

# Agent of Whirling Disease Meets Orphan Worm: Phylogenomic Analyses Firmly Place Myxozoa in Cnidaria

Maximilian P. Nesnidal<sup>1</sup>, Martin Helmkamp<sup>1‡</sup>, Iris Bruchhaus<sup>2</sup>, Mansour El-Matbouli<sup>3</sup>, Bernhard Hausdorf<sup>1\*</sup>

**1** Zoological Museum, University of Hamburg, Hamburg, Germany, **2** Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany, **3** Clinical Division of Fish Medicine, University of Veterinary Medicine, Vienna, Austria

## Abstract

Myxozoa are microscopic obligate endoparasites with complex live cycles. Representatives are *Myxobolus cerebralis*, the causative agent of whirling disease in salmonids, and the enigmatic “orphan worm” *Buddenbrockia plumatellae* parasitizing in Bryozoa. Originally, Myxozoa were classified as protists, but later several metazoan characteristics were reported. However, their phylogenetic relationships remained doubtful. Some molecular phylogenetic analyses placed them as sister group to or even within Bilateria, whereas the possession of polar capsules that are similar to nematocysts of Cnidaria and of minicollagen genes suggest a close relationship between Myxozoa and Cnidaria. EST data of *Buddenbrockia* also indicated a cnidarian origin of Myxozoa, but were not sufficient to reject a closer relationship to bilaterians. Phylogenomic analyses of new genomic sequences of *Myxobolus cerebralis* firmly place Myxozoa as sister group to Medusozoa within Cnidaria. Based on the new dataset, the alternative hypothesis that Myxozoa form a clade with Bilateria can be rejected using topology tests. Sensitivity analyses indicate that this result is not affected by long branch attraction artifacts or compositional bias.

**Citation:** Nesnidal MP, Helmkamp M, Bruchhaus I, El-Matbouli M, Hausdorf B (2013) Agent of Whirling Disease Meets Orphan Worm: Phylogenomic Analyses Firmly Place Myxozoa in Cnidaria. PLoS ONE 8(1): e54576. doi:10.1371/journal.pone.0054576

**Editor:** Axel Janke, BiK-F Biodiversity and Climate Research Center, Germany

**Received:** October 22, 2012; **Accepted:** December 12, 2012; **Published:** January 30, 2013

**Copyright:** © 2013 Nesnidal et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was funded by the priority program “Deep Metazoan Phylogeny” of the Deutsche Forschungsgemeinschaft (<http://www.dfg.de>; grant number: HA 2763/5). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: [hausdorf@zoologie.uni-hamburg.de](mailto:hausdorf@zoologie.uni-hamburg.de)

‡ Current address: School of Life Sciences, Arizona State University, Tempe, Arizona, United States of America

## Introduction

Myxozoa are microscopic obligate endoparasites with complex live cycles that differ between the two subgroups Myxosporea and Malacosporea [1–3]. Myxosporea includes about 2,180 species, with a myxospore phase resulting in the production of myxospores in lower vertebrates, typically in fish, rarely in amphibians and reptiles, and an actinospore phase that engages in sexual reproduction and results in the production of actinospores, in annelids. Just four malacosporean species are known, all parasitizing freshwater bryozoans and some having life stages parasitizing in fish [4,5]. The enigmatic “orphan worm” *Buddenbrockia plumatellae*, which displays a worm-like trophic stage, belongs to this group [6,7]. Several Myxozoa are economically important pathogens such as the myxosporean *Myxobolus cerebralis*, the causative agent for whirling disease in salmonids.

Originally, Myxozoa were classified as protists [8]. Later several metazoan characteristics, namely multicellularity of some life stages, tight junctions, and collagen were detected [1,2,9]. Phylogenetic analyses of 18S rDNA sequences confirmed the placement of Myxozoa within Metazoa and placed them as sister group to or even within Bilateria [6,10–17]. This placement was further confirmed by bilaterian-like *Hox* genes [18,19] and by the identification of *Buddenbrockia* as a myxozoan with longitudinal muscles [6,7].

However, all Myxozoa possess polar capsules that are similar to nematocysts of Cnidaria in ultrastructure and ontogeny and are

used for host attachment [9,20]. Based on these structures, a close relationship between Myxozoa and Cnidaria has been postulated [9,20]. Specifically, some phylogenetic analyses of 18S rDNA sequences of myxosporeans indicated a sister group relationship between Myxozoa and the parasitic cnidarian *Polypodium* [9,21,22]. The name Endocnidozoa has been established for this clade [15]. A clade including Myxozoa and Cnidaria has also been supported by the presence of a minicollagen gene in the malacosporean *Tetracapsuloides bryosalmonae* [23]. Minicollagens are cnidarian-specific constituents of nematocyst walls. Furthermore, phylogenetic analyses of EST data of the malacosporean *Buddenbrockia* [19] placed Myxozoa as sister group of Medusozoa within Cnidaria. However, a topology test could not reject a bilaterian origin for *Buddenbrockia* [19].

Molecular phylogenetic analyses of the relationships of Myxozoa are complicated by their extremely high substitution rates, possibly resulting in long-branch attraction artifacts [11,12,14–17,22]. Evans et al. [17] demonstrated that there are conflicting signals in the phylogenomic data of *Buddenbrockia* and that a removal of just a few sites from this dataset changes the placement of Myxozoa from within Cnidaria to the alternative hypothesis at the base of Bilateria in the maximum likelihood tree.

We sequenced a part of the genome of *Myxobolus cerebralis* (Myxosporea) and compiled a large dataset for phylogenomic analyses to more robustly resolve the relationships of Myxozoa. We constructed a phylogenetic tree based on a dataset including 128 genes, investigated in how far the tree may be affected by long

branch attraction artifacts or compositional bias and tested alternative hypotheses concerning the position of Myxozoa with a topology test.

## Methods

### Sampling and DNA Extraction

Triactinomyxon spores of *Myxobolus cerebralis* were concentrated from the water over a highly infected *Tubifex tubifex* laboratory culture as described by El-Matbouli & Soliman [24]. The triactinomyxon spores were preserved in ethanol and cleaned with Pasteur pipettes. The samples were homogenized by passing through a blunt 20-gauge needle. Genomic DNA was extracted using the QIAamp DNA Mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions with the exception that the samples were incubated with proteinase K over night at 37°C.

### Genomic Sequencing

Whole-genome shotgun data were obtained by running one PicoTiterPlate on the GS FLX Titanium platform (454 Life Sciences, Branford, CT). The library for 454 sequencing was prepared with the GS FLX Titanium General Library Preparation Kit (Roche, Mannheim, Germany), immobilized on beads and clonally amplified using the GS FLX Titanium LV emPCR Kit (Lib-L) (Roche, Mannheim, Germany). The library was then sequenced with the GS FLX Titanium Sequencing Kit XLR70 (Roche, Mannheim, Germany) and the GS FLX Titanium PicoTiterPlate Kit (Roche, Mannheim, Germany). All kits were used according to the manufacturer's protocol. Sequence assembly was performed using MIRA version 3 rc4 (development version) [25].

The genomic sequences have been deposited in the Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/Traces/sra/>) under accession number SRA062243.

### Data Assembly

We used Exonerate version 2.2 [26] to extract orthologs of the 128 protein genes selected by Philippe et al. [27] from the genomic data of *Myxobolus cerebralis* and the EST data of *Buddenbrockia plumatella* [19] with a query set of these protein genes from *Nematostella vectensis*. The sequences were translated using EMBOSS Transeq [28]. The orthology relationships of the *Myxobolus* sequences were confirmed by constructing gene trees in case of doubt. We added the sequences of *M. cerebralis* and *B. plumatella* to a previously assembled phylogenomic dataset with intensively sampled basal metazoans including nine sponges, nine cnidarians, three ctenophores, one placozoan, and 22 slowly evolving bilaterian taxa [27].

### Alignment, Alignment Masking and Gene Selection

The amino acid sequences of the individual ortholog groups were aligned with MAFFT using the most sensitive option L-INS-i [29,30]. To increase the signal-to-noise ratio sections with only random sequence similarity were identified with ALISCORE version 1.0 [31,32] and subsequently excluded with ALICUT (<http://www.utilities.zfmk.de>). This step removed 5.482 of originally 38.415 amino acid positions. All masked alignments were concatenated to a superalignment comprising 32,933 amino acid positions, which has been deposited at TreeBASE (<http://www.treebase.org>, accession number S13690). The protein genes available for the individual taxa are listed in Table S1. Overall 30% of the amino acids in the matrix are missing.

## Phylogenetic Analyses

We performed maximum likelihood analyses using a parallel Pthreads-based version of RAxML version 7.3.0 [33,34] with the LG+F+G model [35]. Based on the complete alignment, we computed 10 maximum likelihood trees using 10 distinct randomized maximum parsimony trees and choose the tree with the highest likelihood. Confidence values for edges of the maximum likelihood tree were computed by rapid bootstrapping [36] (100 replications).

We performed a Bayesian inference analysis with the CAT model that adjusts for site-specific amino acid frequencies [37] as implemented in PhyloBayes version 3.3b (<http://megasun.bch.umontreal.ca/People/lartillot/www/index.htm>). Four independent chains were run for 30,000 points. 5,000 points were discarded as burn-in. The largest discrepancy observed across all bipartitions (maxdiff) is 0.19. Taking every 10th sampled tree, a 50%-majority rule consensus tree was computed using all chains.

To test predefined phylogenetic hypotheses, we calculated the maximum-likelihood tree for a specified hypothesis by resolving multifurcations in a constrained tree with RAxML. Next, we computed per-site log likelihood scores for the global and the constrained maximum-likelihood trees with RAxML and performed an approximately unbiased test [38] using CONSEL [39] to investigate whether the alternative hypotheses can be rejected.

### Influence of Compositional Heterogeneity and Rate Heterogeneity Among Lineages on the Phylogenetic Analyses

Effects of compositional heterogeneity between lineages on the phylogenetic analyses can be mitigated by the exclusion of data partitions that show compositional heterogeneity between lineages [40,41]. To facilitate the identification of such partitions, we scored the compositional heterogeneity of all single protein alignments using relative composition variability, the average variability in composition between taxa [42], but using frequencies instead of absolute numbers [43]. We adapted the formula of relative composition frequency variability (*RCFV*) to amino acids instead of nucleotides:

$$RCFV = \sum_{i=1}^{20} \sum_{j=1}^n |A_{ij} - A_i|/n$$

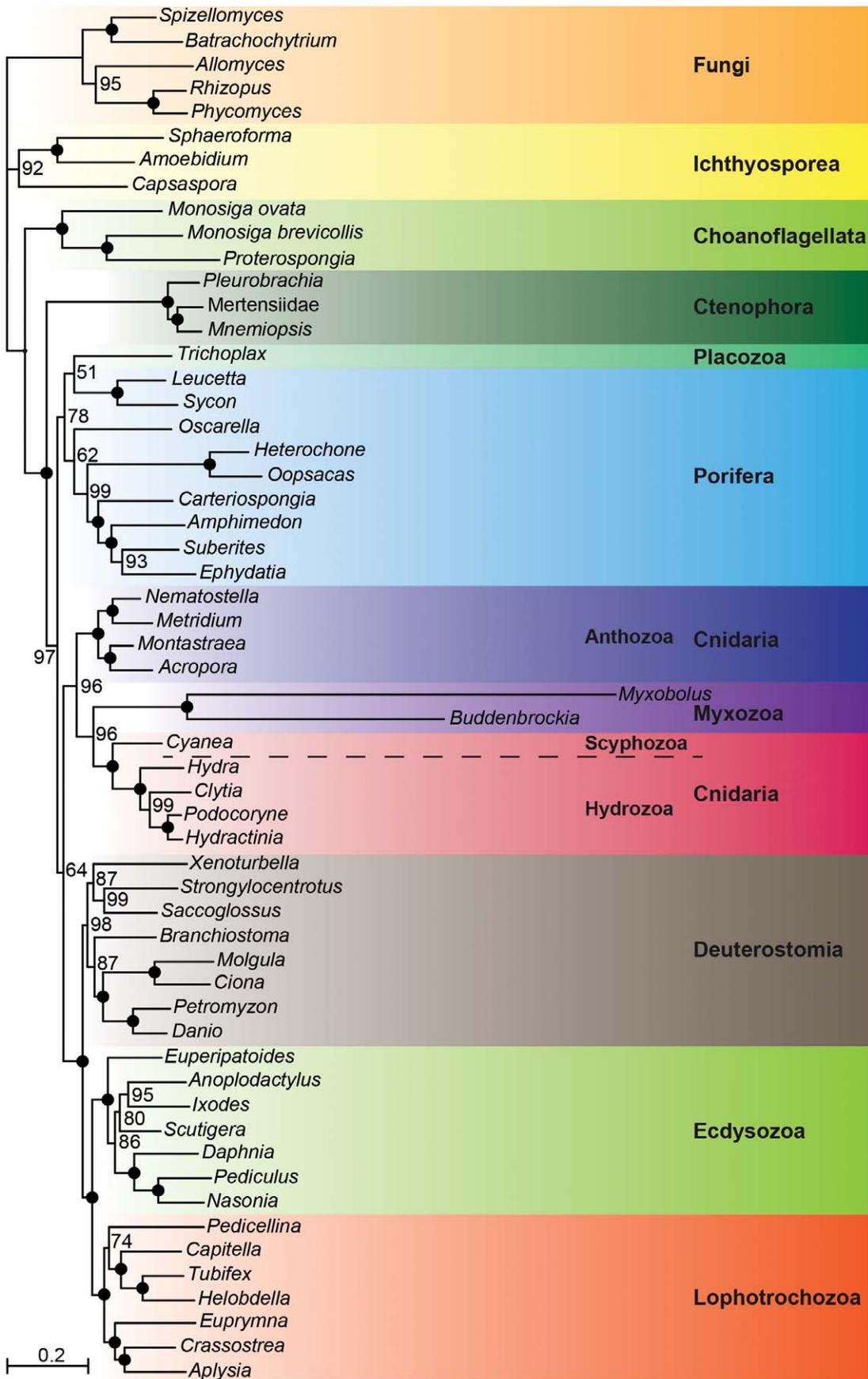
where  $A_{ij}$  is the frequency of amino acid  $i$  in taxon  $j$ , and  $A_i$  is the average frequency of amino acid  $i$  across the  $n$  taxa. Constant sites were excluded from the *RCFV* calculations.

Similarly, effects of rate heterogeneity between lineages can be mitigated by the exclusion of fast evolving character subsets [44,45]. We ranked all individual proteins according to their relative substitution rates estimated as the sum of all pair-wise maximum likelihood distances between the sequences of the 15 species from which sequences of all used proteins were available (Table S1). The ML distances were computed with RAxML and the LG+G+F model.

To evaluate the influence of compositional heterogeneity and rate heterogeneity among lineages on the phylogenetic analyses, we constructed 8 additional datasets by excluding 0, 10 and 33% of the most heterogeneous protein sequences and/or 0, 10 and 33% of the fastest evolving proteins.

## Results and Discussion

Seventy-nine of the 128 genes of the dataset compiled by Philippe et al. [27] were extracted from the new genomic data of the myxosporean *Myxobolus cerebralis*. Moreover, 60 of these genes



**Figure 1. Maximum likelihood tree calculated with the LG+G+F model based on 32,933 amino acid positions derived from 128 proteins of 57 taxa.** Bootstrap values larger than 50% are shown to the right of the nodes; 100% bootstrap values are indicated by black circles. doi:10.1371/journal.pone.0054576.g001

were extracted from EST data of the malacosporean *Buddenbrockia plumatella* [19]. A maximum likelihood analysis of the complete dataset including 32,933 amino acid positions considering the two myxozoan, 44 additional metazoan and 11 outgroup taxa revealed a sister group relationship of *Myxobolus* and *Buddenbrockia* (100% bootstrap support; Fig. 1). The two myxozoans form the sister group of Medusozoa within Cnidaria (96% bootstrap support). The two myxozoans also formed a clade with Medusozoa in a Bayesian inference analysis with the CAT model (Figure S1). However, this clade is only very weakly supported (0.64 posterior probability) and the relationships within this clade are unresolved. It has been shown that phylogenetic analyses with CAT model may be misled by heterogeneity of substitution profiles, the equilibrium frequencies over the twenty amino acids at the individual sites, over time [46]. Actually, such a systematic error affected the reconstruction of the phylogenetic relationships among diploblasts based on mitochondrial protein sequences [46] and it may also be the cause of the weak support for the relationships of myxozoans within diploblasts in our analysis.

Evans et al. [17] suggested that the substantial amount of missing data may have affected the phylogenetic placement of Myxozoa in the phylogenomic study of Jiménez-Guri et al. [19] and that the effects of missing data should further be explored. To test the potential influence of missing data on the phylogenetic placement of Myxozoa, we produced a trimmed dataset by excluding all positions from the amino acid alignment for which no sequence information from at least one of the myxozoans was available. The maximum likelihood analysis of the resulting dataset including 18,273 amino acid positions confirmed the sister group relationship between Myxozoa and Medusozoa (97% bootstrap support; Figure S2).

Our analyses corroborate the placement of Myxozoa in Cnidaria as has been suggested based on the possession of polar capsules that are similar to nematocysts of Cnidaria [9,20], the presence of a minicollagen gene otherwise known only from Cnidaria [23] and the phylogenetic analyses of 18S rDNA sequences of myxosporeans [9,22] and EST data of the malacosporean *Buddenbrockia* [19]. However, previous sequence datasets were not sufficient to reject the alternative hypothesis that Myxozoa (or Endocnidozoa) are the sister group of Bilateria or are even a branch within Bilateria [6,10–17]. With the newly compiled sequence data, this hypothesis could be significantly rejected using the approximately unbiased test (Table 1). Thus, the new genomic data unambiguously support the evolution of the parasitic Myxozoa from Cnidaria.

There are also different hypotheses concerning the relationships of Myxozoa within Cnidaria. Some authors [9,21,22] concluded

that Myxozoa are the sister taxon of *Polypodium*, which has been classified in the hydrozoan clade Narcomedusae (Trachylina) at that time. However, Evans et al. [16] suggested that *Polypodium* is either the sister taxon of Hydrozoa or part of the hydrozoan clade Leptothecata (Hydroidolina). In any case, this would mean that Myxozoa form a clade with Hydrozoa. In contrast, Myxozoa are the sister taxon of Medusozoa in the phylogenomic analysis of Jiménez-Guri et al. [19]. The latter hypothesis is supported by our phylogenetic analyses (Fig. 1). If we constrain the monophyly of Myxozoa+Hydrozoa, Myxozoa are placed as sister group of all Hydrozoa in the constrained maximum likelihood tree. They do not group with *Clytia*, a representative of the Leptothecata in our dataset, as we would expect if Myxozoa and *Polypodium* are part of Leptothecata. Rather, the topology test with our dataset permits to reject the hypothesis that Myxozoa form a clade with Hydrozoa (Table 1).

The topology tests have shown that the placement of Myxozoa as the sister taxon of Medusozoa within Cnidaria cannot be explained by random errors. However, systematic errors resulting from model violations like heterogeneity in substitution rates or amino acid composition among lineages are not considered in topology tests and bootstrap analyses and can confound accurate tree reconstruction [20,44,47,48]. Just as *Buddenbrockia plumatella*, *Myxobolus* is characterized by a very high substitution rate. Whereas the average distance from the base of the maximum likelihood tree (Fig. 1) to the terminal nodes is 0.466, the distance of *Myxobolus* is 1.521. The average pairwise amino acid distance calculated with the LG+F+G model in the complete data set is 0.492, the average pairwise distance between *Myxobolus* and other taxa is 1.111. The highest pairwise distance has been recorded between *Myxobolus* and the mertensiid ctenophore (1.286). In fact, it has been supposed that the placement of myxozoans, might be affected by long branch attraction [11,12,14–17].

We therefore assessed the influence of proteins with varying evolutionary rates and amino acid compositions among lineages on the inferred phylogenetic relationships of the myxozoans via a step-wise exclusion analysis. We removed the 10% or 33% fastest evolving proteins and/or the 10% or 33% proteins with the highest relative composition variability from the dataset and repeated the phylogenetic reconstruction. The monophyly of Myxozoa was confirmed in all of the resulting 8 trees with 100% bootstrap support (Figures S3–S10). The support for the monophyly of Cnidaria inclusive Myxozoa and of Myxozoa+Medusozoa was high (88–99%) in all of these analyses (Table 2; Figures S3–S10). Thus, these sensitivity analyses corroborate the robustness of the placement of Myxozoa as sister group to Medusozoa within Cnidaria.

**Table 1. Results of the approximately unbiased test.**

phylogenetic hypothesis	references claiming the hypothesis	likelihood	$\Delta$ likelihood <sup>a</sup>	AU <sup>b</sup>
ML tree		−985,545	0	0.998
Myxozoa form a clade with Bilateria	6, 10–17	−985,636	91	0.002*
Myxozoa form a clade with Hydrozoa	9, 22	−985,647	102	0.000*

<sup>a</sup> $\Delta$  Likelihood: differences between the likelihood of a constrained tree and the maximum likelihood tree.

<sup>b</sup>AU: approximately unbiased test ( $p$ -values). Values for topologies significantly rejected at the 0.05 level are indicated by an asterisk.

doi:10.1371/journal.pone.0054576.t001

**Table 2.** Sensitivity of the phylogenetic analysis to rate heterogeneity and compositional heterogeneity among lineages.

Approach	no exclusion of heterogeneous proteins	exclusion of 10% most heterogeneous proteins	exclusion of 33% most heterogeneous proteins
no exclusion of fast proteins	96	98	99
	<i>96</i>	<i>98</i>	<i>99</i>
exclusion of 10% fastest proteins	98	99	96
	<i>98</i>	<i>99</i>	<i>96</i>
exclusion of 33% fastest proteins	90	88	93
	<i>90</i>	<i>88</i>	<i>94</i>

Bootstrap support values for the monophyly of Cnidaria inclusive Myxozoa (upper values) and of Myxozoa+Medusozoa (lower values in italic).  
doi:10.1371/journal.pone.0054576.t002

The next step will be the generation of genomic data for *Polypodium* so that the hypothesis that this parasitic cnidarian represents the sister group of Myxozoa [9,21,22] can be tested in phylogenomic analyses.

## Conclusion

The phylogenetic placement of Myxozoa within Cnidaria might have important implications for parasitologists and veterinarians seeking to develop strategies for avoiding infection with these economically important parasites and combating the resulting diseases. Our results expose the evolutionary plasticity of the cnidarian bauplan, which gave rise to obligatory parasitic organisms reminiscent of protozoa or bilaterian worms from free-living, radially symmetrical origins. The loss of eumetazoan apomorphies like epithelial tissue layers, gut, nerve and sensory cells, shows that Myxozoa are a primary example for the extreme reduction of complexity that comes with the evolution of a parasitic lifestyle. More detailed analyses of myxozoan genomes in the future might further strengthen our understanding of metazoan evolution by revealing the genetic underpinnings that drive these profound changes during myxozoan development.

## Supporting Information

**Figure S1 Bayesian inference reconstructions with the CAT model based on 32,933 amino acid positions derived from 128 proteins of 57 taxa.** Bayesian posterior probabilities are shown to the right of the nodes.  
(PDF)

**Figure S2 Maximum likelihood tree calculated with the LG+G+F model based on 18,273 amino acid positions derived from 128 proteins (after excluding all positions for which no sequence information is available from myxozoans).** Bootstrap values larger than 50% are shown to the right of the nodes.  
(PDF)

**Figure S3 Maximum likelihood tree calculated with the LG+G+F model based on 31,488 amino acid positions derived from 115 proteins (after excluding the 10% most heterogeneous proteins).** Bootstrap values larger than 50% are shown to the right of the nodes.  
(PDF)

**Figure S4 Maximum likelihood tree calculated with the LG+G+F model based on 26,974 amino acid positions derived from 85 proteins (after excluding the 33% most heterogeneous proteins).** Bootstrap values larger than 50% are shown to the right of the nodes.

(PDF)

**Figure S5 Maximum likelihood tree calculated with the LG+G+F model based on 30,565 amino acid positions derived from 115 proteins (after excluding the 10% fastest evolving proteins).** Bootstrap values larger than 50% are shown to the right of the nodes.

(PDF)

**Figure S6 Maximum likelihood tree calculated with the LG+G+F model based on 29,712 amino acid positions derived from 106 proteins (after excluding the 10% fastest evolving proteins and the 10% most heterogeneous proteins).** Bootstrap values larger than 50% are shown to the right of the nodes.

(PDF)

**Figure S7 Maximum likelihood tree calculated with the LG+G+F model based on 26,760 amino acid positions derived from 84 proteins (after excluding the 10% fastest evolving proteins and the 33% most heterogeneous proteins).** Bootstrap values larger than 50% are shown to the right of the nodes.

(PDF)

**Figure S8 Maximum likelihood tree calculated with the LG+G+F model based on 23,553 amino acid positions derived from 85 proteins (after excluding the 33% fastest evolving proteins).** Bootstrap values larger than 50% are shown to the right of the nodes.

(PDF)

**Figure S9 Maximum likelihood tree calculated with the LG+G+F model based on 23,289 amino acid positions derived from 81 proteins (after excluding the 33% fastest evolving proteins and the 10% most heterogeneous proteins).** Bootstrap values larger than 50% are shown to the right of the nodes.

(PDF)

**Figure S10 Maximum likelihood tree calculated with the LG+G+F model based on 21,643 amino acid positions derived from 67 proteins (after excluding the 33% fastest evolving proteins and the 33% most heterogeneous proteins).** Bootstrap values larger than 50% are shown to the right of the nodes.

(PDF)

**Table S1 Protein genes available from the studied taxa.**  
(XLSX)

## Acknowledgments

We are grateful to M. Kube and R. Reinhardt (Max Planck Institute for Molecular Genetics, Berlin) for the construction and sequencing of genomic libraries of *Myxobolus cerebralis*, to Malik Alawi for a script for calculating the RCFV values and to H. Philippe for providing a previously compiled phylogenomic dataset.

## References

- Kent ML, Andree KB, Bartholomew JL, El-Matbouli M, Desser SS, et al. (2001) Recent advances in our knowledge of the Myxozoa. *J Eukaryot Microbiol* 48: 395–413.
- Canning EU, Okamura B (2004) Biodiversity and evolution of the Myxozoa. *Adv Parasitol* 56: 44–131.
- Lom J, Dyková I (2006) Myxozoan genera: definition and notes on taxonomy, life-cycle terminology and pathogenic species. *Folia Parasitologica* 53: 1–36.
- Tops S, Baxa DV, McDowell TS, Hedrick RP, Okamura B (2004) Evaluation of malacosporean life cycles through transmission studies. *Dis Aquat Organ* 60: 109–121.
- Grabner DS, El-Matbouli M (2010) Experimental transmission of malacosporean parasites from bryozoans to common carp (*Cyprinus carpio*) and minnow (*Phoxinus phoxinus*). *Parasitology* 137: 629–639.
- Monteiro AS, Okamura B, Holland PWH (2002) Orphan worm finds a home: *Buddenbrockia* is a myxozoan. *Mol Biol Evol* 19: 968–971.
- Okamura B, Curry A, Wood TS, Canning EU (2002) Ultrastructure of *Buddenbrockia* identifies it as a myxozoan and verifies the bilaterian origin of the Myxozoa. *Parasitology* 124: 215–223.
- Lom J (1990) Phylum Myxozoa. In *Handbook of Protozoa*. Edited by Margulis L, Corliss JO, Melkonian M, Chapman DJ. Boston: Jones and Bartlett. 36–52.
- Siddall ME, Martin DS, Bridge D, Desser SS, Cone DK (1995) The demise of a phylum of protists: phylogeny of Myxozoa and other parasitic Cnidaria. *J Parasitol* 81: 961–967.
- Smothers JF, von Dohlen CD, Smith LH, Spall RD (1994) Molecular evidence that the myxozoan protists are metazoans. *Science* 265: 1719–1721.
- Cavalier-Smith T, Allsopp MTEP, Chao EE, Boury-Esnault N, Vacelet J (1996) Sponge phylogeny, animal monophyly and the origin of the nervous system: 18S rRNA evidence. *Can J Zool* 74: 2031–2045.
- Hanelt B, van Schyndel D, Adema CM, Lewis LA, Loker ES (1996) The phylogenetic position of *Rhopalura ophiocoma* (Orthonectida) based on 18S ribosomal DNA sequence analysis. *Mol Biol Evol* 13: 1187–1191.
- Schlegel M, Lom J, Stechmann A, Bernhard D, Leipe D, et al. (1996) Phylogenetic analysis of complete small subunit ribosomal RNA coding region of *Myxidium lieberkuehni*: evidence that Myxozoa are Metazoa and related to the Bilateria. *Arch Protistenkd* 147: 1–9.
- Kim JH, Kim W, Cunningham CW (1999) A new perspective on lower metazoan relationships from 18S rDNA sequences. *Mol Biol Evol* 16: 423–427.
- Zrzavý J, Hypša V (2003) Myxozoa, *Polypodium*, and the origin of the Bilateria: The phylogenetic position of “Endocnidozoa” in light of the rediscovery of *Buddenbrockia*. *Cladistics* 19: 164–169.
- Evans N, Lindner A, Raikova E, Collins A, Cartwright P (2008) Phylogenetic placement of the enigmatic parasite, *Polypodium hydriforme*, within the phylum Cnidaria. *BMC Evol Biol* 8: 139.
- Evans NM, Holder MT, Barbeitos MS, Okamura B, Cartwright P (2010) The phylogenetic position of Myxozoa: Exploring conflicting signals in phylogenomics and ribosomal data sets. *Mol Biol Evol* 27: 2733–2746.
- Anderson CL, Canning EU, Okamura B (1998) A triploblast origin for Myxozoa? *Nature* 392: 346–347.
- Jiménez-Guri E, Philippe H, Okamura B, Holland PWH (2007) *Buddenbrockia* is a cnidarian worm. *Science* 317: 116–118.
- Lom J, Dyková I (1997) Ultrastructural features of the actinosporean phase of Myxosporaea (phylum Myxozoa): a comparative study. *Acta Protozool* 36: 83–103.
- Zrzavý J, Mihulka S, Kepka P, Bezděk A, Tietz D (1998) Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. *Cladistics* 14: 249–285.
- Siddall ME, Whiting MF (1999) Long-branch aberrations. *Cladistics* 15: 9–24.
- Holland JW, Okamura B, Hartikainen H, Secombes CJ (2011) A novel mimicollagen gene links cnidarians and myxozoans. *Proc R Soc B* 278: 546–553.
- El-Matbouli M, Soliman H (2006) Construction and screening of a cDNA library from the triactinomyxon spores of *Myxobolus cerebralis*, the causative agent of salmonid Whirling Diseases. *Parasitology* 132: 467–477.

## Author Contributions

Conceived and designed the experiments: BH. Performed the experiments: MPN MH IB. Analyzed the data: MPN BH. Contributed reagents/materials/analysis tools: ME IB. Wrote the paper: BH MPN MH IB ME.

- Chevreur B, Wetter T, Suhai S (1999) Genome sequence assembly using trace signals and additional sequence information. *Proc German Conference Bioinformatics, GCB '99*. 45–56.
- Slater GSC, Birney E (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6: 31.
- Philippe H, Derelle R, Lopez P, Pick K, Borchellini C, et al. (2009) Phylogenomics revives traditional views on deep animal relationships. *Curr Biol* 19: 706–712.
- Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet* 16: 276–277.
- Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucl Acids Res* 30: 3059–3066.
- Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* 9: 286–298.
- Misof B, Misof K (2009) A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: A more objective means of data exclusion. *Syst Biol* 58: 21–34.
- Kück P, Meusemann K, Dambach J, Thormann B, Reumont BM, et al. (2010) Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Front Zool* 7: 10.
- Stamatakis A (2006) RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
- Ott M, Zola J, Aluru S, Stamatakis A (2007) Large-scale maximum likelihood-based phylogenetic analysis on the IBM BlueGene/L. *Proceedings of the ACM/IEEE conference on Supercomputing 2007*. Reno, Nevada. ACM.
- Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* 25: 1307–1320.
- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAXML Web servers. *Syst Biol* 57: 758–771.
- Lartillot N, Philippe H (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol Biol Evol* 21: 1095–1109.
- Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. *Syst Biol* 51: 492–508.
- Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246–1247.
- Collins TM, Fedrigo O, Naylor GJP (2005) Choosing the best genes for the job: The case for stationary genes in genome-scale phylogenetics. *Syst Biol* 54: 493–500.
- Nesnidal MP, Helmkampf M, Bruchhaus I, Hausdorf B (2010) Compositional heterogeneity and phylogenomic inference of metazoan relationships. *Mol Biol Evol* 27: 2095–2104.
- Phillips MJ, Penny D (2003) The root of the mammalian tree inferred from whole mitochondrial genomes. *Mol Phylogenet Evol* 28: 171–185.
- Zhong M, Hansen B, Nesnidal M, Golombek A, Halanych KM, Struck TH (2011) Detecting the sympleiomorphy trap: a multigene phylogenetic analysis of terebelliform annelids. *BMC Evol Biol* 11: 369.
- Delsuc F, Brinkmann H, Philippe H (2005) Phylogenomics and the reconstruction of the tree of life. *Nat Rev Genet* 6: 361–375.
- Jeffroy O, Brinkmann H, Delsuc F, Philippe H (2006) Phylogenomics: the beginning of incongruence? *Trends Genet* 22: 225–231.
- Roure B, Philippe H (2011) Site-specific time heterogeneity of the substitution process and its impact on phylogenetic inference. *BMC Evol Biol* 11: 17.
- Jermiin LS, Ho SYW, Ababneh F, Robinson J, Larkum AWD (2004) The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. *Syst Biol* 53: 638–643.
- Rodríguez-Espeleta N, Brinkmann H, Roure B, Lartillot N, Lang BF, Philippe H (2007) Detecting and overcoming systematic errors in genome-scale phylogenies. *Syst Biol* 56: 389–399.