RESEARCH NOTE

Revisiting the phylogeny of phylum Ctenophora: a molecular perspective [version 1; peer review: 1 approved with reservations, 3 not approved]

Luis A. Arteaga-Figueroa¹, Valentina Sánchez-Bermúdez¹,
Nicolás D. Franco-Sierra ¹⁻³

¹Semillero de Biología Computacional, Departamento de Ciencias Biológicas, Escuela de Ciencias, Universidad EAFIT, Medellín, Colombia

²Grupo CIBIOP (Ciencias Biológicas y Bioprocesos), Departamento de Ciencias Biológicas, Escuela de Ciencias, Universidad EAFIT, Medellín, Colombia

³Grupo BEC (Biodiversidad, Evolución y Conservación), Departamento de Ciencias Biológicas, Escuela de Ciencias, Universidad EAFIT, Medellín, Colombia

V1 First published: 20 Dec 2016, 5:2881
<https://doi.org/10.12688/f1000research.10426.1>
Latest published: 21 Aug 2017, 5:2881
<https://doi.org/10.12688/f1000research.10426.2>

Abstract

The phylogenetic relationships of deep metazoans, specifically in the phylum Ctenophora, are not totally understood. Previous studies have been developed on this subject, mostly based on morphology and single gene analyses (rRNA sequences). Several loci (protein coding and ribosomal RNA) from taxa belonging to this phylum are currently available on public databases (e.g. GenBank). Here we revisit Ctenophora molecular phylogeny using public sequences and probabilistic methods (Bayesian inference and maximum likelihood). To get more reliable results multi-locus analyses were performed using 5.8S, 28S, ITS1, ITS2 and 18S, and IPNS and GFP-like proteins. Best topologies, consistent with both methods for each data set, are shown and analysed. Comparing the results of the phylogenetic reconstruction with previous research, most clades showed the same relationships as the ones found with morphology and single gene analyses, consistent with hypotheses made in previous research. There were also some unexpected relationships clustering species from different orders.

Keywords

Ctenophora , Phylogenetic reconstruction , Ribosomal subunits , Non-fluorescent protein (GFP-like) , Bayesian Inference , Maximum Likelihood , Isopenicillin-n-synthase (IPNS)

Open Peer Review

Approval Status ?

	1	2	3	4
version 2 (revision) 21 Aug 2017		 view		
version 1 20 Dec 2016	 view	 view	 view	 view

- Martin Dohrmann**, Ludwig-Maximilians-Universität München, Munich, Germany
- Steven H.D. Haddock**, Monterey Bay Aquarium Research Institute, Moss Landing, USA
- Kevin M. Kocot**, The University of Alabama, Tuscaloosa, USA
- D Timothy J. Littlewood** , Natural History Museum, London, UK

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the **Phylogenetics** collection.

Corresponding author: Nicolás D. Franco-Sierra (nfranco@eafit.edu.co)

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2016 Arteaga-Figueroa LA *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 work is properly cited.

How to cite this article: Arteaga-Figueroa LA, Sánchez-Bermúdez V and Franco-Sierra ND. **Revisiting the phylogeny of phylum Ctenophora: a molecular perspective [version 1; peer review: 1 approved with reservations, 3 not approved]** F1000Research 2016, 5:2881 <https://doi.org/10.12688/f1000research.10426.1>

First published: 20 Dec 2016, 5:2881 <https://doi.org/10.12688/f1000research.10426.1>

Introduction

The relationships among deep metazoans (Cnidaria and Ctenophora) and Parazoa (Porifera and Placozoa) are not totally clear¹. In this paper we try to reconstruct the phylogeny inside the phylum Ctenophora with state of the art methods and compare our results with previous work^{2,3}. Ctenophores are a key phylum for the understanding of the development of organ systems, triploblastic animals and bilateral symmetry⁴. Our goal was to reconstruct the phylogeny of previous studies using as many sequences as possible available on GenBank. The use of these sequences allowed us to perform a multilocus analysis (MLSA), instead of the single gene analyses previously performed. The sequences selected for this study have never been used for phylogenetic analysis exclusively of this phylum^{5,6}.

Our research consists of 1) the analysis of ribosomal genes (5.8S; 28S; ITS1; ITS2 and 18S) and 2) the analysis of two ortholog genes found in ctenophores (a GFP-like non-fluorescent protein, and isopenicillin-N-synthase FYY1).

The ribosomal genes were analysed using partitioned nucleotide substitution models while the ortholog genes were analysed using partitioned amino acid substitution models. We compared the findings of our two approaches (ribosomal and ortholog genes) to each other and against the previously reported phylogenetic trees obtained from molecular data³ and morphological data².

Methods

All sequences corresponding to Ctenophora (Taxonomy ID: 10197) were retrieved from GenBank's nucleotide database. Short

sequences (those shorter than 150 base pairs) or ambiguously labeled sequences (those not assigned to a specific species) were discarded. This criterion was used to obtain an almost complete matrix including most loci available of reported taxa across the phylum.

Seven loci were chosen for analysis; five corresponded to ribosomal RNA regions (5.8S; 28S; ITS1; ITS2 and 18S). The other two corresponded to ortholog genes: a putative non-fluorescent protein (GFP-like protein), and isopenicillin-N-synthase FYY1 (IPNS), a protein involved in the bioluminescence process. The sequences were extracted using the annotation of the retrieved records using Biopython 1.67⁷.

The taxa present for each analysis is listed in Table 1 and Table 2. All sequences used (with corresponding accession numbers) and scripts used for analysis are available at <http://doi.org/10.5281/zenodo.193080>¹⁶.

Given the phylogenetic distance between the different taxa of this phylum, for the protein coding genes, we decided to work at the amino acid sequence level due to high sequence saturation at the nucleotide level.

The sequences corresponding to the ortholog genes were translated *in silico* using DNA2PEP 1.1⁸ with standard genetic code, and aligned using MAFFT 7.222⁹. A MLSA was performed using these two loci. Alignments were concatenated using Python scripts and partitioned by gene to be analyzed for amino acid model and

Table 1. Protein coding genes used for phylogenetic reconstruction. (1) Sequence was available on GenBank. (-) sequence was not available on GenBank or not reported.

Order	Family	Species	IPNS	GFP-like
Beroida	Beroidae	<i>Beroe abyssicola</i>	1	1
Beroida	Beroidae	<i>Beroe forskalii</i>	1	1
Cestida	Cestidae	<i>Cestum veneris</i>	-	1
Cestida	Cestidae	<i>Velamen parallelum</i>	1	1
Cydippida	Bathyctenidae	<i>Bathytecna chuni</i>	1	-
Cydippida	Mertensiidae	<i>Charistephane fugiens</i>	1	1
Cydippida	Dryodoridae	<i>Dryodora glandiformis</i>	1	1
Cydippida	Euplokamididae	<i>Euplokamis dunlapae</i>	1	1
Cydippida	Haeckelidae	<i>Haeckelia beehleri</i>	-	1
Cydippida	Haeckelidae	<i>Haeckelia rubra</i>	1	1
Cydippida	Pleurobrachiidae	<i>Hormiphora californensis</i>	1	1
Cydippida	Lampeidae	<i>Lampea lactea</i>	1	1
Lobata	Bathocyroidae	<i>Bathocyroe fosteri</i>	1	1
Lobata	Bolinopsidae	<i>Bolinopsis infundibulum</i>	1	1
Lobata	Deiopeidae	<i>Deiopea kaloktenota</i>	-	1
Lobata	Lampoctenidae	<i>Lampoctis cruentiventer</i>	1	1
Lobata	Leucotheidae	<i>Leucothea pulchra</i>	-	1
Lobata	Ocyropsidae	<i>Ocyropsis maculata</i>	1	1
Thalassocalycida	Thalassocalycidae	<i>Thalassocalyce inconstans</i>	1	1

Table 2. Ribosomal RNA genes used for phylogenetic reconstruction. (1) Sequence was available on GenBank. (-) sequence was not available on GenBank or not reported.

Order	Family	Species	18S	28S	ITS1	5.8S	ITS2
Beroida	Beroidae	<i>Beroe cucumis</i>	1	1	1	1	1
Beroida	Beroidae	<i>Beroe forskalii</i>	1	1	1	1	1
Beroida	Beroidae	<i>Beroe ovata</i>	1	1	1	1	1
Cestida	Cestidae	<i>Cestum veneris</i>	1	1	1	-	-
Cestida	Cestidae	<i>Velamen parallelum</i>	1	1	1	-	1
Cydippida	Mertensiidae	<i>Charistephane fugiens</i>	1	-	-	-	-
Cydippida	Haeckelidae	<i>Haeckelia beehleri</i>	1	1	1	1	1
Cydippida	Haeckelidae	<i>Haeckelia rubra</i>	1	1	1	1	1
Cydippida	Pleurobrachiidae	<i>Hormiphora plumosa</i>	1	1	1	1	1
Cydippida	Lampeidae	<i>Lampea pancerina</i>	1	-	-	-	-
Cydippida	Mertensiidae	<i>Mertensia ovum</i>	1	-	1	1	-
Cydippida	Pleurobrachiidae	<i>Pleurobrachia bachei</i>	1	1	1	-	-
Cydippida	Pleurobrachiidae	<i>Pleurobrachia brunnea</i>	1	1	-	-	-
Cydippida	Pleurobrachiidae	<i>Pleurobrachia globosa</i>	1	1	-	-	-
Cydippida	Pleurobrachiidae	<i>Pleurobrachia pileus</i>	1	1	-	1	-
Lobata	Bolinopsidae	<i>Bolinopsis infundibulum</i>	1	1	1	1	1
Lobata	Deiopeidae	<i>Deiopea kaloktenota</i>	1	-	-	-	-
Lobata	Lampoctenidae	<i>Lampoctis cruentiventer</i>	1	-	-	-	-
Lobata	Leucotheidae	<i>Leucothea multicornis</i>	1	-	-	-	-
Lobata	Leucotheidae	<i>Leucothea pulchra</i>	-	1	-	1	1
Lobata	Bolinopsidae	<i>Mnemiopsis leidyi</i>	1	1	1	1	1
Lobata	Ocyropsidae	<i>Ocyropsis crystallina</i>	1	1	1	1	1
Lobata	Ocyropsidae	<i>Ocyropsis maculata</i>	1	1	1	1	1
Platyctenida	Coeloplanidae	<i>Coeloplana bocki</i>	1	-	1	-	-
Platyctenida	Coeloplanidae	<i>Coleoplana bannwarthii</i>	1	1	1	1	1
Platyctenida	Coeloplanidae	<i>Vallicula multiformis</i>	1	1	1	1	1
Thalassocalycida	Thalassocalycidae	<i>Thalassocalyce inconstans</i>	1	1	-	-	-

best partition scheme using PartitionFinderProtein 1.1.1¹⁰ Model adjustment was assessed using Bayesian information criterion (BIC). The best model found by PartitionFinderProtein 1.1.1 for IPNS partition was LG + G + I, and LG + G had better adjustment for GFP-Like partition. Phylogenetic reconstruction for the ortholog genes was carried out by maximum likelihood (ML) and Bayesian inference (BI) methods, using both Garli 2.01¹¹ and MrBayes 3.2.6¹², with the proper amino acid substitution model parameters for each partition.

For the ML analysis, using Garli, a total of 5 independent ML searches were performed and supported with 65 bootstrap pseudoreplicates. For BI analysis, using MrBayes, two independent MCMC runs (four chains for each) were carried out for 1,000,000 generations, using a relative burn-in discard of 35% of total sampled trees (sampling frequency of 100 generations).

For the five rRNA loci, the automated pipeline PhyPipe¹³ (available at: <https://gitlab.com/cibiop/phyPIPE/>) was used for phylogenetic

reconstruction by BI and ML methods. A regular PhyPipe run comprises DNA sequence alignment with MAFFT 7.222, partition analysis with PartitionFinder, and phylogenetic reconstruction with RAxML 8.2.8¹⁴, MrBayes 3.2.6¹² or Garli 2.01¹¹. For this analysis, MrBayes was executed under the following parameters: two independent MCMC runs, four chains, 1,000,000 generations, 35% of relative burn-in and sampling frequency of 100. For ML analysis, Garli was executed doing first a ML search (5 independent searches), then 1000 bootstrap pseudoreplicates were performed and mapped to the best ML topology using SumTrees from DendroPy 4.1.0 package¹⁵.

Results

The majority of the phylum analysed in this study show a standard grouping condition; the organisms that are related in one of the analyses are also related in the other. This is more evident comparing at family level, where the individuals of the same family, and in some cases order, grouped with other organisms of the same order. Exceptions are discussed below.

In the reported trees there are families represented by several species while complete orders are represented by just one species. For the purpose of clarity, from now on the families represented by several species will be discussed at the family level while the orders represented by just one species will be discussed using the representing species.

Thalassocalyce inconstans and *Lamproteus cruentiventer* are unexpectedly grouped together in both analyses, but the position is not the same in comparison with the other clades. The support values in the ribosomal tree are very low compared to the ortholog genes tree.

The order Cydippida is divided in five different clades or sub-groups, shown in different tones of blue in Figure 1. We confirm that this group is paraphyletic as reported previously in 2,3. The species *Bathytene chuni* (Bathyteneidae, Cydippida) is grouped with *Ocyropsis maculata* (Ocyropsidae, Lobata) in amino acid

analysis, but there is no information on the ribosomal sequences of *Bathytene chuni*, so it was not possible to compare.

In the amino acid tree, Pleurobrachiidae, Mertensiidae, Lampeidae, Euplokamididae form a clade; but the relationships between them are not clear, and the bootstrap values and posterior probability are low in this group. In the ribosomal analysis we could include the Platyteneida order, which grouped with high support in the clade formed by the mentioned families and order. In the ribosomal analysis, we see some shared features with one of Harbison's trees³. Also the Thalassocalyce-Lamproteus clade is related to this group, but the position varies depending of the analysis. *Dryodora glandiformis* groups with the clade formed by Lobata species, but this result is only evident in the ortholog genes tree, due to the lack of rRNA sequences for this particular taxon.

All the trees were rooted using *Beroidea* as the outgroup, following the hypothesis that this is the most basal group. The same choice

Figure 1. Phylogenetic trees based on protein sequences of IPNS and GFP-like genes (on left) and rRNA loci (on right). Trees were constructed using Bayesian inference and maximum likelihood methods and consistent topologies were found within methods. Support values are shown at nodes in the form of posterior probability/bootstrap value.

of root was made by Harbison². Additionally, the Beroe genus is a good outgroup because it belongs to the class Nuda while the other studied species belong to the class Tentaculata (our ingroup).

Bathocyroe fosteri is present in an unexpected position in the tree. It should have been included in the Lobata clade. Instead, it was placed outside the subgroup containing the Lobata, Pleurobrachiiidae, Mertensiidae, Lampeidae and Thalassocalycidae families. This finding is not compared with rDNA loci analysis since ribosomal sequences for *Bathocyroe fosteri* were not available.

In the research performed by Harbison², the Lobata group was placed below the Cestida group. Later, this finding was not discussed by Podar *et al.* 2001³ as in the ribosomal data they used both groups are in a polytomy. Our finding, using rDNA, is in concordance with the findings of Podar *et al.* 2001³, but using the ortholog genes the finding is contrary to what was proposed by Harbison. We found, on the ortholog genes analysis, that Cestida group is the one derived from Lobata and not vice versa as suggested by Harbison. This finding has a high bootstrap value and posterior probability support.

Discussion

The Haeckelidae family preserves its position in the phylogenetic trees placed as sister group of all the other Tentaculata taxa analysed, with high support, according to previous studies².

The lack of reported DNA sequences of few groups, like the orders Cryptoboliferida, Cambodjiida, Ganeshida; several families, like Eurhamplaeidae and some records reported as Ctenophora *incertae sedis* (Tentaculata *incertae sedis*), make it harder to have an entire vision of the phylogenetic relationships inside the phylum. The order Ganeshida is grouped with Lobata, according to Harbison² and the lack of this group may have caused a misplacement of the Thalassocalycida representant.

To improve the results, Coeloplanidae should be included in the protein phylogenetic analysis. Further studies including more sequences from families such as Leucotheidae, Lampoctenidae and Thalassocalycidae are also needed to solve the resulting polytomies and to obtain better support to confirm the relationship between *Thalassocalyce inconstans* and *Lampoctis cruentiventer* in the rRNA analysis.

Data availability

The raw data and scripts used for this project are available in Zenodo, DOI [10.5281/zenodo.193080](https://doi.org/10.5281/zenodo.193080)¹⁶.

Author contributions

LAAF conceived the study, performed the sequence compilation and literature revision. LAAF, VSB and NDFS carried out the phylogenetic reconstructions and analysed the results. All authors were involved in writing the manuscript and have agreed to its final content.

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no grants were involved in supporting this work.

Acknowledgments

We specially thank Sergio Pulido-Tamayo for stimulating discussions and critical review of the manuscript, Juan F. Díaz-Nieto, Javier C. Alvarez and Diana Rincón T. for their guidance and valuable comments. We also want to thank Lizette I. Quan-Young and Steve Haddock for providing useful bibliography and sequences for this analysis, respectively.

References

- Lynch M: **The Age and Relationships of the Major Animal Phyla.** *Evolution*. 1999; **53**(2): 319–325.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Harbison GR: **On the classification and evolution of ctenophora.** In: *The origin and relationship of lower invertebrates*. S Conway Morris, JD George, R Gilson, and HM Platt, ed. 1985.
- Podar M, Haddock SH, Sogin ML, *et al.*: **A molecular phylogenetic framework for the phylum Ctenophora using 18S rRNA genes.** *Mol Phylogenet Evol*. 2001; **21**(2): 218–230.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Ryan JF, Pang K, Schnitzler CE, *et al.*: **The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution.** *Science*. 2013; **342**(6164): 1242592.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Francis WR, Christianson LM, Powers ML, *et al.*: **Non-excitable fluorescent protein orthologs found in ctenophores.** *BMC Evol Biol*. 2016; **16**(1): 167.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Francis WR, Shaner NC, Christianson LM, *et al.*: **Occurrence of Isopenicillin-N-Synthase Homologs in Bioluminescent Ctenophores and Implications for Coelenterazine Biosynthesis.** *PLoS One*. 2015; **10**(6): e0128742.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cock PJ, Antao T, Chang JT, *et al.*: **Biopython: freely available Python tools for computational molecular biology and bioinformatics.** *Bioinformatics*. 2009; **25**(11): 1422–1423.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Wernersson R, Pedersen AG: **RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences.** *Nucleic Acids Res*. 2003; **31**(13): 3537–3539.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Katoh K, Standley DM: **MAFFT multiple sequence alignment software version 7: improvements in performance and usability.** *Mol Biol Evol*. 2013; **30**(4): 772–80.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lanfear R, Calcott B, Kainer D, *et al.*: **Selecting optimal partitioning schemes for phylogenomic datasets.** *BMC Evol Biol*. 2014; **14**(1): 82.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bazinnet AL, Zwickl DJ, Cummings MP: **A gateway for phylogenetic analysis powered by grid computing featuring GARLI 2.0.** *Syst Biol*. 2014; **63**(5): 812–818.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ronquist F, Teslenko M, van der Mark P, *et al.*: **MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space.** *Syst Biol*. 2012; **61**(3): 539–42.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Franco-Sierra ND, Gómez-Zuluaga M, Díaz-Nieto JF, *et al.*: **PhyPipe: an automated pipeline for phylogenetic reconstruction from multilocus sequences [v1; not peer reviewed].** *F1000Res*. 2016; **5**(ISCB Com): 1609 (poster).
[Publisher Full Text](#)
- Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.** *Bioinformatics*. 2014; **30**(9): 1312–1313.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Sukumaran J, Holder MT: **DendroPy: a Python library for phylogenetic computing.** *Bioinformatics*. 2010; **26**(12): 1569–1571.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Arteaga-Figueroa LA, Sánchez-Bermúdez V, Franco-Sierra ND: **Revisiting the phylogeny of phylum Ctenophora: a molecular perspective.** *Zenodo*. 2016.
[Data Source](#)

Open Peer Review

Current Peer Review Status: ? × × ×

Version 1

Reviewer Report 03 April 2017

<https://doi.org/10.5256/f1000research.11235.r19434>

© 2017 Littlewood D. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

D Timothy J. Littlewood

Department of Life Sciences, Natural History Museum, London, UK

Unfortunately I must concur with the major concerns highlighted by other reviewers.

The data set is essentially a reassessment (meta-analysis) of previously published data. The phylogeny is functionally rooted against an in-group taxon without explanation. The analyses could have been improved by combining amino acid and nucleotide data (rather than solely treating these data separately). Due reference to similar articles from which these data have been derived was omitted. Other key references are missing. There is confusion within the article over the utility of some multi-copy genes and so interpretation and veracity of results is compromised.

In combination and considering the lack of sufficient novelty of data, approach or interpretation the publication falls short of achieving its goal. The phylogeny is revisited and with some investment of time from the authors, but little additional clarity and few insights are forthcoming to merit acceptance in its current state.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 28 February 2017

<https://doi.org/10.5256/f1000research.11235.r20287>

© 2017 Kocot K. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Kevin M. Kocot

Department of Biological Sciences, The University of Alabama, Tuscaloosa, AL, USA

This study is a metaanalysis of available ctenophore sequence data. My thoughts largely echo those of the previous two reviewers.

The methods seem reasonable (with the exception of the methodological problems with analysis of the Homiphora IPNS gene raised by Steve Haddock) but rooting with Beroida as the outgroup is inappropriate as no molecular studies have supported this in the past, key references are absent, and the English of the manuscript needs significant improvement.

Available ctenophore transcriptome data could be used to expand sampling of the protein-coding genes. If that were done, a concatenated analysis of all of the markers used with only taxa sampled for 18S (so all taxa overlap for at least part of the alignment) with the addition of appropriate outgroups would be an interesting improvement.

Competing Interests: I am also actively studying the evolution of Ctenophora but I can honestly say that this does not impact my view on the present work.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 13 February 2017

<https://doi.org/10.5256/f1000research.11235.r20040>

© 2017 Haddock S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Steven H.D. Haddock

Monterey Bay Aquarium Research Institute, Moss Landing, CA, USA

This article performs a meta-analysis of molecular phylogenetics within the Ctenophores using published sequences. I have previously had amiable correspondence with the authors and sent them data. Although the protein-coding sequences and some of the ribosomal RNA data came from my lab, I nonetheless feel I can give an unbiased assessment of their subsequent use.

Regrettably, I do not see enough original intellectual contribution or additional scientific value to justify its publication. At best, it is a minor contribution, in which case the interpretation needs to be improved, and at worst, a good portion of it is a re-publishing of work already published by other authors (Simion [1] and Podar [4] in particular).

A large part of the analysis is building trees using previously published ribosomal RNA datasets. This recapitulation does not add anything to the discussion of ctenophore internal relationships,

and in fact, by rooting the tree with *Beroë*, they obscure the true evolution of the group as shown repeatedly since at least 2001 [4]. They also fail to cite Simion, *et al.* [1], which is the source of some of the data. Merely adding that citation would not solve the fundamental issue, which is that there is no added value to their re-building of the same phylogeny.

The "novel" aspect of the paper is building trees based on two protein-coding genes which were also published previously in two separate papers [2,3]. The trees based on their [our] protein datasets do not give any additional insights into ctenophore relationships except in that some species are present in those trees that are not represented in the 18S phylogeny. This taxonomic coverage does not reveal any particular insight. These data also already appeared in trees (albeit not limited to ctenophores only) in the original publication.

There is some confusion because there is no *Hormiphora* IPNS gene in their [our] dataset, yet it is listed in the table of genes and that species is present in the tree, apparently based on a GFP-like gene that was found. Furthermore, the IPNS genes are not single-copy[2], so are not reliable for phylogeny building.

There is a misspelling of *Bathocytena* in Table 1 and 2 and of *Lampocteis* in Table 2.

In summary, two gene trees, of which one gene which was found to be absent in a ctenophore lineage, does not seem to be sufficient basis for a paper. The title itself is a vast overstatement of the content of this study.

References

1. Simion P, Bekkouche N, Jager M, Quéinnec E, et al.: Exploring the potential of small RNA subunit and ITS sequences for resolving phylogenetic relationships within the phylum Ctenophora. *Zoology (Jena)*. 2015; **118** (2): 102-14 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Francis WR, Shaner NC, Christianson LM, Powers ML, et al.: Occurrence of Isopenicillin-N-Synthase Homologs in Bioluminescent Ctenophores and Implications for Coelenterazine Biosynthesis. *PLoS One*. 2015; **10** (6): e0128742 [PubMed Abstract](#) | [Publisher Full Text](#)
3. Francis WR, Christianson LM, Powers ML, Schnitzler CE, et al.: Non-excitable fluorescent protein orthologs found in ctenophores. *BMC Evol Biol*. 2016; **16** (1): 167 [PubMed Abstract](#) | [Publisher Full Text](#)
4. Podar M, Haddock SH, Sogin ML, Harbison GR: A molecular phylogenetic framework for the phylum Ctenophora using 18S rRNA genes. *Mol Phylogenet Evol*. 2001; **21** (2): 218-30 [PubMed Abstract](#) | [Publisher Full Text](#)

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 08 February 2017

<https://doi.org/10.5256/f1000research.11235.r19979>

© 2017 Dohrmann M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Martin Dohrmann

Department of Earth & Environmental Sciences & GeoBio-Center, Ludwig-Maximilians-Universität München, Munich, Germany

Introduction

1st paragraph:

- "deep metazoans" is an odd term; also "Parazoa" is no longer accepted as a valid name. All 4 taxa are clearly metazoans; they are best summarized as "non-bilaterian metazoans".

- The citation after the 1st sentence is from 1999; this should be replaced with something more recent as a lot of research in this area has been done since then.

- While the statement of the 1st sentence is true, reconstructing the phylogenetic relationships within Ctenophora does not help much to solve these issues, i.e. finding the position of Ctenophora in the animal tree of life is a separate issue that this study is unable to address.

- The "previous work" cited in the 2nd sentence is very old. There is a study from 2015 (Simion *et al.*, *Zoology* 118: 102-114) that also reconstructed internal relationships of Ctenophora based on multigene analyses. It is crucial that the authors interpret their results in light of that study. It is actually quite puzzling that the paper is not cited, especially because the authors used sequences originally reported in Simion *et al.* (2015).

- The statement in the following sentence is highly debatable. As long as the phylogenetic position of ctenophores is not resolved (see e.g. Dohrmann & Wörheide 2013 *Integr. Comp. Biol.* 53: 503-511; Pisani *et al.* 2015 *PNAS* 112: 15402-15407), it is totally unclear how relevant they are to answering these questions.

- In the next sentence, "previous studies" should be replaced with "ctenophores".

- The following 2 sentences suggest that this paper represents the first multilocus analysis addressing internal phylogeny of Ctenophora. As mentioned above, this is not true. In this study, the authors used 2 protein-coding genes and the 28S gene in addition to the 18S and ITS/5.8S markers already used by Simion *et al.* (2015). This is what sets their study apart from the previous paper, and this has to be clearly communicated. The paper should focus on discussing differences to the results of Simion *et al.* in light of expanding the set of markers (but also addressing the different taxon sampling in the 2 studies).

- The abbreviation "MLSA" is introduced for "multilocus analysis" – what does the "S" stand for? Maybe it should read "multilocus sequence analysis"?

2nd paragraph:

- "ribosomal genes" should read "ribosomal RNA genes" (also elsewhere in the MS), since there are also genes coding for ribosomal proteins.

- "ortholog" should be replaced with "protein-coding" (also elsewhere in the MS), since ribosomal RNA genes are also orthologs.

3rd paragraph:

- I think the taxonomic overlap between the protein-coding and the ribosomal RNA datasets is sufficient to conduct a combined analysis, to infer a tree based on all the evidence simultaneously. As far as I recall, using mixed nucleotide and amino-acid data is possible with RAxML and MrBayes; alternatively, the protein-coding partition could be analyzed on nucleotide level (possibly excluding 3rd codon positions if they are oversaturated).

Methods

- It is unclear how ambiguously alignable regions were treated. These have to be excluded prior to analysis, but a quick glance at the concatenated matrices provided in the data supplement (concat_matrix and concat_prot_corrected) suggests otherwise. Difficult-to-align regions can bias phylogenetic inference, so this is an important point to address.

- Information about the lengths of loci and concatenated alignments should be given.

- Information about how the trees were rooted should be given in this section. In the Results section it is mentioned that Beroida was used as the outgroup to all other ctenophores. However, this is poorly justified. For example, Simion *et al.* (2015) found this group deeply nested within ctenophores. In general, I suggest following closely the methodological protocol of Simion *et al.* to make the 2 studies truly comparable.

3rd paragraph:

- Replace "is" with "are"

- The accession numbers are buried in some text files in the data supplement, which is quite inconvenient for the reader. I suggest providing them directly in Tables 1 and 2.

7th paragraph:

- Information about the substitution models used has to be provided here. I highly recommend using partitioning by gene and incorporating secondary structure information, for full comparability with Simion *et al.* (secondary structure models are available both in RAxML and MrBayes).

Results, Discussion

- These sections have to be rewritten after reanalysis of the data and comparison with Simion *et al.* (2015).

References

1. Simion P, Bekkouche N, Jager M, Quéinnec E, et al.: Exploring the potential of small RNA subunit and ITS sequences for resolving phylogenetic relationships within the phylum Ctenophora. *Zoology (Jena)*. 2015; **118** (2): 102-14 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Dohrmann M, Wörheide G: Novel scenarios of early animal evolution--is it time to rewrite textbooks?. *Integr Comp Biol*. 2013; **53** (3): 503-11 [PubMed Abstract](#) | [Publisher Full Text](#)
3. Pisani D, Pett W, Dohrmann M, Feuda R, et al.: Genomic data do not support comb jellies as the sister group to all other animals. *Proc Natl Acad Sci U S A*. 2015; **112** (50): 15402-7 [PubMed Abstract](#) | [Publisher Full Text](#)

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 12 Feb 2017

Nicolás D. Franco-Sierra, Universidad EAFIT, Colombia

Thanks for your comments and suggestions on our work, they are really helpful to improve our analysis. We agree with the observations you pointed out above and we are currently working on a revised version of our manuscript.

We sincerely apologize for not including the respective comparison with the work performed by Simion *et al* (2015). We are working on the readjustments in order to make our analysis fully comparable with Simion *et al* (2015).

Competing Interests: No competing interests were disclosed.

Comments on this article

Version 1

Reader Comment 05 Jan 2017

Paul Simion, Université Montpellier, France

Dear authors,

While I am always glad to see new studies on ctenophore phylogeny, I am very surprised that you did not cite Simion *et al*. 2014 (of which I am the first author) for two reasons :

1. You used all the data sequenced in that study.

2. Both study are very similar in topic and design, and should therefore be compared.

Please find a link to the study :

<http://www.sciencedirect.com/science/article/pii/S0944200614000816>

Sincerely,
Paul Simion

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research