Report

Current Biology

A Transcriptomic-Phylogenomic Analysis of the Evolutionary Relationships of Flatworms

Highlights

- Phylogenomics provide insights into the interrelationships of Platyhelminthes
- Macrostomorpha are the basalmost rhabditophorans
- Polycladida are sister group of Lecithoepitheliata/Prorhynchida
- Bothrioplanida are the free-living sister group of Neodermata

Authors

Bernhard Egger, François Lapraz, ..., Carolina Noreña, Maximilian J. Telford

Correspondence

m.telford@ucl.ac.uk

In Brief

The interrelationships of the flatworms (phylum Platyhelminthes) are poorly resolved. Egger et al. assembled a phylogenomic dataset of >107,000 aligned amino acids with less than 28% missing data from 27 flatworm taxa in 11 orders covering all major clades and reconstruct a well-resolved tree with high confidence, revealing several unexpected clades.

Accession Numbers

SRX871300	SRX872314
SRX871445	SRX883021
SRX872404	SRX872398
SRX871533	SRX872347
SRX872402	SRX872356
SRX872327	SRX872362
SRX872365	SRX872414
SRX871508	SRX872416
SRX872321	SRX879690
SRX875739	SRX872410
SRX875742	SRX874324
SRX872403	SRX875881

Egger et al., 2015, Current Biology *25*, 1347–1353
May 18, 2015
http://dx.doi.org/10.1016/j.cub.2015.03.034

CellPress

Current Biology
Report

A Transcriptomic-Phylogenomic Analysis of the Evolutionary Relationships of Flatworms

Bernhard Egger,[1,2,10] François Lapraz,[1,10,11] Bartłomiej Tomiczek,[1,10] Steven Müller,[1] Christophe Dessimoz,[1,3] Johannes Girstmair,[1] Nives Škunca,[4] Kate A. Rawlinson,[1] Christopher B. Cameron,[5] Elena Beli,[5,6] M. Antonio Todaro,[7] Mehrez Gammoudi,[8] Carolina Noreña,[9] and Maximilian J. Telford[1,*]

[1]Department Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, UK
[2]Institute of Zoology, University of Innsbruck, Technikerstrasse 25, 6020 Innsbruck, Austria
[3]Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK
[4]ETH Zurich, Department of Computer Science, Universitätsstrasse 19, 8092 Zurich, Switzerland
[5]Université de Montréal, Département de Sciences Biologiques, Pavillon Marie-Victorin, CP 6128, Succ. Centre-ville, Montréal, QC H3C 3J7, Canada
[6]Dipartimento di Scienze e Tecnologie Biologiche ed Ambientali, Università del Salento, 73100 Lecce, Italy
[7]Università degli Studi di Modena e Reggio Emilia, Via Campi 213/d, 41100 Modena, Italy
[8]Université Tunis El-Manar Campus Universitaire, 2092 Tunis, Tunisia
[9]Museo Nacional de Ciencias Naturales (CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain
[10]Co-first author
[11]Present address: CNRS, CBD UMR5547, Université de Toulouse, UPS, Centre de Biologie du Développement, Bâtiment 4R3, 118 Route de Narbonne, 31062 Toulouse, France
*Correspondence: m.telford@ucl.ac.uk
http://dx.doi.org/10.1016/j.cub.2015.03.034

SUMMARY

The interrelationships of the flatworms (phylum Platyhelminthes) are poorly resolved despite decades of morphological and molecular phylogenetic studies [1, 2]. The earliest-branching clades (Catenulida, Macrostomorpha, and Polycladida) share spiral cleavage and entolecithal eggs with other lophotrochozoans. Lecithoepitheliata have primitive spiral cleavage but derived ectolecithal eggs. Other orders (Rhabdocoela, Proseriata, Tricladida and relatives, and Bothrioplanida) all have derived ectolecithal eggs but have uncertain affinities to one another. The orders of parasitic Neodermata emerge from an uncertain position from within these ectolecithal classes. To tackle these problems, we have sequenced transcriptomes from 18 flatworms and 5 other metazoan groups. The addition of published data produces an alignment of >107,000 amino acids with less than 28% missing data from 27 flatworm taxa in 11 orders covering all major clades. Our phylogenetic analyses show that Platyhelminthes consist of the two clades Catenulida and Rhabditophora. Within Rhabditophora, we show the earliest-emerging branch is Macrostomorpha, not Polycladida. We show Lecithoepitheliata are not members of Neoophora but are sister group of Polycladida, implying independent origins of the ectolecithal eggs found in Lecithoepitheliata and Neoophora. We resolve Rhabdocoela as the most basally branching euneoophoran taxon. Tricladida, Bothrioplanida, and Neodermata constitute a group that appears to have lost both spiral cleavage and centrosomes. We identify Bothrioplanida as the long-sought closest free-living sister group of the parasitic Neodermata. Among parasitic orders, we show that Cestoda are closer to Trematoda than to Monogenea, rejecting the concept of the Cercomeromorpha. Our results have important implications for understanding the evolution of this major phylum.

RESULTS AND DISCUSSION

We assembled coding sequence data from 55 animal species, including 27 species of platyhelminth. We identified 1,348 orthologous genes and produced a large (>107,000 positions) and taxonomically broad phylogenomic dataset (27 flatworm species from 11 orders) for the analysis of the phylogeny of this important and diverse group of animals. The dataset contains very few missing data (average 72% complete, measured as the percentage of positions with data present within the total alignment), especially in the case of the newly sequenced taxa (average 82% complete, all but two >68% complete). We used site-heterogeneous Bayesian tree reconstruction (PhyloBayes CAT+GTR+G4 [3] model, which has site-specific equilibrium frequency profiles; Figure 1) and site-homogenous maximum-likelihood (ML) approaches (PhyML LG+G4 [4] and RAxML CATGTR [5], which have homogenous equilibrium frequency profiles; Figures S1 and S2) to reconstruct the phylogeny based on these data. Most relationships within Platyhelminthes are robustly resolved as shown by concordance between different analyses, Bayesian posterior probabilities (Figure 1), jackknife resampling (Figure 2), and phylogenetic signal dissection (Figure 3).

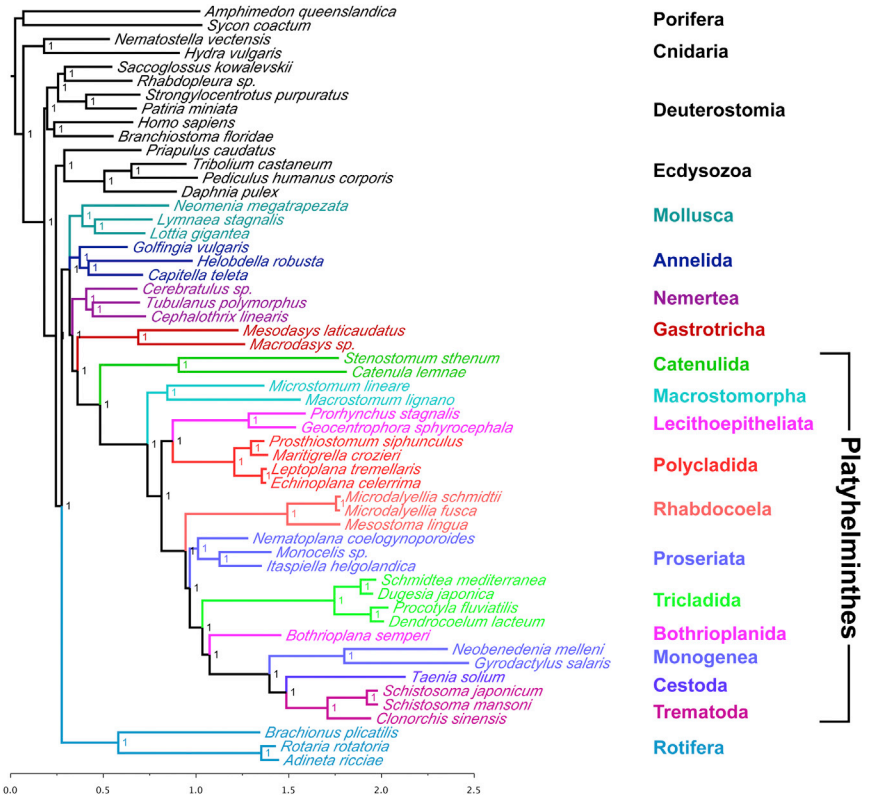


Figure 1. Phylogeny Produced Using PhyloBayes with the Site-Heterogeneous CAT+GTR+G4 Model on the Full 107,659 Amino Acid Alignment

There is support for a sister group relationship between Gastrotricha and Platyhelminthes, which are members of an unresolved clade including mollusks, annelids, and nemerteans, contrary to the concept of the Platyzoa. Platyhelminthes are monophyletic. Macrostomorpha is the earliest-branching rhabditophoran clade. Lecithopepitheliata and Polycladida are sister groups. Rhabdocoels are the sister clade to all other neoophoran orders, including proseriates, but are separated from other Euneoophora by a very short internode. *Bothrioplana* is the closest free-living relative of the parasitic Neodermata. Values at nodes indicate posterior probabilities. Scale bar indicates number of substitutions per site. MaxDiff = 1.0; MeanDiff = 0.00934579. Lophotrochozoan groups in Figures 1, 2, and 3 are indicated by colored labels.

Platyhelminthes Are a Monophyletic Group of Lophotrochozoans

Our tree supports the now canonical view of Platyhelminthes as members of Lophotrochozoa, which was first shown using 18S rDNA data [6] and has subsequently received strong support from multigene phylogenies (e.g., [7]). Of perhaps greater interest is the finding of a strongly supported sister group relationship between the two species representing the order Catenulida and the remaining Platyhelminthes: the Rhabditophora (Figures 1, 2, 3, and 4). Rhabditophora share the convincing molecular synapomorphy of two changes in mitochondrial genetic code [8], and we provide phylogenomic confirmation of the monophyly of Platyhelminthes (Catenulida+Rhabditophora). Surprisingly, a convincing phenotypic synapomorphy of Platyhelminthes is still lacking [9, 10]. We have not considered the xenacoelomorphs, originally part of Platyhelminthes, as they have been shown by various means not to be part of the protostomes [11].

Support for Platyzoa May Derive from a Long-Branch Attraction Artifact

While our ML tree supports Platyzoa ((Platyhelminthes, Gnathifera)(Gastrotricha)) [12–14] (Figure S1), with the rotiferans representing the larger group of Gnathifera, our Bayesian analyses, in common with two recent well-sampled phylogenomic studies of lophotrochozoan relationships [7, 11], show largely consistent support for Gastrotricha and Platyhelminthes being grouped together with Nemertea, Annelida, and Mollusca; Rotifera are outside of this clade (Figure 1). In our CAT+GTR+G4 analysis, Nemertea are sister group of Platyhelminthes+Gastrotricha. The exception to this finding is in our jackknife analysis, where the position of gastrotrichs relative to platyhelminths and other lophotrochozoan phyla is unresolved (Figure 2). In the signal dissection experiment, the fastest-evolving quartile of the data (most susceptible to long-branch attraction [LBA] or LBA artifact) supports Platyzoa (Figure 3, Q4), and this gives credence to the view of Platyzoa as arising from such a systematic error. Adopting measures to counter this problem with selected slowly evolving genes and well-fitted models (CAT+GTR+G4) rejects Platyzoa (Figure 3, Q1 and Q2).

A Biflagellate Sperm Unites All Rhabditophora except Macrostomorpha

To date, the identity of the basalmost branching group of Rhabditophora has not been settled, with Macrostomorpha and Polycladida vying for this position [2, 7, 14–16]. Members of both of these groups possess the likely primitive character of spiral cleavage (absent in many more derived groups, see Figure 4) and also have entolecithal eggs, again a likely primitive character. Polycladida have a larval stage (present in both major clades of polyclads) that some consider homologous to the trochophore seen in several other lophotrochozoan phyla [17]. Macrostomorpha have aflagellate sperm; this contrasts with the remaining Rhabditophora, including Polycladida, which typically have a biflagellate sperm with a 9 × 2 + "1" pattern of microtubules and on this basis have been grouped as Trepaxonemata [1, 18]. Our data strongly support Macrostomorpha as sister group of all other rhabditophoran orders. Macrostomorpha are excluded from the monophyletic Trepaxonemata with posterior probability of 1.0 (Figure 1), jackknife support of 1.0 (Figure 3), and PhyML "SH-like" support [4] of 1.0 (Figure S1) as well as being found with pp = 1.0 in all four quartiles of the signal dissection experiment (Figure 3).

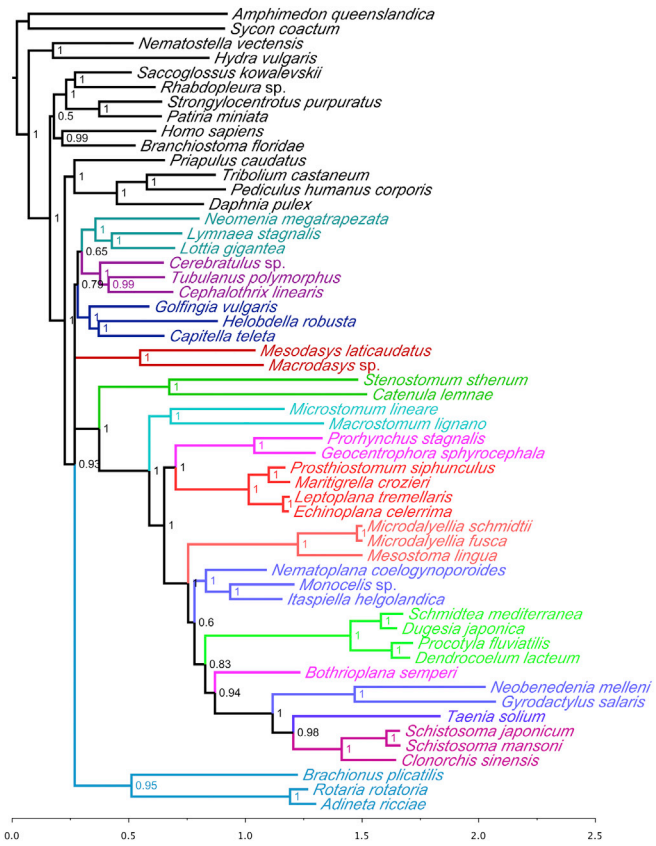


Figure 2. Jackknife Analysis of 100 Datasets of 20,000 Amino Acids Each, Produced Using the PhyloBayes CAT+GTR+G4 Model

Values at nodes indicate proportion of replicates in which the node is found (1 corresponds to 100% jackknife). The topology is largely the same as the full analysis shown in Figure 1, and most clades receive high support. Relatively low support for the sister group relationship of rhabdocoels and other euneoophorans is observed. There is no clear support for or against Platyzoa, indicated by the polytomy at the base of the Lophotrochozoa. Scale bar indicates number of substitutions per site.

Independent Evolution of Ectolecithal Eggs in Lecithoepitheliata and Euneoophora

Apart from Macrostomorpha and Polycladida, all rhabditophoran groups, including Lecithoepitheliata, are distinguished by ectolecithal eggs (yolk not incorporated into the embryonic blastomeres) and the associated characteristic (absent in Lecithoepitheliata) of an ovary structured into separate germary and vitellary areas. This assemblage of Rhabditophora with ectolecithal eggs is generally considered to constitute a clade called Neoophora [1, 19]. Lecithoepitheliata have been reconstructed as sister group of other Neoophora based on morphological characters [1] and limited marker molecular data [2, 20]. Lecithoepitheliata are split into freshwater-dwelling Prorhynchida and marine Gnosonesimida and may in fact be para- or polyphyletic [1, 2, 21]. In the only molecular study involving members of both taxa, they are presented as being grouped with other ectolecithal Platyhelminthes (i.e., members of the Neoophora), but Prorhynchida were found to be sister group of all other Neoophora, and Gnosonesimida as sister group of all other Neoophora except Prorhynchida [2]. This topology led these authors to support the monophyly of Neoophora and the single origin of ectolecithality.

Our study includes two members of Prorhynchida and, in striking contrast to most previous studies, places them not in Neoophora but as sister group of Polycladida, in accordance with [16]. The monophyly of Polycladida and Lecithoepitheliata/Prorhynchida is given maximum support in all analyses (Figures 1, 2, 3, S1, and S2).

This result has important implications for our understanding of the evolution of ectolecithality within flatworms: eggs with extraembryonic yolk cells would have evolved at least twice independently, once in Lecithoepitheliata (at least Prorhynchida) and once in the common ancestor of the remaining Neoophora. Neoophora excluding Lecithoepitheliata were named Euneoophora [2]. In further support of the monophyly of Euneoophora, we found that the parahox gene *Caudal/Cdx* was detectable in the transcriptomes of the lophotrochozoan phyla we sampled and also in the different orders of archoophorans, i.e., catenulids, macrostomorphans, polyclads, and lecithoepitheliates, but was undetectable in the transcriptomes of all included euneoophorans (Figure 4).

Rhabdocoela, Not Proseriata, Are Likely to Be the Basalmost Euneoophoran Clade

The least confidently resolved part of the flatworm portion of our tree involves the relative positions of proseriates, rhabdocoels, and the remaining euneoophorans. In our CAT+GTR+G4 phylogeny of our complete dataset, Rhabdocoela are sister group of all other Euneoophora (pp = 1.0) (Figure 1), but with low jackknife support of 0.6 (Figure 2). Other analyses (ML) instead support Proseriata in this position (Figures S1 and S2), and this is in common with most previous analyses involving one or a few genes [2, 16, 20, 22–25].

On balance, we suggest that the basal Rhabdocoela solution is the most likely for two reasons. The first reason is the support it receives from the typically better-performing CAT+GTR+G4 model analysis over the PhyML analysis. The site-heterogeneous CAT model has been repeatedly shown to fit real data better than simpler models such as the site-homogenous model used in the ML analyses, and to be better able to overcome systematic error [11]. The second reason is that we observed stronger support for basal Rhabdocoela when analyzing the slowly evolving genes (Q1 and Q2); the more rapidly evolving genes support an association of Rhabdocoela and Tricladida/Bothrioplanida/Neodermata (Q3) or Rhabdocoela and Tricladida/Neodermata (Q4) (Figure 3). The support that this particular

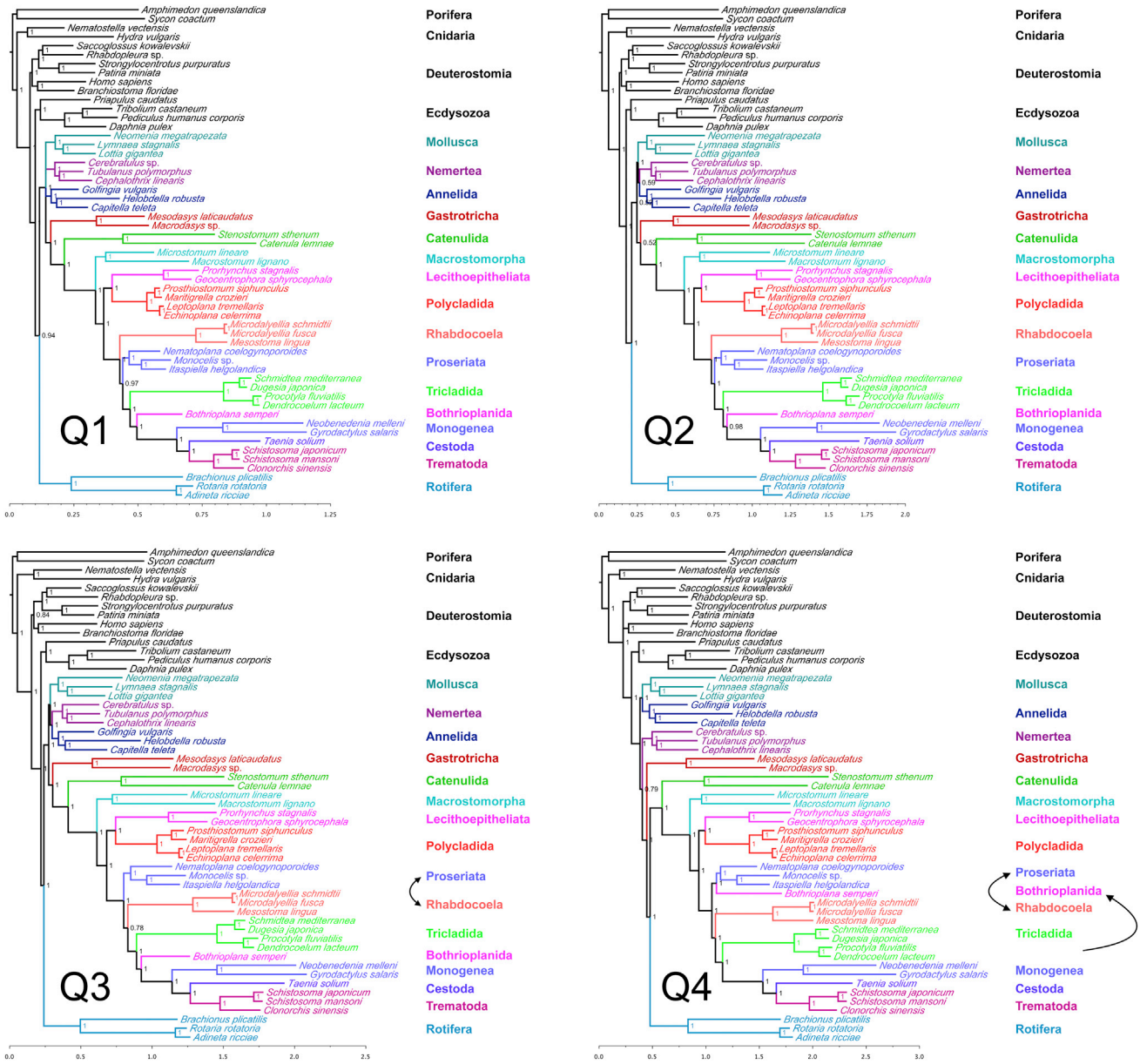


Figure 3. Phylogenetic Signal Dissection: Gene Rate Ranking to Look for Possible LBA Artifacts

The trees shown were produced using PhyloBayes' CAT+GTR+G4 model on four equal-sized datasets (quartiles Q1 to Q4) containing genes evolving at increasingly rapid rates (Figure S3). Q1 is slowest and expected to be least susceptible to long-branch attraction (LBA); Q4 is fastest evolving and, a priori, most susceptible to LBA. The trees of the slowest two quartiles are identical in all important respects to the topology found using the full dataset. In the faster-evolving quartiles, the positions of the long-branched rhabdocoels and short-branched proseriates are reversed. In Q4, the short-branched *Bothrioplana* groups with short-branched Proseriata and the long-branched Rhabdocoela and Neodermata are grouped together. In the slower-evolving Q1–Q3, no support for Platyzoa is observed. In Q4, support switches to Platyzoa (Rotifera, Gastrotricha, and Platyhelminthes), presumably due to LBA effects. Relative substitution rates: Q1 = 1.14, Q2 = 1.33, Q3 = 1.42, Q4 = 1.54. Percent missing data: Q1 = 27%, Q2 = 26%, Q3 = 28%, Q4 = 29%. MaxDiff/MeanDiff: Q1 = 1.0/0.0192412, Q2 = 0.928747/0.02347, Q3 = 0.425926/0.00683628, Q4 = 0.647856/0.00764029. Scale bars indicate number of substitutions per site.

grouping receives in the analyses of more rapidly evolving genes seems likely to be due to an LBA artifact that leads to an incorrect association between the rhabdocoels and neodermatans, both of which have long branches. LBA is exacerbated by rapidly evolving genes [26]. In the fastest quartile of data (Q4; a priori most susceptible to LBA) the long-branched rhabdocoels move even closer to the long-branched Tricladida/Neodermata than the short-branched *Bothrioplana* (Figure 3).

Loss of Centrosomes Defines a Group Including Planarians and Parasites

More reliably resolved is the position of Tricladida, which has strong support for a position closer to Neodermata and

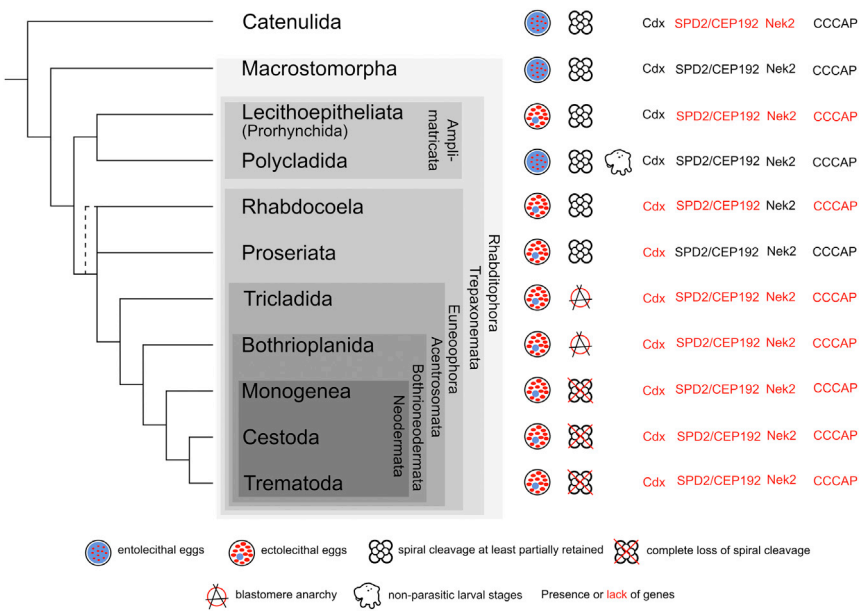


Figure 4. Consensus Tree of Relationships of Eleven Platyhelminth Orders with Important Morphological and Genetic Characters Mapped

The less-reliably resolved branches involve the Rhabdocoela and the Proseriata, although our results suggest that basally branching Rhabdocoela (indicated by dashed line) is likely the correct solution. Developmental features, such as egg and cleavage type, planktotrophic larvae, and gene presence/absence patterns are indicated to the right of the tree. If polyclad larvae are homologous with the trochophores of annelids and mollusks, primary larval stages must have been lost in Catenulida, Macrostomorpha, Lecithoepitheliata, and the Euneoophora. Entolecithal eggs as found in Catenulida, Macrostomorpha, and Polycladida are an ancestral character. Ectolecithal eggs are independently present in Lecithoepitheliata and Euneoophora. The parahox gene *Cdx* is undetectable in all Euneoophora. Spiral cleavage has been lost in Acentrosomata, and the three centrosome-associated genes shown are undetectable in this group.

Bothrioplanida than either Rhabdocoela or Proseriata. None of the triclads, bothrioplanids, or neodermatans show any sign of spiral cleavage in their early embryogenesis, and the loss of this trait is a persuasive morphological character uniting this group (Figure 4). Recent studies have noted that genes including *SPD-2/Cep192*, *Nek2*, and *CCCAP*, which have an evolutionarily conserved role in centrosome formation across Metazoa, were missing from the planarian *Schmidtea mediterranea* as well as from the neodermatan *Schistosoma mansoni* yet were present in the macrostomorphan *Macrostomum lignano* [27]. This gene loss correlates with the loss of the centrosome in *Schmidtea* and possibly also in *Schistosoma*, and it was suggested that this loss of centrosomal genes is also implicated in the loss of the highly regulated spiral cleavage [27]. Thanks to our taxonomically broad sample of transcriptomes, we have been able to extend this analysis and show that three genes associated with centrioles, *SPD-2/Cep192*, *Nek2*, and *CCCAP*, are at least partly present in most of the more basally branching platyhelminth taxa for which we have transcriptomes but are undetectable in any of the Tricladida, Bothrioplanida, or Neodermata (Figure 4). The evidence for absence of a gene based on inevitably partial transcriptomes must not be overinterpreted, however, and we note that none of these three genes are found in the transcriptomes that we have produced for two lecithoepitheliates (Figure 4), which show a rather conserved spiral cleavage pattern [19].

Identifying the Free-Living Ancestor of the Parasitic Neodermata

The monophyly of Neodermata with well-characterized apomorphies such as a secondary unciliated syncytial epidermis is undisputed [1]. It has long been clear that "Turbellaria" is a paraphyletic group and that the wholly parasitic Neodermata emerged from among free-living forms [1]. That said, the identity of the closest free-living relative of Neodermata has proven elusive. In early morphological phylogenies, Rhabdocoela (or members of Rhabdocoela) were considered to be sister group of Neodermata [1, 9, 28]. This relationship was not supported in subsequent molecular phylogenies using one or a few genes, in which a bewildering selection of higher flatworm taxa, e.g., Fecampiida, Prolecithophora, and Tricladida [29] or Rhabdocoela, Fecampiida, Prolecithophora, and Tricladida [16, 20, 23–25], were proposed as sister group of Neodermata. In a recent study using four genes and many taxa [2], and now in our own study using 1,347 genes, Bothrioplanida, previously considered close to or part of the Proseriata [1], are shown to be sister group of Neodermata (Figure 1).

Relationships among the Neodermatan Groups

Neodermata comprise Monogenea, Cestoda, and Trematoda [1]. The interrelationships of these taxa has been debated, with Cestoda being considered sister group of either Monogenea (= Cercomeromorpha) or Trematoda. The Cercomeromorpha hypothesis was rejected by phylogenetic analyses using 18S and 28S sequences [16, 25], and the alternative sister group relationship between Trematoda and Cestoda was supported by studies employing whole mitochondrial gene phylogenies [30, 31], by a microRNA study [32], by a multigene phylogeny using 312 gene models [33], and now by our own study. Surprisingly, a recent phylogenetic study using four genes and a large number of flatworm species supports the Cercomeromorpha hypothesis [2].

Old and New Systematic Names

With the sister group relationship between Polycladida and Lecithoepitheliata/Prorhynchida demonstrated by our phylogenetic analysis (Figure 1), the taxon Neoophora, defined as encompassing all flatworms with ectolecithal eggs [1], has become polyphyletic and should therefore be noted with quotation marks, "Neoophora." "Neoophora," excluding Prorhynchida, are monophyletic in our analyses, and this clade has previously been named Euneoophora, characterized by the presence of

ectolecithal eggs and by germaria and vitellaria as spatially separated organs [2].

We propose the name Amplimatricata *new taxon* for Polycladida+Lecithoepitheliata, based on the tendency in both groups for possession of an ample extracellular matrix [34]. Taking into account the remaining uncertainty over the monophyly of Lecithoepitheliata [2], Amplimatricata encompasses at least Polycladida+Prorhynchida. Acentrosomata *new taxon* is a clade consisting of Tricladida and its closely related taxa Prolecithophora and Fecampiida (all three taxa making up Adiaphanida [35]), Bothrioplanida, and Neodermata (Figure 4). The name is based on the implied absence of centrosomes in all of these taxa (Figure 4). Lacking strong similarities to serve as a clade-defining synapomorphy between Bothrioplanida and Neodermata, we use the name Bothrioneodermata *new taxon* to identify this monophyletic group (Figure 4).

Conclusions

We have presented new transcriptomic data from 22 new species and produced a large and taxonomically complete dataset for assessing the relationships of Platyhelminthes. The majority of our conclusions are robust and are supported by different methods of analysis, high Bayesian posterior probabilities, and high jackknife support. The two instances of lower support concern the position of Platyhelminthes relative to other lophotrochozoan phyla and the early-branching position of rhabdocoels relative to other Euneoophora. The evidence against Platyzoa and support for early-branching Rhabdocoela by site-heterogeneous analyses (Figure 1) and by the slowest-evolving quartiles of the total dataset (Figure 3, Q1 and Q2) suggest that the alternatives, which are supported by the less-well-fitting site-homogenous analyses (Figures S1 and S2) and the faster-evolving quartiles of the data (Figure 3, Q3 and Q4), are the result of LBA. The suggested monophyly of Proseriata+Acentrosomata (Figure 4) might be tested further by the addition of the two additional members of Adiaphanida, Fecampiida and Prolecithophora, as well as the second clade of Lecithoepitheliata, Gnosonesimida, as there has been evidence that Lecithoepitheliata may be paraphyletic [1, 2].

ACCESSION NUMBERS

The NCBI Sequence Read Archive experiment accession numbers for the data reported in this paper are SRX871300, SRX871445, SRX872404, SRX871533, SRX872402, SRX872327, SRX872365, SRX871508, SRX872321, SRX875739, SRX875742, SRX872403, SRX872314, SRX883021, SRX872398, SRX872347, SRX872356, SRX872362, SRX872414, SRX872416, SRX879690, SRX872410, SRX874324, and SRX875881.

SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures, one table, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2015.03.034.

AUTHOR CONTRIBUTIONS

B.E., C.B.C., C.N., E.B., F.L., K.A.R., J.G., M.A.T., and M.G. collected samples for RNA extraction. B.E., F.L., J.G., and K.A.R. extracted RNA. B.E. and M.J.T. assembled transcriptomes. B.T., C.D., N.S., and S.M. produced the OMA analyses. M.J.T. assembled the sets of orthologs and conducted phylogenetic analyses. B.E. and M.J.T. wrote the manuscript and designed the figures. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Claus Nielsen for help identifying *Cerebratulus* and *Golfingia*, and we thank Katrien De Mulder, Lena Egger, Angie Opitz, Petra Derler, Helena Derler, and Kiyono Sekii for assistance with sampling. This work was supported by a Biotechnology and Biological Sciences Research Council grant (BB/H006966/1) (F.L.), a Leverhulme Trust grant (F/07 134/DA) (B.E.), the European Research Council (ERC-2012-AdG 322790) (S.M.), Marie Curie ITN NEPTUNE (J.G.), and the Research Focal Point Scientific Computing at the University of Innsbruck. M.J.T. is supported by a Royal Society Wolfson Research Merit Award.

Received: January 20, 2015
Revised: February 27, 2015
Accepted: March 19, 2015
Published: April 9, 2015

REFERENCES

1. Ehlers, U. (1985). Das Phylogenetische System der Plathelminthes, First Edition. (Gustav Fischer Verlag).
2. Laumer, C.E., and Giribet, G. (2014). Inclusive taxon sampling suggests a single, stepwise origin of ectolecithality in Platyhelminthes. Biol. J. Linn. Soc. Lond. *111*, 570–588.
3. Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics *25*, 2286–2288.
4. Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.
5. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.
6. Aguinaldo, A.M.A., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A., and Lake, J.A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. Nature *387*, 489–493.
7. Struck, T.H., Wey-Fabrizius, A.R., Golombek, A., Hering, L., Weigert, A., Bleidorn, C., Klebow, S., Iakovenko, N., Hausdorf, B., Petersen, M., *et al*. (2014). Platyzoan paraphyly based on phylogenomic data supports a noncoelomate ancestry of spiralia. Mol. Biol. Evol. *31*, 1833–1849.
8. Telford, M.J., Herniou, E.A., Russell, R.B., and Littlewood, D.T.J. (2000). Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. Proc. Natl. Acad. Sci. USA *97*, 11359–11364.
9. Smith, J.P.S., Tyler, S., and Rieger, R.M. (1986). Is the Turbellaria polyphyletic? Hydrobiologia *132*, 13–21.
10. Egger, B., Steinke, D., Tarui, H., De Mulder, K., Arendt, D., Borgonie, G., Funayama, N., Gschwentner, R., Hartenstein, V., Hobmayer, B., *et al*. (2009). To be or not to be a flatworm: the acoel controversy. PLoS ONE *4*, e5502.
11. Philippe, H., Brinkmann, H., Copley, R.R., Moroz, L.L., Nakano, H., Poustka, A.J., Wallberg, A., Peterson, K.J., and Telford, M.J. (2011). Acoelomorph flatworms are deuterostomes related to *Xenoturbella*. Nature *470*, 255–258.
12. Cavalier-Smith, T. (1998). A revised six-kingdom system of life. Biol. Rev. Camb. Philos. Soc. *73*, 203–266.
13. Hejnol, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G.W., Edgecombe, G.D., Martinez, P., Baguñà, J., Bailly, X., Jondelius, U., *et al*. (2009). Assessing the root of bilaterian animals with scalable phylogenomic methods. Proc. Biol. Sci. *276*, 4261–4270.
14. Wey-Fabrizius, A.R., Herlyn, H., Rieger, B., Rosenkranz, D., Witek, A., Welch, D.B.M., Ebersberger, I., and Hankeln, T. (2014). Transcriptome

data reveal syndermatan relationships and suggest the evolution of endoparasitism in Acanthocephala via an epizoic stage. PLoS ONE *9*, e88618.

15. Mallatt, J., Craig, C.W., and Yoder, M.J. (2010). Nearly complete rRNA genes assembled from across the metazoan animals: effects of more taxa, a structure-based alignment, and paired-sites evolutionary models on phylogeny reconstruction. Mol. Phylogenet. Evol. *55*, 1–17.
16. Mallatt, J., Craig, C.W., and Yoder, M.J. (2012). Nearly complete rRNA genes from 371 Animalia: updated structure-based alignment and detailed phylogenetic analysis. Mol. Phylogenet. Evol. *64*, 603–617.
17. Nielsen, C. (2012). Animal Evolution, Third Edition. (Oxford University Press).
18. Justine, J.-L. (2001). Spermatozoa as phylogenetic characters for the Platyhelminthes. In Interrelations of the Platyhelminthes (The Systematics Association Special Volume Series 60), D.T.J. Littlewood and R.A. Bray, eds. (Taylor & Francis), pp. 231–238.
19. Martín-Durán, J.M., and Egger, B. (2012). Developmental diversity in free-living flatworms. Evodevo *3*, 7.
20. Joffe, B.I., and Kornakova, E.E. (2001). Flatworm phylogeneticist: Between molecular hammer and morphological anvil. In Interrelations of the Platyhelminthes (The Systematics Association Special Volume Series 60), D.T.J. Littlewood and R.A. Bray, eds. (Taylor & Francis), pp. 279–291.
21. Karling, T.G. (1968). On the genus *Gnosonesima* Reisinger (Turbellaria). Sarsia *33*, 81–108.
22. Littlewood, D.T.J., Rohde, K., and Clough, K.A. (1999). The interrelationships of all major groups of Platyhelminthes: phylogenetic evidence from morphology and molecules. Biol. J. Linn. Soc. Lond. *66*, 75–114.
23. Baguñà, J., Carranza, S., Paps, J., Ruiz-Trillo, I., and Riutort, M. (2001). Molecular taxonomy and phylogeny of the Tricladida. In Interrelations of the Platyhelminthes (The Systematics Association Special Volume Series 60), D.T.J. Littlewood and R.A. Bray, eds. (Taylor & Francis), pp. 49–56.
24. Larsson, K., and Jondelius, U. (2008). Phylogeny of Catenulida and support for Platyhelminthes. Org. Divers. Evol. *8*, 378–387.
25. Lockyer, A.E., Olson, P.D., and Littlewood, D.T.J. (2003). Utility of complete large and small subunit rRNA genes in resolving the phylogeny of the Neodermata (Platyhelminthes): implications and a review of the cercomer theory. Biol. J. Linn. Soc. Lond. *78*, 155–171.
26. Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. *27*, 401–410.
27. Azimzadeh, J., Wong, M.L., Downhour, D.M., Sánchez Alvarado, A., and Marshall, W.F. (2012). Centrosome loss in the evolution of planarians. Science *335*, 461–463.
28. Karling, T.G. (1974). On the anatomy and affinities of the turbellarian orders. In Biology of the Turbellaria, N.W. Riser and M.P. Morse, eds. (McGraw-Hill), pp. 1–16.
29. Willems, W.R., Wallberg, A., Jondelius, U., Littlewood, D.T.J., Backeljau, T., Schockaert, E.R., and Artois, T.J. (2006). Filling a gap in the phylogeny of flatworms: relationships within the Rhabdocoela (Platyhelminthes), inferred from 18S ribosomal DNA sequences. Zool. Scr. *35*, 1–17.
30. Park, J.-K., Kim, K.-H., Kang, S., Kim, W., Eom, K.S., and Littlewood, D.T.J. (2007). A common origin of complex life cycles in parasitic flatworms: evidence from the complete mitochondrial genome of *Microcotyle sebastis* (Monogenea: Platyhelminthes). BMC Evol. Biol. *7*, 11.
31. Perkins, E.M., Donnellan, S.C., Bertozzi, T., and Whittington, I.D. (2010). Closing the mitochondrial circle on paraphyly of the Monogenea (Platyhelminthes) infers evolution in the diet of parasitic flatworms. Int. J. Parasitol. *40*, 1237–1245.
32. Fromm, B., Worren, M.M., Hahn, C., Hovig, E., and Bachmann, L. (2013). Substantial loss of conserved and gain of novel MicroRNA families in flatworms. Mol. Biol. Evol. *30*, 2619–2628.
33. Hahn, C., Fromm, B., and Bachmann, L. (2014). Comparative genomics of flatworms (Platyhelminthes) reveals shared genomic features of ecto- and endoparastic Neodermata. Genome Biol. Evol. *6*, 1105–1117.
34. Rieger, R.M., Tyler, S., Smith, J.P.S., and Rieger, G.E. (1991). Platyhelminthes: Turbellaria. In Microscopic Anatomy of Invertebrates, Volume 3: Platyhelminthes and Nemertinea, F.W. Harrison and B.J. Bogitsh, eds. (Wiley), pp. 7–140.
35. Noren, M., and Jondelius, U. (2002). The phylogenetic position of the Prolecithophora (Rhabditophora, 'Platyhelminthes'). Zool. Scr. *31*, 403–414.

Current Biology

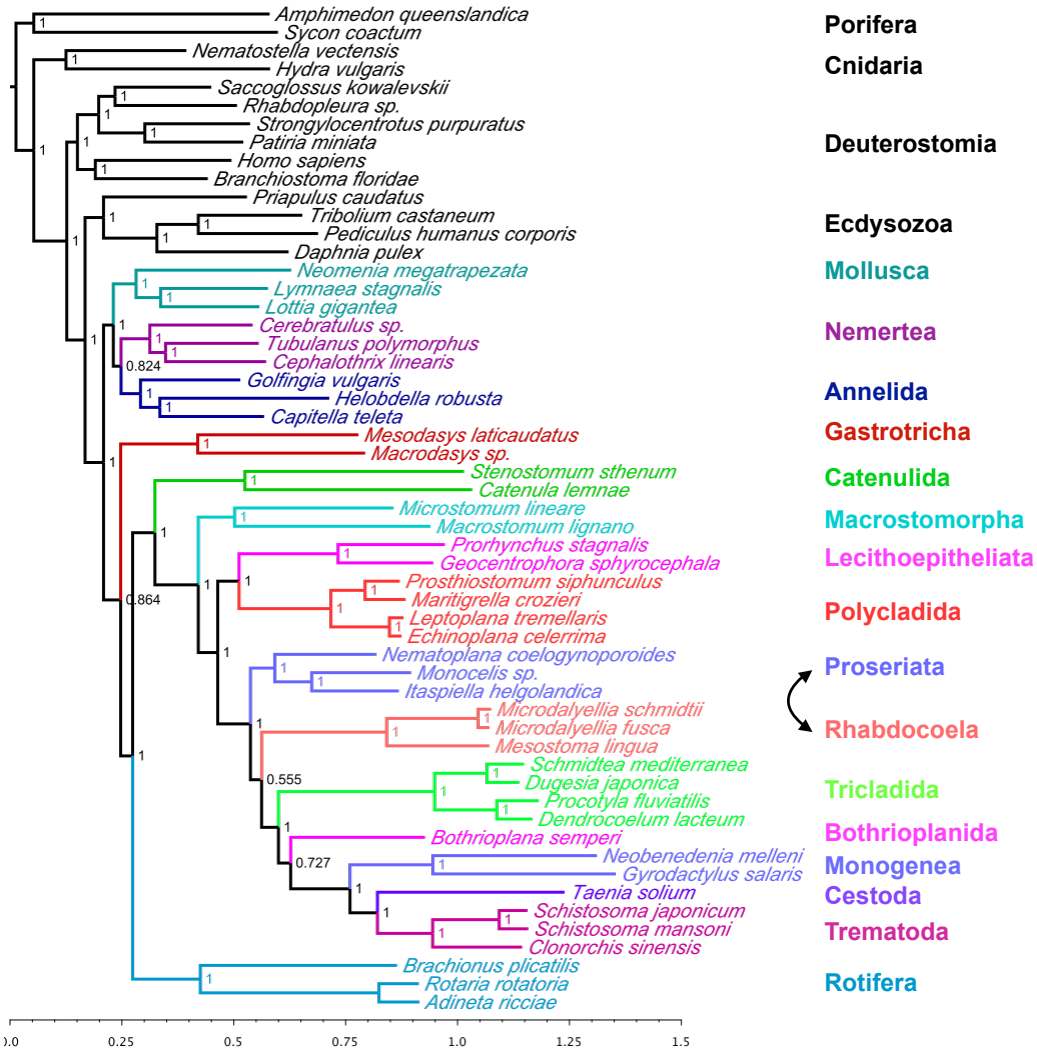
Supplemental Information

A Transcriptomic-Phylogenomic Analysis of the Evolutionary Relationships of Flatworms

Bernhard Egger, François Lapraz, Bartłomiej Tomiczek, Steven Müller, Christophe Dessimoz, Johannes Girstmair, Nives Škunca, Kate A. Rawlinson, Christopher B. Cameron, Elena Beli, M. Antonio Todaro, Mehrez Gammoudi, Carolina Noreña, and Maximilian J. Telford

Supplemental Information

Figure S1: Phylogeny produced using PhyML with the site-homogenous LG+G4 model on full 107,659 amino acids alignment. Values at nodes indicate SH-like support [4]. Major differences compared to Fig. 1 are the clade of Platyhelminthes, Gastrotricha and Rotifera (Platyzoa) and the reversed positions of Rhabdocoela and Proseriata. This may be due to



Long Branch Attraction between groups of 'platyzoans' and between Rhabdocoela and Neodermata. LBA has been shown to be more prevalent with the site-homogenous model used here. Scale bar indicates number of substitutions per site.

Figure S2: Phylogeny produced using RAxML with the site-homogenous LG+G4 model on full 107,659 amino acids alignment [5]. Major differences compared to Fig. 1 are the clade of Platyhelminthes, Gastrotricha and Rotifera (Platyzoa) and the reversed positions of Rhabdocoela and Proseriata. This may be due to Long Branch Attraction between groups of 'platyzoans' and between Rhabdocoela and Neodermata. LBA has been shown to be more prevalent with the site-homogenous model used here. Scale bar indicates number of substitutions per site.

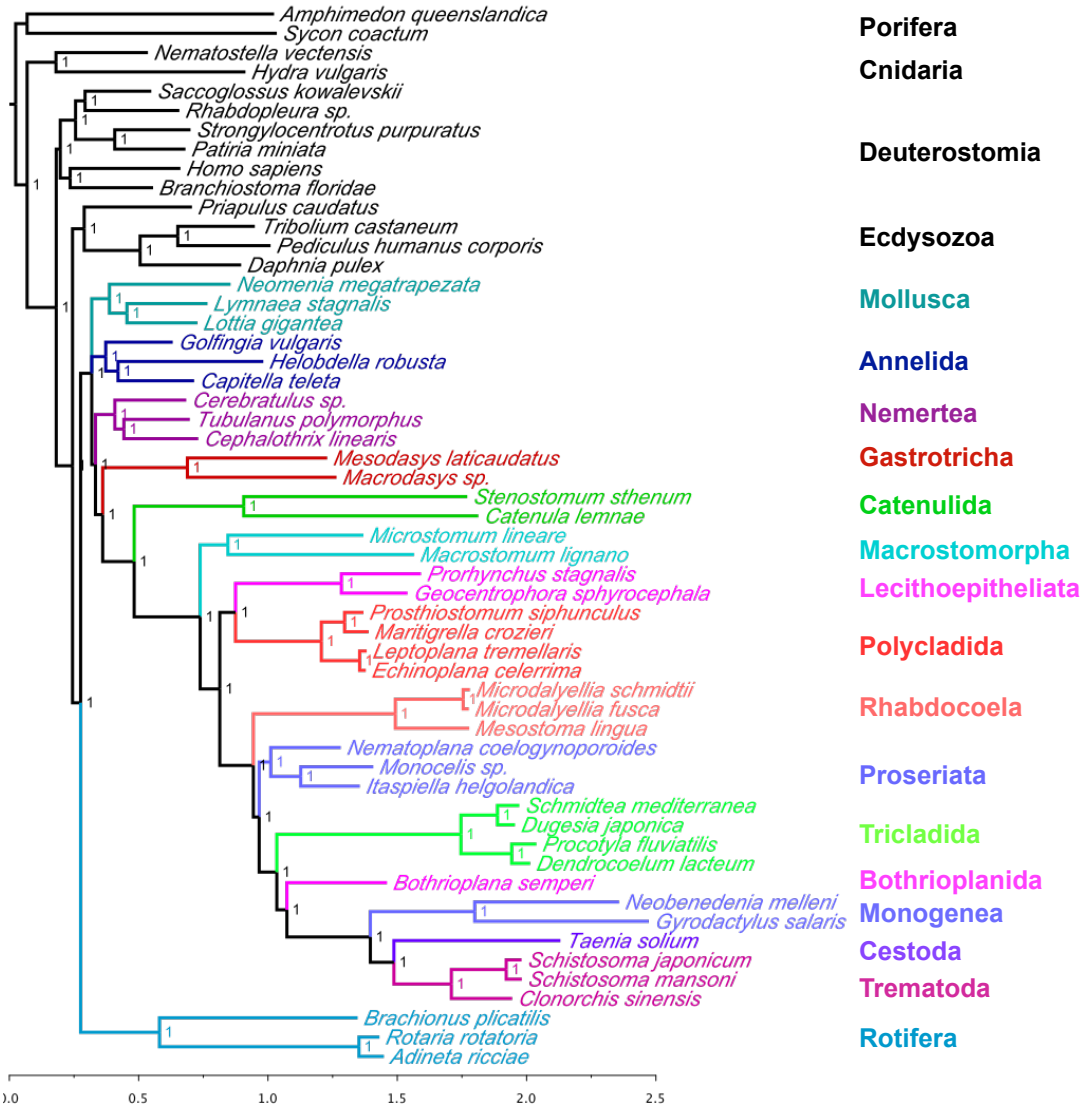
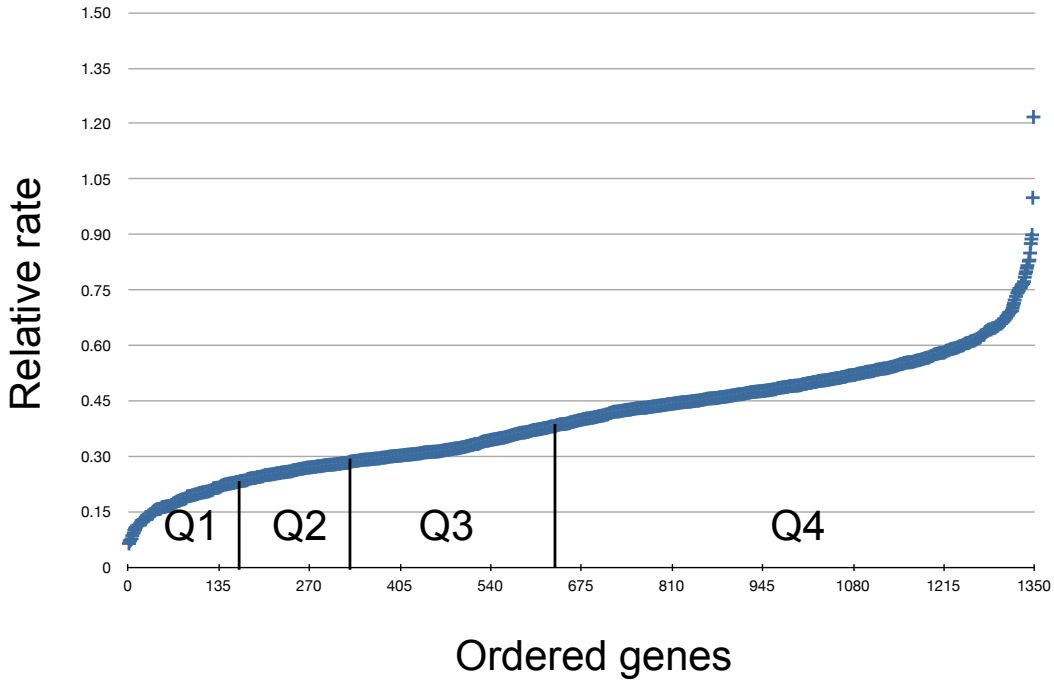


Figure S3: Graph showing the relative rates of substitution of all 1348 genes ordered by rate as used in the phylogenetic signal dissection experiment. The division of genes into four quartiles is indicated. The number of genes making up the faster evolving quartiles is larger because the final datasets were produced after deleting sites with more than a set percentage of missing data. The faster evolving genes had more missing data (or were less easy to align into regions conserved across all taxa). The final 4 datasets all contained the same number of positions.



Supplemental Experimental procedures

Specimen collection and determination

Polyclad flatworms were collected from their substrate with a soft brush and transferred into a water-filled container [S1]. For marine interstitial flatworms (*Itaspiella helgolandica*, *Monocelis* sp., *Nematoplana coelogynoporoides*) sand samples were collected into lockable plasticware containers and later extracted in the laboratory with a 1:1 solution of $MgCl_2*6H_2O$ and sea water and filtered through 40-100 µm meshes [S2]. The content of the meshes was flushed with seawater into petri dishes and animals of interest were sorted into embryo dishes under a binocular microscope. Freshwater samples – containing water plants, mud or sand – were poured into petri dishes, further diluted with water if necessary and animals of interest were searched for under a binocular microscope.

Extraction of *Mesodasys laticaudatus* (Macrodasyida, Cephalodasyidae) was done by the narcotization-decantation technique using a 7% magnesium chloride solution. Animals were allowed to recover for 2 hours in sea water.

Species determination was carried out with live animals using either whole animals under a binocular or squeeze preparations under a compound microscope [S3]. Literature used for species determination were taxonomic guides [S4-S7], monographs on particular groups [S8-S10] and specific taxonomic accounts of considered species [S11-S12], all facilitated by the excellent Turbellarian Database (http://turbellaria.umaine.edu/).

RNA extraction and sequencing

For one sample (*Prorhynchus stagnalis*, BioProject PRJNA275072) kept in RNAlater (Life Technologies), a Nucleospin RNA XS kit (Macherey-Nagel, Düren, Germany) was used. For all other samples a TRIzol Reagent (Life Technologies, Carlsbad, CA)/TRI Reagent (Sigma-Aldrich, St. Louis, MO) based RNA extraction protocol was used on live animals or dissected tissues, following manufacturers' protocols. Total RNA was stored at -80 °C until sent for sequencing (The Centre for Applied Genomics: The Hospital for Sick Children, Toronto, Canada), where cDNA libraries were prepared using Illumina TruSeq kits (Illumina Inc., San Diego, CA). For three samples with low amounts of RNA (*Catenula lemnae*, *Geocentrophora sphyrocephala* and the second *Prorhynchus stagnalis* sample, BioProject PRJNA275317), SMART mRNA amplification kits (Clontech Laboratories Inc., Mountain View, CA) followed by Nextera XT kits (Illumina) were used. The libraries were sequenced on an Illumina HiSeq 2000/2500 producing 100 bp paired end reads. In total, 18 flatworms, a nemertean, a sipunculid, a gastrotrich, a priapulid and a pterobranch were newly sequenced for this study, for all accession numbers see Table S1.

Transcriptome assembly and peptide prediction

After quality assessment with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) it was determined using PRINSEQ lite [S13] that the first 12 nucleotides needed to be trimmed off the 100 bp reads. Assembly of the trimmed paired reads was done using Trinity v20130225 [S14] using the flag '--min_kmer_cov 2' in addition to default parameters. To test for the presence of cross contamination between libraries run on the same flow cell, we used the bowtie software (http://bowtie-bio.sourceforge.net) and a custom script to identify any assembled transcript with fewer than four read matches which were discarded. In addition we discarded all transcripts in which the number of reads from the intended species matching the transcript was not at least 5 times greater than the number of matches to the transcript from reads from any of the other potentially contaminating species.

For peptide predictions for all nucleotide data sets (i.e. including those publically available), the Trinity script 'transcripts_to_best_scoring_ORFs.pl' was run on the nucleotide data set, keeping all ORFs >100aa. For all peptide datasets cd-hit [S15] was used to reduce redundancy by clustering sequences with a global sequence identity of >95%. All subsequent analyses, including the phylogenetic analyses were based on amino acid sequences.

OMA analysis of homologous genes.

In brief,

i) We processed the proteomes from 52 species using the OMA software (http://omabrowser.org) to identify 3,164 sets of orthologous proteins with at least 28 representative sequences.

ii) for each of these 3,164 sets of orthologs we identified additional orthologs from 35 new species to give 87 species in total.

iii) we selected 55 species from the total of 87 species, eliminating taxa with lots of missing data.

iv) we produced one tree per set of orthologs and kept only gene sets with large sets of monophyletic platyhelminths: 1,348 sets of orthologs.

Non-redundant peptide datasets from 52 species including 27 species of platyhelminths, 8 species of non-platyhelminth lophotrochozoans, and 17 other ecdysozoan, deuterostome, diploblast and sponge species as outgroups (see Table S1) were processed by the OMA software using default settings to identify sets of Orthology Groups (= OG; sets of genes in which all representatives are orthologous to all other members). The all-against-all comparisons of sequences were run in parallel on the UCL Computer Science cluster. Using OMA we were able to identify 230,759 OGs. From these we selected 3,164 OGs with a minimum of 28 species represented (>50% of species with a member of the OG).

Assembling larger sets of orthologs for phylogenetic analyses

As we were able to add newly available data from our own sequenced transcriptomes as well as recently available public data we devised a pipeline for adding new sequences to our existing orthology groups. Running the OMA all-against-all is extremely computationally

intensive and the time taken increases quadratically with respect to the number of species, we decided, therefore, to follow a considerably quicker approach that focussed only on genes that match the 3,164 previously identified OGs.

Using custom perl scripts, for each existing OG we aligned the OG sequences (OG-ALIG) and from this alignment produced a Hidden Markov Model (OG-HMM) using Hmmer1.3b1 [S16]. We next searched the set of sequences from the initial OG with its own OG-HMM to find the score of the lowest sequence match, this lowest score then provided an OG specific cutoff for searching for additional sequences from other sets of peptides.

For each species, an HMM search using each OG-HMM was conducted and the top 3 sequences with a match greater than the OG specific cutoff were kept. Next, for each OG, the standalone version of OMA was run on the collection of potential hits from all new species plus the original constituents of the OG.

OMA standalone found the orthology groups present in each collection of sequences derived from the initial HMM search as well as additional OGs present thanks to the relatively low cutoff used. Sometimes, more than one of the new OGs produced in this way contained identical sequences because the low cutoff meant members of one orthology group could also be picked up by a second. To disentangle these, all instances of OGs in which any sequence also appears in another OG were merged and OMA standalone run again on the merged set of sequences. This approach allowed all paralogs to be disentangled into separate and unique OGs. The end result was 8,424 new OGs, the increase in total number of OGs is due to the presence of paralogs for some genes, typically each paralog had a small number of sequences when compared to the original OG it was based on.

At this stage we had data from 87 taxa with different levels of completeness. To improve the overall quality of the concatenated alignment the 32 least complete or redundant taxa were now deleted (e.g. we kept only the most complete species of the genus *Brachionus*) to leave 55 species. Some lower quality flatworms were retained due to their interest within the scope of the project (see Table S1). The OGs were now selected for further analysis only if they contained sequences from at least 25 species. 2,528 OGs were kept at this stage.

To reduce the likelihood of contaminating sequences or the presence of paralogs within the Rhabditophora, using custom PERL scripts, we cut any rhabditophoran sequences which did not cluster with the main clade of Rhabditophora on a tree constructed for each OG. Each tree was built based on a muscle alignment, trimmed with trimAl [S17] and analysed using MrBayes v 3.2.2 [S18] (settings: prset applyto = (all) aamodelpr = mixed; lset rates = gamma ngammacat = 5; nruns = 2 nchains = 2 ngen = 50000 samplefreq = 10 Diagnfreq = 1000 Burninfrac = 0.5 stoprule = yes Stopval = 0.1 Starttree = parsimony). The OG was only kept if there were more than 10 rhabditophoran sequences clustered in a monophyletic grouping on the tree and if this largest clade contained at least 4 times as many species as the next

largest rhabditophoran clade. All rhabditophoran sequences not members of this largest clade were deleted as potential contaminants/paralogs. After this cleaning procedure there remained 1,348 OGs. We did not include the Catenulida within this requirement in order to allow us to test the monophyly of the Platyhelminthes.

Using the alignment for each remaining OG, a maximum-likelihood tree was calculated using PhyML [4]. PhyML settings used were -o tlr (tlr: tree topology (t), branch length (l) and rate parameters (r) are optimised) -a e (alpha parameter of gamma distribution is estimated) -c 5 (5 gamma rate categories), the substitution model was LG. The total length of that tree (in estimated substitutions per position across all branches) was divided by the number of taxa on the tree to give an estimate of the rate of evolution for each gene. Genes were concatenated in order of their evolutionary rates (see section 'Phylogenetic signal dissection') to produce an overall alignment of 563,188 positions. This was processed to keep only those individual positions with a minimum of 60% occupancy producing a final alignment with 107,659 positions. The overall completeness of the species in the trimmed alignment ranged from 15%-97% with an average of 73% completeness (see Table S1).

Phylogenetic tree reconstruction

Trees were constructed using PhyloBayes 3.3e [3]. The site heterogeneous CAT+GTR+G4 mixture model was used. This model has repeatedly been shown using cross-validation to be optimal for large datasets such as that presented which has the capacity to provide estimates of the large number of parameters required [11,S19-S20].

Two independent runs were performed with a total length of >4000 cycles. To construct the tree, the first 500 cycles were discarded as burn-in, and the topology and posterior consensus support was computed on the remaining trees (Fig. 1).

Trees were also reconstructed using the maximum likelihood approach using PhyML v 3.0. The LG substitution model was selected, the proportion of invariable sites was estimated and a gamma distribution with 4 categories used. An approximate likelihood ratio test using SH-like supports [4] was conducted to provide estimates of support for clades on the best tree (Suppl. Fig. 1).

Trees were also reconstructed using the maximum likelihood approach using RAxML 8 [5]. The CAT GTR substitution model was selected. n.b. the RAXML 'CAT' has no relation to the phylobayes 'CAT' model, it is instead related to the gamma correction.

Jackknifing

In order to provide an alternative estimate of the support for the clades within the tree we used a jackknife approach. 100 jackknife samples were produced by sampling 20,000 positions at random from the full data set. Each data set was analysed for 300 cycles using

the CAT+GTR+G4 analysis of PhyloBayes and a consensus tree produced for each sample using a burn-in of 200 cycles. The 100 consensus trees produced in this way were collated and a master consensus was produced which represents the overall consensus jackknife tree (Fig. 2).

Phylogenetic signal dissection

To gauge the effects of using data sets with different evolutionary rates on the support for different clades in our tree we divided the total alignment, for which genes had been ordered based on rate of evolution from slowest to fastest, into four quartiles, Q1-Q4. Q1 contains the 25% of positions from the slowest evolving genes in the alignment, Q2 the next 25% etc. These 4 quartiles were each used to reconstruct a phylogenetic tree as previously described (Fig. 3).

Supplemental References

S1. Prudhoe, S. (1985). A Monograph on Polyclad Turbellaria. First Edition (Oxford: Oxford University Press).

S2. Pfannkuche, O., and Thiel, H. (1988). Sample processing. In Introduction to the study of meiofauna, R.P. Higgins, and H. Thiel, eds. (Washington: Smithsonian Institution), pp. 134-145.

S3. Westheide, W., and Purschke, G. (1988). Organism processing. In Introduction to the study of meiofauna, R.P. Higgins, and H. Thiel, eds. (Washington: Smithsonian Institution), pp. 146-160.

S4. Luther, A. (1960). Die Turbellarien Ostfennoskandiens I. Acoela, Catenulida, Macrostomida, Lecithoepitheliata, Prolecithophora, und Proseriata. Fauna Fenn. Hels. 7, 1-155.

S5. Young, J.O. (2001). Keys to the freshwater microturbellarians of Britain and Ireland with notes on their ecology. Freshw. Biol. Ass. Sci. Publ. 59, 1-142.

S6. Ax, P. (2008). Plathelminthes aus Brackgewässern der Nordhalbkugel. First Edition (Mainz: Franz Steiner Verlag).

S7. Streble, H., and Krauter, D. (2008). Das Leben im Wassertropfen. Eleventh Edition (Stuttgart: Franckh-Kosmos Verlag).

S8. Lang, A. (1884). Die Polycladen (Fauna und Flora des Golfes von Neapel). First Edition (Leipzig: W. Engelmann).

S9. Luther, A. (1955). Die Dalyelliiden (Turbellaria Neorhabdocoela). Acta Zool. Fenn. 87, 1-337.

S10. Faubel, A. (1983). The Polycladida, Turbellaria. Proposal and establishment of a new system. Part 1. The Acotylea. Mitt. Hamb. Zool. Mus. Inst. 80, 17-121.

S11. Nuttycombe, J.W. (1956). The Catenula of the Eastern United States. Amer. Midl. Nat. 55, 419-433.

S12. Borkott, H. (1970). Geschlechtliche Organisation, Fortpflanzungsverhalten und Ursachen der sexuellen Vermehrung von *Stenostomum sthenum* nov. spec. (Turbellaria,

Catenulida). Mit Beschreibung yon 3 neuen *Stenostomum*-Arten. Z. Morph. Tiere 67, 183-262.

S13. Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. Bioinformatics 27, 863-864.

S14. Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Prot. 8, 1494-1512.

S15. Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658-1659.

S16. Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucl. Acids Res. Web Server Issue 39, W29-W37.

S17. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972-1973.

S18. Ronquist, F., and Huelsenbeck J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

S19. Telford, M.J., and Copley, R.R. (2011). Improving animal phylogenies with genomic data. Trends Genet. 27, 186-195.

S20. Telford, M.J., Lowe, C.J., Cameron, C.B., Ortega-Martinez, O., Aronowicz, J., Oliveri, P., and Copley, R.R. (2014). Phylogenomic analysis of echinoderm class relationships supports Asterozoa. Proc. R. Soc. Lond. B Biol. Sci. 281, 20140479.