, but no MLBP support. This change in topology on the addition of taxa that divide the echinoid branch is consistent with the idea that the attraction between the cephalochordate and the echinoid results from a systematic error; the addition of non-problematic taxa tends to improve the reliability of a phylogeny21. We compared the support for the optimal topology to that for the alternative topology, which groups the cephalochordates with Ambulacraria and Xenoturbella using the Shimodaira–Hasegawa (SH) tests and Bayes factors. The SH test did not reject the alternative topology as significantly worse than the optimal topology (SH P = 0.476), but the alternative topology was rejected as significantly worse using the Bayes factors test (2loge(B10) = 24), although this test might not be conservative. The inclusion of data from the other major taxon of cephalochordates—the Epigonichthysidae—will divide the cephalochordate branch and, we predict, will further increase support for the monophyletic chordates.

A parallel possibility is that rapidly evolving urochordates, along with their sister group the vertebrates, have been attracted towards the base of the tree by long-branch effects22. Two lines of evidence support this possibility. First, in the analyses described above, we used diploblast metazoans (ctenophore and cnidarian) as close outgroups; the inclusion of more distant outgroups (yeast and choanoflagellates), as in ref. 2, results in a more basal position for Olfactores with the corollary that cephalochordates group with echinoderms. This result is observed even when the asteroid, hemichordate
and Xenoturbella data are included. Second, exclusion of the long-branched urochordates always results in cephalochordates grouping with vertebrates with or without our additional taxa and even in the presence of distant outgroups (see Supplementary Information).

We also considered the phylogeny of the other deuterostome groups. The hemichordate is the sister group of the echinoderms (asteroid and echinoid), which gives weight to a monophyletic Ambulacraria. Of greater interest is our finding that Xenoturbella is related to but outside the monophyletic Ambulacraria. The clade (Xenoturbella plus Ambulacraria) has BPP = 100%, TF = 99% and MLBP = 68% whereas ambulacarian monophyly excluding Xenoturbella has BPP = 100%, TF = 99.3% and MLBP = 100%. We compared the support for the optimal topology to that for alternative possibilities in which Xenoturbella is grouped with hemichordates or echinoderms using the SH tests and Bayes factors. Both tests rejected as significantly less well supported the grouping of Xenoturbella with hemichordates (SH P = 0.012, Bayes factors 2log(B10) = 96) or Xenoturbella with echinoderms (SH P = 0.037, Bayes factors 2log(B10) = 118).

To test our findings further, we sequenced the entire mitochondrial genome (15,234 nucleotides) of Xenoturbella bocki. We have used the amino-acid sequences of all 13 protein-coding loci to examine the position of Xenoturbella within the deuterostomes. Our initial results placed Xenoturbella at the base of the deuterostomes rather than as the sister group of Ambulacraria. We were aware, however, that the mitochondrial genetic code changes in the lineage leading to Ambulacraria and an additional change in the echinoderms might have led to compositional bias that could artefactually distance them from Xenoturbella. Indeed, the proportions of amino acids that were affected by the changes of genetic code (methionine and isoleucine, asparagine and lysine (M, I, N, K)) in the echinoderms differ from the average seen in deuterostomes and metazoans (Table 1). We tested the effects of this imbalance on the position of Xenoturbella in two ways. First we excluded the most biased taxa—the echinoderms—leaving the hemichordates as representatives of Ambulacraria. The result of excluding the echinoderms is to move Xenoturbella next to the hemichordates, as predicted if its basal position were due to compositional bias in the echinoderms (see Supplementary Information). Second, we avoided the influence of the problematic amino acids (M, I, N, K) by recoding each instance of them as ‘missing data’, which affected 33.9% of our reliably aligned amino acids. The effect of this was again to move Xenoturbella to become the sister group of Ambulacraria. Xenoturbella plus Ambulacraria monophyly has BPP = 100%, TF = 83% and MLBP = 82% whereas ambulacarian monophyly excluding Xenoturbella has BPP = 100%, TF = 68% and MLBP = 70% (Fig. 1c).

We note that according to our mitochondrial data set, the cephalochordates and vertebrates form a highly supported monophyletic group (BPP = 100%, TF = 96.5% and MLBP = 91%), which further diminishes the credibility of the cephalochordate–echinoderm link. In common with previous analyses, our tree reconstructions place the long-branched urochordates as a sister group to all other Bilateria.

What conclusions can we draw from this knowledge of the relationships of all the main deuterostome clades? Our analyses indicate that the relationship between cephalochordates and echinoderms probably results from systematic error. This makes excellent sense from a morphological standpoint because chordates share several characteristics (somites, dorsal nerve chord, notochord and hypophysis) that are absent from Ambulacraria, Xenoturbella and non-deuterostome phyla. Although hemichordates have in the past been considered to possess homologues of the chordate notochord and dorsal nerve cord, these similarities do not stand up to close scrutiny on the basis of ultrastructure or shared expression of orthologous patterning genes.

We have also shown compelling evidence that Xenoturbella should be positioned within the deuterostomes as a sister group to Ambulacraria. This result has important consequences for our understanding of deuterostome evolution. First, it means that Xenoturbella is not a neotenic member of the hemichordates but an independent lineage—Xenoturbellida—which constitutes a newly recognized deuterostome phylum. We propose the name Xenambulacraria for the highly supported monophyletic grouping of Xenoturbellida plus Ambulacraria.

The second implication of this result is that characteristics shared by the hemichordates and Xenoturbella can be parsimoniously reconstructed as having been present in the last common ancestor of Xenambulacraria. Characteristics of this ancestor must have included those characteristics shared by all deuterostomes (deuterostomy, radial cleavage, enterocoely); those shared by hemichordates and chordates (gill slits and endostyle); and the basiepithelial nervous system and unusual ciliary structure shared by the hemichordates and Xenoturbella. Evidence for most of these characteristics remains to be found in the little-studied Xenoturbella.

Although our results, which support a monophyletic chordate clade, conform to a reasonable view of unique chordate characters, we are, nevertheless, returned to a state of relative ignorance concerning the characteristics and evolution of the earliest deuterostomes. Perhaps the biggest question is whether the ancestral deuterostome nervous system was centralized, as in chordates, or diffuse, as in Xenambulacraria. The central nervous system that is prevalent in the protostomes—the closest sister group of the deuterostomes—suggests the former, but a diffuse nervous system is found in the cidarians and acoelomorph flatworms, which are widely thought to be the sister group of the protostomes plus deuterostomes. Further progress in understanding the origins and evolution of deuterostome and chordate characteristics is likely to come from the integration of palaeontological, comparative embryological and genomic approaches.

### METHODS

**Data set assembly.** Representative sequences from the data set of Philippe et al. were searched against the est others subset of GenBank, supplemented with EST sequences from Saccoglossus (206,715 sequences), Solaster (5,614 sequences) and Xenoturbella (1,372 sequences) using tblastn. Hits from our EST data sets plus those from the lamprey Petromyzon marinus, hagfish Eptatretus burgeri and the tunicates MolgulaLECT1111 and Oikopleura dioica were then assembled and added to these alignments. Additional data sets were assembled using a starting set of Xenoturbella proteins that had an orthologue in Homo, Drosophila and Caenorhabditis (see Supplementary Information). The Xenoturbella mitochondrial genome was sequenced as described in the Supplementary Information. Additional mitochondrial genomes came from the OGRE website (drake.physics.mcmaster.ca/ogre/index.shtml).

**Phylogenetic analyses.** The concatenated amino-acid alignments were analysed using bayesian and maximum-likelihood approaches and maximum-likelihood
non-parametric bootstrapping. Likelihood-based tests of alternative topologies used SH tests and Bayes factors (see Supplementary Information).

Received 8 June; accepted 11 September 2006.

Published online 18 October 2006.


Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements R.R.C., T.J., M.J.T. and S.J.B. are supported by the EU Marie Curie RTN Zoonet. S.J.B. is also funded by the BBSRC. H.N. was supported by the Human Frontier Science Program Long-Term Fellowship. L.L.M. was supported by NIH and NSF grants as well as the McKnight Brain Research Foundation and the University of Florida opportunity funds.

Author Contributions S.J.B. sequenced the mitochondrial genome. C.J.L., R.F., J.A., M.K. and E.S.L. generated the EST data. M.J.T. analysed the data and led the write-up.

Author Information The Xenoturbella mitochondrial genome sequence has GenBank Accession number DQ832701. Xenoturbella ESTs used in this project have accession numbers EE111315–EE122968. Novel Solaster ESTs used in this project have accession numbers EE122969–EE123339. Alignments used are available for download as Supplementary Information or on request from M.J.T. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.J.T. (m.telford@ucl.ac.uk).