



Comparative Genomics Analysis of Ciliates Provides Insights on the Evolutionary History Within “Nassophorea–Synhymenia–Phyllopharyngea” Assemblage

Bo Pan^{1,2†}, Xiao Chen^{3†}, Lina Hou^{4†}, Qianqian Zhang^{5†}, Zhishuai Qu^{1,6†}, Alan Warren⁷ and Miao Miao^{4*}

OPEN ACCESS

Edited by:

Ludmila Chistoserdova,
University of Washington,
United States

Reviewed by:

Mann Kyoon Shin,
University of Ulsan, South Korea
Zhenzhen Yi,
South China Normal University, China

*Correspondence:

Miao Miao
miaomiao@ucas.ac.cn

†These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 20 August 2019

Accepted: 20 November 2019

Published: 12 December 2019

Citation:

Pan B, Chen X, Hou L, Zhang Q,
Qu Z, Warren A and Miao M (2019)
Comparative Genomics Analysis
of Ciliates Provides Insights on
the Evolutionary History Within
“Nassophorea–Synhymenia–
Phyllopharyngea” Assemblage.
Front. Microbiol. 10:2819.
doi: 10.3389/fmicb.2019.02819

¹ Institute of Evolution and Marine Biodiversity, Ocean University of China, Qingdao, China, ² Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, ³ Department of Genetics and Development, Columbia University Medical Center, New York, NY, United States, ⁴ Savaid Medical School, University of Chinese Academy of Sciences, Beijing, China, ⁵ Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, China, ⁶ Ecology Group, Technical University of Kaiserslautern, Kaiserslautern, Germany, ⁷ Department of Life Sciences, Natural History Museum, London, United Kingdom

Ciliated protists (ciliates) are widely used for investigating evolution, mostly due to their successful radiation after their early evolutionary branching. In this study, we employed high-throughput sequencing technology to reveal the phylogenetic position of Synhymenia, as well as two classes Nassophorea and Phyllopharyngea, which have been a long-standing puzzle in the field of ciliate systematics and evolution. We obtained genomic and transcriptomic data from single cells of one synhymenian (*Chilodontopsis depressa*) and six other species of phyllopharyngeans (*Chilodochona* sp., *Dysteria derouxi*, *Hartmannula sinica*, *Trithigmostoma cucullulus*, *Trochilia petrani*, and *Trochilia* sp.). Phylogenomic analysis based on 157 orthologous genes comprising 173,835 amino acid residues revealed the affiliation of *C. depressa* within the class Phyllopharyngea, and the monophyly of Nassophorea, which strongly support the assignment of Synhymenia as a subclass within the class Phyllopharyngea. Comparative genomic analyses further revealed that *C. depressa* shares more orthologous genes with the class Nassophorea than with Phyllopharyngea, and the stop codon usage in *C. depressa* resembles that of Phyllopharyngea. Functional enrichment analysis demonstrated that biological pathways in *C. depressa* are more similar to Phyllopharyngea than Nassophorea. These results suggest that genomic and transcriptomic data can be used to provide insights into the evolutionary relationships within the “Nassophorea–Synhymenia–Phyllopharyngea” assemblage.

Keywords: ciliated protozoa, evolution, introns, phylogenetic relationship, stop codon usage

INTRODUCTION

Members of the order Synhymeniida are characterized by a cytopharyngeal basket of well-developed nematodesmata bound proximally by the presence of a band of somatic dikinetid cilia (Kivimaki et al., 2009; **Figure 1**). Corliss assigned the Synhymeniida to the subclass Hypostomata and placed the nassulids and microthoracids in a separate order (Nassulida) within the same subclass (Corliss, 1979; **Table 1**). Subsequently, Synhymeniida was assigned to the class Nassophorea because it possesses a cytopharyngeal basket (nasse), a character shared by the typical nassophorean groups (nassulids and microthoracids) (Small and Lynn, 1985; Lynn and Small, 2002; Lynn, 2008; **Table 1**). However, phylogenetic studies based mainly on the small subunit ribosomal DNA (SSU rDNA) gene and other related genes (e.g., LSU rDNA, ITS, alpha-tubulin), as well as a comparison of ultrastructural synapomorphies and autapomorphies, suggested that the subclass Synhymenia should be moved from the class Nassophorea to Phyllopharyngea (Gong et al., 2009; Zhang et al., 2014; Gao et al., 2016; **Table 1**). Based on a study of the ultrastructure of the typical synhymenian species *Zosterodasys agamalievi* Kivimaki et al. (1997) concluded that the synhymenians cannot be assigned to either Phyllopharyngea or Nassophorea, because they have a combination of characters

that precludes their inclusion in either. A recent study used a phylogenomic approach to test evolutionary relationships within the class Nassophorea, however, the dataset did not represent a broad sampling of all major ciliate clades (Lynn et al., 2018). Genomes of typical synhymeniids, for example, have yet to be reported (Lynn and Small, 2002).

The subclasses Cyrtophoria, Rhynchodia, Chonotrichia, and Suctoria each constitutes a clade within the monophyletic class Phyllopharyngea (de Puytorac, 1994; Lynn and Small, 1997; Lynn, 2008; **Table 1**). The chonotrichians, which are sessile symbionts typically found on the mouthparts of crustaceans, have long been recognized as phyllopharyngeans (Corliss, 1979; Lynn, 2016). Previous studies have revealed strong similarities between the cortical microtubular components of chonotrichians (represented by *Chilodochona*) with the cortical components of free-living cyrtophorians (Dobrzapnaska-Kaczanowska, 1963; Grain and Batisse, 1974; Fahrni, 1982). However, molecular phylogenies for chonotrichians were lacking (Gao et al., 2012, 2014). Cyrtophorians are characterized by a glabrous dorsal side and prominent cytopharyngeal basket composed of a large number of microtubules (Lynn, 2008). Although several cyrtophorians have been studied recently (Chen et al., 2016; Pan et al., 2016; Zhang et al., 2018), the systematics of this group remains ambiguous. Therefore, the class

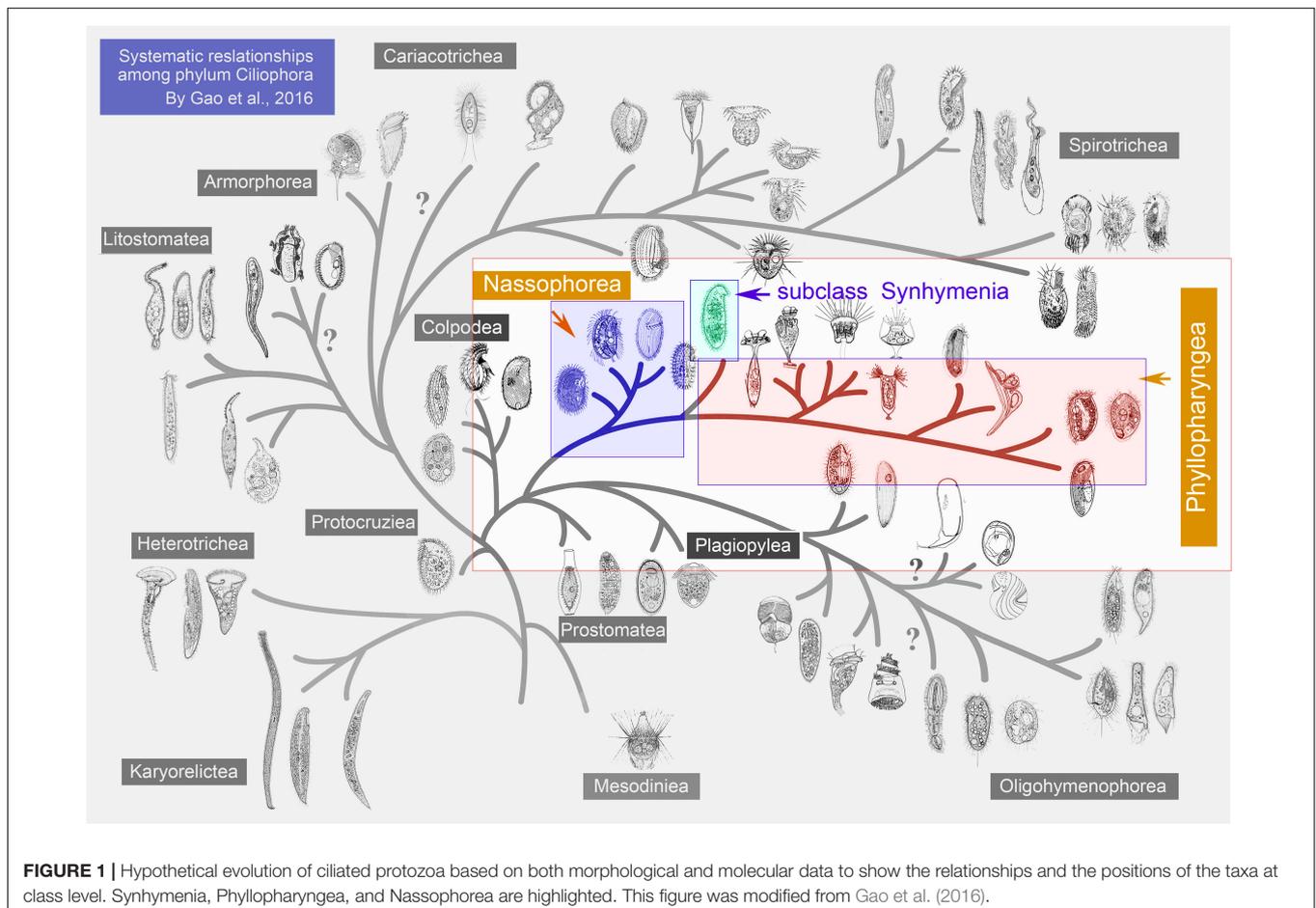


TABLE 1 | Taxonomic schemes for the classification of 10 species ciliates in four systematic schemes.

Corliss, 1979	de Puytorac, 1994	Lynn, 2008	Adl et al., 2019
Kinerofragminophora	Nassophorea	Nassophorea	Nassophorea
Hypostomata	Nassulina	Synhymeniida	Nassulida
Synhymeniida	Synhymeniida	Scaphidiodontidae	<i>Nassula</i>
<i>Chilodontopsis</i>	<i>Chilodontopsis</i>	<i>Chilodontopsis</i>	<i>Furgasonia</i>
Nassulida	Nassulida	Nassulida	Microthoracida
Nassulina	<i>Nassula</i>	Nassulidae	<i>Pseudomicrothorax</i>
Nassulidae	<i>Furgasonia</i>	<i>Nassula</i>	Phyllopharyngea
<i>Nassula</i>	Microthoracida	Furgasoniidae	Synhymeniida
Furgasoniidae	<i>Pseudomicrothorax</i>	<i>Furgasonia</i>	<i>Chilodontopsis</i>
<i>Furgasonia</i>	Phyllopharyngea	Microthoracida	Cyrtophoria
Microthoracina	Cyrtophoria	Leptopharyngidae	<i>Chilodonella</i>
Leptopharyngidae	Chilodonellida	<i>Pseudomicrothorax</i>	<i>Trithigmostoma</i>
<i>Pseudomicrothorax</i>	Chilodonellidae	Phyllopharyngea	<i>Dysteria</i>
Cyrtophorida	<i>Trithigmostoma</i>	Cyrtophoria	<i>Trochilia</i>
Chlamydodontina	<i>Chilodonella</i>	Chlamydodontida	<i>Hartmannula</i>
Chilodonellidae	Dysteriida	Chilodonellidae	Chonotrichia
<i>Chilodonella</i>	Hartmannulina	<i>Chilodonella</i>	<i>Chilodochona</i>
<i>Trithigmostoma</i>	Hartmannulidae	<i>Trithigmostoma</i>	
Dysteriina	<i>Hartmannula</i>	Dysteriida	
Hartmannulidae	Dysteriina	Dysteriidae	
<i>Hartmannula</i>	Dysteriidae	<i>Dysteria</i>	
Dysteriidae	<i>Dysteria</i>	<i>Trochilia</i>	
<i>Dysteria</i>	Trochiliidae	Hartmannulidae	
<i>Trochilia</i>	<i>Trochilia</i>	<i>Hartmannula</i>	
Chonotrichida		Chonotrichia	
Exogemmina		Exogemmida	
Chilodochonidae		Chilodochonidae	
<i>Chilodochona</i>		<i>Chilodochona</i>	

Phyllopharyngea needs to be re-evaluated based on more data. In the present study, genomic and transcriptomic data were used to analyze the evolutionary relationships among the synhymenians, nassophoreans, and phyllopharyngeans.

MATERIALS AND METHODS

Single-Cell Sample Preparation

Six freshwater ciliate species (*Chilodontopsis depressa*, *Trithigmostoma cucullulus*, *Trochilia petrani*, *Trochilia* sp., *Dysteria derouxi*, and *Hartmannula sinica*) were collected from Xiaoxihu Lake, Qingdao, China (120.35°E, 36.06°N). For each species, cells were washed in phosphate-buffered saline (PBS) buffer (without Mg²⁺ or Ca²⁺), and genomic DNA extracted from a single cell was amplified using a Single Cell WGA Kit (Yikon, YK001A) based on MALBAC technology (Lu et al., 2012). The RNA of *C. depressa* and *D. derouxi* was extracted from a single cell using the RNeasy kit (Qiagen, Hilden, Germany) and digested with DNase. The rRNA fraction was depleted using GeneRead rRNA Depletion Kit (Qiagen, Hilden, Germany). The raw data of single-cell genome sequencing of *Chilodochona* sp. were supplied by Dr. Denis H. Lynn (University of British Columbia, Vancouver, Canada).

High-Throughput Sequencing and Data Processing

Illumina libraries of 300 bp were prepared from amplified single-cell genomic DNA using Nextera DNA Flex Library Prep Kit by random PCR with primers (Illumina #20018704, United States) according to manufacturer's instructions. Paired-end sequencing (150 bp read length) was performed using an Illumina HiSeq4000 sequencer. The sequencing adapter was trimmed and low quality reads (reads containing >10% Ns or 50% bases with Q-value ≤ 5) were filtered out by Trimmomatic v0.36 (using parameters: LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) (Bolger et al., 2014).

Single-cell genomes of the seven species were assembled using SPAdes v3.11.1 (default parameters and $-k = 21, 33, 55, 77$) (Bankevich et al., 2012; Nurk et al., 2013). Mitochondrial genomic peptides of ciliates along with bacterial and human genome sequences were downloaded from NCBI GenBank as BLAST databases to remove contamination caused by mitochondria, bacteria, or factitious sources (BLAST + v2.6.0, E-value cutoff = $1e-5$). CD-HIT v4.7 (CD-HIT-EST, $-c 0.98 -n 10 -r 1$) was employed to eliminate the redundancy of contigs (with sequence identity threshold = 98%) (Fu et al., 2012). Poorly supported contigs (coverage < 1 and length < 400 bp) were discarded by a custom Perl script,

considering the inherent characteristics of the single-cell genomic sequencing technique (Zong et al., 2012). QUASt v4.6.1 was used to evaluate the resulting contigs ($-M = 0$) (Gurevich et al., 2013).

For the three nassophorean species, *Nassula variabilis*, *Furgasonia blochmanni*, and *Pseudomicrothorax dubius*, RNA sequences were obtained from the published dataset (NCBI accession number PRJNA434361). The adapters of reads were trimmed by Trimmomatic v0.36 (parameters as described above), and the clean reads were used for transcriptome assembly by Trinity v2.6.6 (default parameters) (Grabherr et al., 2011). Any redundancy and contamination of transcriptome assemblies were removed by BLAST and CD-HIT, as described above.

RNA-seq transcriptome data for another 21 ciliate species were downloaded from the Marine Microbial Eukaryote Transcriptome Sequencing Project¹ plus six ciliate species whose high-quality data were acquired from public databases (data sources are summarized in **Supplementary Table S1**).

Intron Length Distribution and Stop Codon Preference Detection

RNA-seq data of *C. depressa* were mapped to the corresponding reference genome assembly by HISAT2 v2.1.0 (Kim et al., 2015) and the transcriptome was assembled by StringTie v1.3.4d (Pertea et al., 2016). The size distribution of introns was analyzed by custom Perl scripts and the intronic boundary motif was visualized by WebLogo3 (Crooks et al., 2004). The frequency of stop codon usage was estimated by a custom Perl script which recognized the three nucleotides following the codon of the last amino acid in each matched complete protein sequences. This is identified as a potential stop codon by BLAST (E -value $\leq 1e-10$, identity $\geq 30\%$, alignment length ≥ 50 amino acids) transcript sequences of investigated species against ciliate protein database in Uniprot² (detail of method was uploaded to the GitHub repository: Bryan0425/Stop_codon_usage). Among different groups of ciliates, Student's t -test was performed in inner-class and outer-class.

Gene Prediction and Annotation

Transcriptomes were analyzed with TransDecoder v5.1.0³ (*Tetrahymena* genetic codes for *N. variabilis*, universal genetic codes for the other three species, *F. blochmanni*, *P. dubius*, and *C. depressa*) to search the longest ORFs, which generated protein predictions longer than 100 amino acids. To validate the potential ORFs, the predicted protein sequences were blasted to the Swiss-Prot database (E -value $< 1e-5$) (UniProt, 2018) and matched to the Pfam-A database (El-Gebali et al., 2018) by HMMER v3.1b2 (Eddy, 2009). Confirmed protein products were further annotated using four databases (CDD, Pfam, SUPERFAMILY, and TIGRFAM) by InterProScan v5.29 (Mitchell et al., 2018), and the ciliate gene database from NCBI GenBank by BLAST + v2.6.0 (E -value cutoff $< 1e-5$). RepeatMasker v4.0.7 (Smit, 1996) was employed to annotate non-coding RNA and repeat sequences.

¹<https://www.imicrobe.us>

²<https://www.uniprot.org/>

³<http://transdecoder.github.io/>

De novo gene prediction was performed by AUGUSTUS v3.3 ($-species = tetrahymena$, rearranging only TGA as stop codon, TAA/TAG as Glutamine) (Keller et al., 2011).

Ortholog Detection and GO Pathway

Orthologs among *C. depressa* and other ciliates were detected by a reciprocal BLAST hit (RBH) approach (E -value $< 1e-5$, identity $\geq 30\%$, alignment length ≥ 50 aa) and pairwise mutual best-hits were identified as putative orthologs. A Venn diagram was generated by the R package VennDiagram (Chen and Boutros, 2011). Gene Ontology (GO) term enrichment analysis was performed by BiNGO v3.0.3 (p.adjust < 0.05) (Maere et al., 2005), which was integrated in Cytoscape v3.4.0 (Shannon et al., 2003). The bubble plots were generated by the R package ggplot2 (Wickham, 2016).

Datasets and Alignments

Newly characterized sequences that combined relevant sequences obtained from the GenBank database were assembled into two datasets: (1) phylogenomics data (59 taxa in total), i.e., the 2 newly sequenced transcriptomes; 5 newly sequenced genomes; and 39 other ciliates, plus 9 apicomplexans and 4 dinoflagellates as outgroups; (2) SSU-rDNA database (56 taxa in total), i.e., 5 newly sequenced genes plus 3 heterotrichs and karyorelictids as outgroups. Sequences information and GenBank accession numbers are shown in **Supplementary Table S2**.

Phylogenetic and Phylogenomic Analyses

Small subunit ribosomal DNA sequences were multiple-aligned and trimmed to be blunt using BioEdit 7.1.3.0 (ClustalW algorithm) with the default parameters. Maximum-likelihood (ML) and Bayesian inference (BI) analyses were carried out on the online server CIPRES Science Gateway (Miller et al., 2010; Penn et al., 2010), using RAxML-HPC2 on XSEDE v8.2.10 (Stamatakis, 2014) with GTRGAMMA model and MrBayes on XSEDE v3.2.6 (Ronquist et al., 2012) with GAMMA model calculated by MrModeltest v2.3 (Nylander, 2004), respectively. In the ML analyses, 1,000 bootstraps were performed to assess the reliability of internal branches. For Bayesian analyses, Markov chain Monte Carlo (MCMC) simulations were run with two sets of four chains for 4,000,000 generations, with sampling every 100 generations and a burn-in of 10,000 (Chen et al., 2015). All the remaining trees were used to calculate posterior probabilities (PP) using a majority rule consensus. MEGA v7.0.20 (Kumar et al., 2016) was used to visualize the tree topologies.

The phylogenomic analyses were carried out by the GPSit pipeline (using parameters $-e 1e-10 -d 50 -g 100000 -f 100$ relaxed masking mode) based on 157 orthologs shared by all 46 ciliates utilizing a "supermatrix" approach (Chen et al., 2018b). The multiple sequence alignment was uploaded to CIPRES Science Gateway. RAxML-HPC2 v8.2.9 (using parameters: LG model of amino acid substitution + Γ distribution + F, four rate categories, 500 bootstrap replicates) and PHYLOBAYES MPI 1.5a (using parameters: CAT-GTR model + Γ distribution,

four independent chains, 10,000 generations for matrix of relaxed masking, or 20,000 generations for matrix of stringent mode, with first 10% of all generations as burn-in, convergence (Maxdiff < 0.3) were performed to generate the ML and BI trees, respectively (Le and Gascuel, 2008; Lartillot et al., 2009; Miller et al., 2010; Stamatakis, 2014; Gentekaki et al., 2017). All trees were visualized by MEGA v7.0.20. To visualize all available phylogenetic signals in the supermatrix, phylogenetic network analyses were calculated with SPLITSTREE v4.14.4 (using parameters: Network = NeighborNet; 1000 bootstrap replicates) (Huson and Bryant, 2005).

RESULTS

Morphology of Synhymenians, Nassophoreans, and Phyllopharyngeans

The main morphological characteristics of the synhymenians, nassophoreans, and phyllopharyngeans are summarized in **Figure 2A**. Their divergence/convergence has been described in detail and discussed previously (Hausmann and Peck, 1978; Eisler, 1988; Chen et al., 2018a). Briefly, synhymenians and nassophoreans have a synhymenium, which is a kinetic structure below the oral apparatus, mainly on the ventral surface. In synhymenians the synhymenium extends from left to right on the ventral surface, but in Nassophorea it is restricted to the left ventral surface. This serves as the main characteristic to distinguish these two groups. With the exception of the synhymenians, this structure is absent in the Phyllopharyngea.

Genomic Features of Classes Phyllopharyngea and Nassophorea

Although the newly sequenced ciliates have differing genome sizes and GC contents, most genome assemblies of phyllopharyngeans are about 40 M (**Table 2**). *Chilodontopsis depressa* has small, AT-rich introns, most of which are 36 bp (**Figure 2B**) with a GT-AG motif and a conservative third A site (**Figure 2C**). Also, of the non-coding RNA types predicted by RepeatMasker, 55% are tRNAs (171), 21% are snRNAs (65), 13% are rRNAs (40), and 11% are other RNAs (34) (**Figure 2D**). The predicted genes were annotated according to cellular component, molecular function, and biological process (**Figure 2E** and **Supplementary Figure S2**). Compared to *Tetrahymena thermophila*, *C. depressa* has a high gene content in its binding and protein-binding pathways, which might be associated with the molecular function of its well-developed cytopharyngeal basket of microtubules and its algivorous mode of nutrition (Tucker, 1968; Weisenberg et al., 1968; Chaaban and Brouhard, 2017).

Orthologous genes were identified among the seven species sequenced in the current work and other previously sequenced ciliates (**Figure 3** and **Supplementary Figure S1**). 703 orthologous genes shared by *Chilodontopsis depressa* and four model ciliates were identified (**Figure 3A**). *C. depressa* shared fewer orthologous genes with phyllopharyngeans (550) than with nassophoreans (1141), which is consistent with its closer

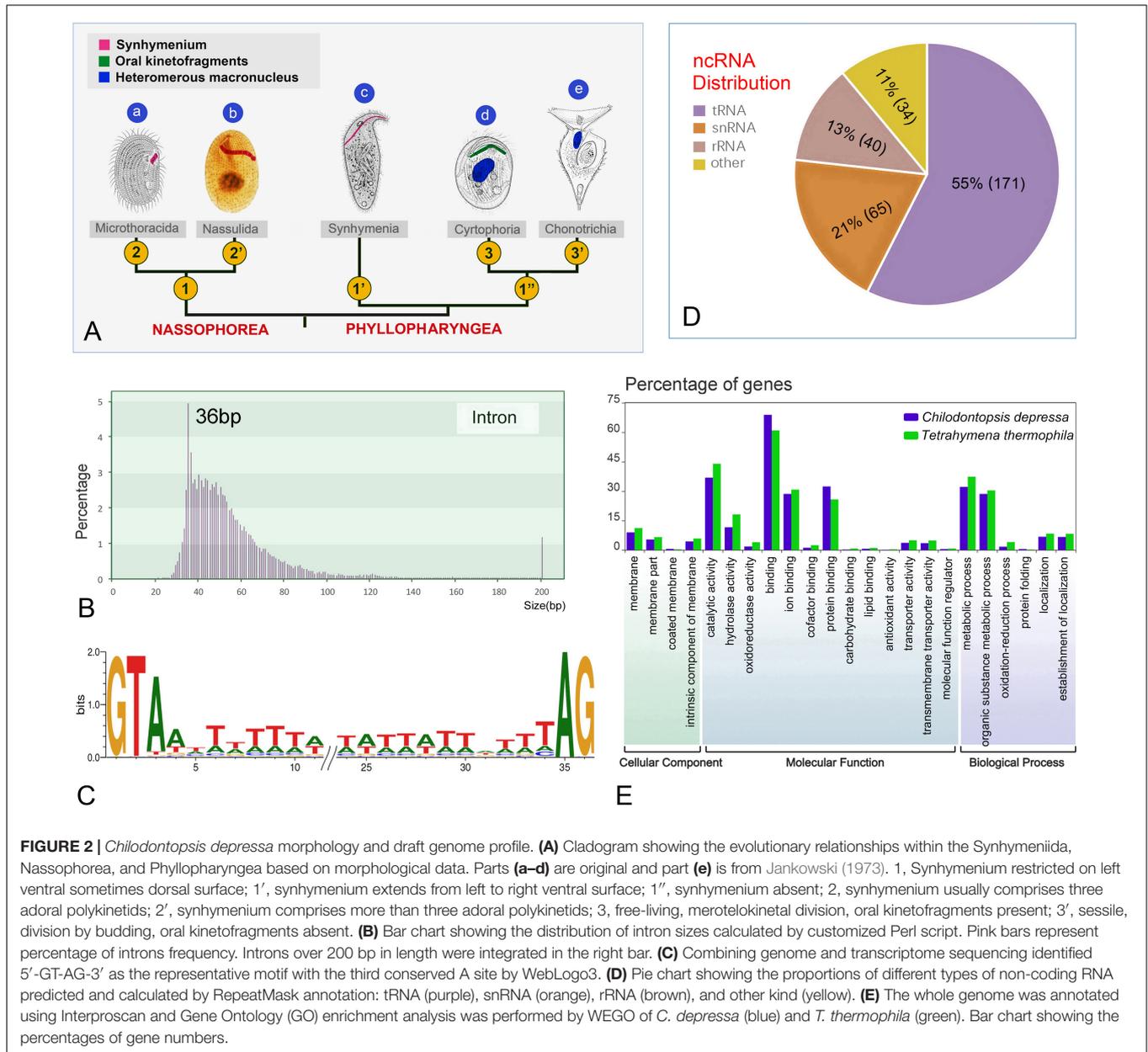
morphological resemblance to the latter, e.g., the presence (vs. absence) of a synhymenium (**Figures 2A, 3B,C**). Moreover, the GO term enrichment results suggested that *C. depressa* shared more metabolic features with phyllopharyngeans than with nassophoreans (**Figure 4A** and **Supplementary Figure S2**), such as regulation of cellular processes, regulation of biological processes, cellular response to stimulus, and biological regulation pathways. Such features tend to be enriched in nassophoreans not phyllopharyngeans.

Based on the homolog search between transcripts of *C. depressa* and protein sequences of other ciliates, the most frequently used stop codon in *C. depressa* was identified as TAA (91.3%, **Figure 4B**). Compared with model ciliates such as *T. thermophila* and *Oxytricha trifallax*, stop codon usages of other newly sequenced phyllopharyngeans and nassophoreans were also predicted. The results showed that the stop codon usage of *C. depressa* was closer to *Chilodonella uncinata* and *D. derouxi* (both of which belong to Phyllopharyngea) (deviation = 1.9) than to Nassophorea species (deviation = 25.6). The preference of stop codon usage was well-conserved within each class (p -value = 0.01538 < 0.05 by t -test).

Phylogenetic and Phylogenomic Analyses

The SSU rDNA dataset provided the most abundant taxon sampling. We carried out phylogenetic analyses based on 56 SSU-rDNA sequences (**Figure 5**). In the resulting trees, *C. depressa* grouped with the typical synhymenians, albeit in the basal position. The subclass Synhymenia grouped with the class Phyllopharyngea (or “Subkinetalia,” a name coined for the class Phyllopharyngea with synhymenians excluded), with high support values (ML/BI, 93/1.00). The monophyly of Subkinetalia is fully supported. The class Nassophorea is paraphyletic in the SSU rDNA tree, with the three orders (Microthoracida, Nassulida, and Discotrichida) separating to various positions within the CONthreeP assemblage (a robust clade comprising the classes Colpodea, Oligohymenophorea, Nassophorea, Prostomatea, Plagiopylea, and Phyllopharyngea) (Lynn, 2008; Adl et al., 2019).

Seven new genomes and two new transcriptomes were generated in the present study. The “omics” data source of orthologous genes for phylogenomic analyses has thus been increased to 46 ciliates, although *C. depressa* remains the only representative synhymenian (**Figure 6**). By using GPSit with relaxed mode (Chen et al., 2018b), we established a database with 157 genes comprising 173,835 amino acid residues. In the resulting phylogenomic tree using the ML algorithm, *C. depressa* clustered with the Subkinetalia with moderate support (ML, 78). The monophyly of the Subkinetalia was strongly supported by both ML and BI analyses. The branching pattern within Subkinetalia is in accordance to the SSU rDNA phylogeny inferences (**Figure 5**), although present taxon sampling is restricted to the subclass Cyrtophoria. For the class Nassophorea, transcriptome data are available for three species representing two orders, i.e., Nassulida and Microthoracida. The three species grouped together with a

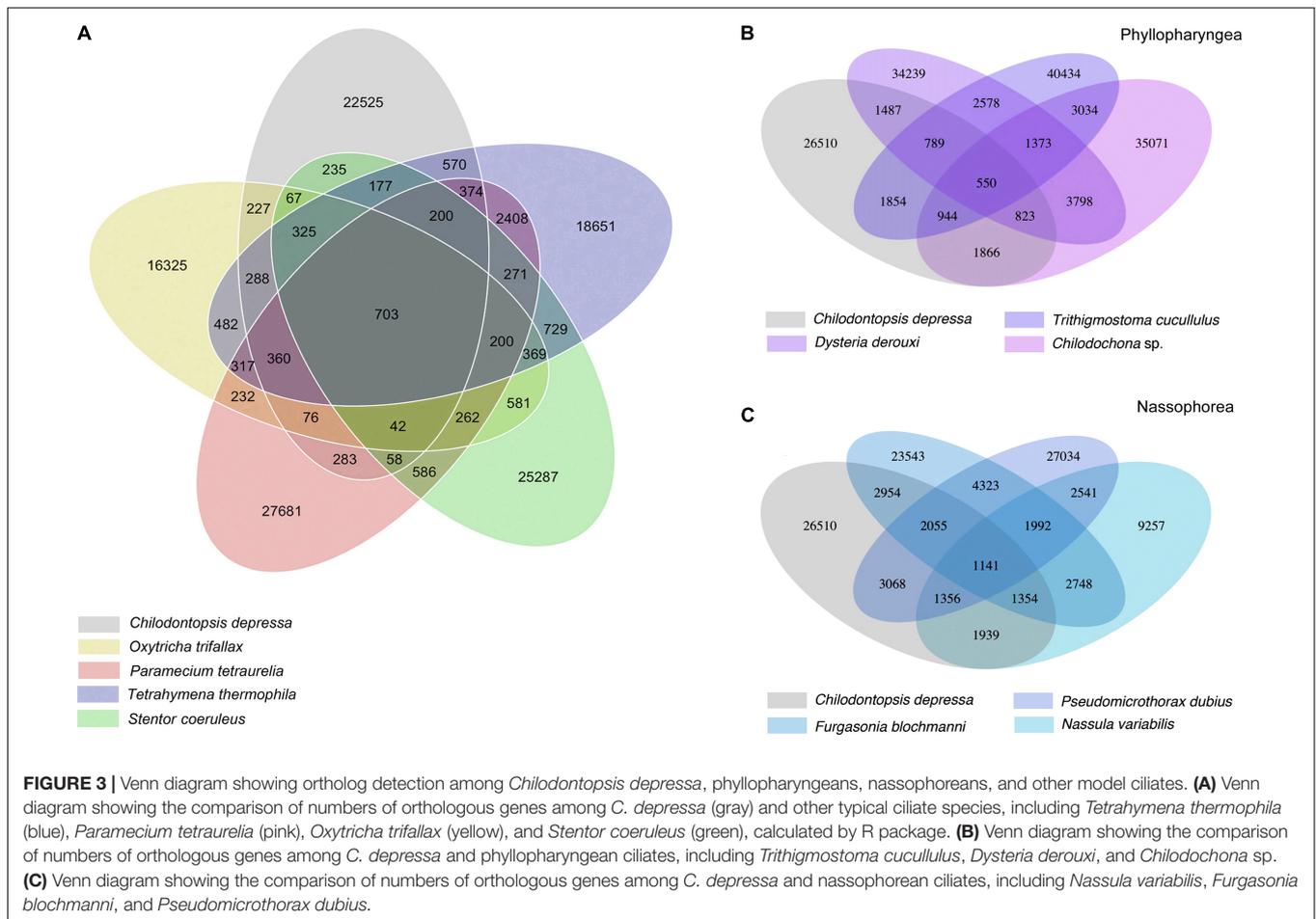


moderate support (ML/BI, 78/0.50). The classes Nassophorea and Colpodea formed a moderately supported clade in the ML tree, which subsequently grouped with the class

Oligohymenophorea. The class Phyllopharyngea was sister to the Colpodea + Nassophorea + Oligohymenophorea clade according to the ML algorithm.

TABLE 2 | Information of genome assembly from single-cell sequencing.

Statistics	<i>Chilodontopsis depressa</i>	<i>Trithigmostoma cucullulus</i>	<i>Trochilia petrani</i>	<i>Trochilia</i> sp.	<i>Chilodochona</i> sp.	<i>Dysteria derouxi</i>	<i>Hartmannula sinica</i>
Assembly size (bp)	129,702,864	48,714,649	41,378,687	119,916,114	45,682,069	46,996,486	41,739,095
Number of contigs	89,876	32,789	22,283	73,516	38,938	23,348	18,986
Largest contig	36,898	27,797	56,445	93,598	37,703	40,963	57,867
N50 (bp)	2,084	2,143	4,035	3,260	1,568	3,365	4,393
GC (%)	34.23	51.62	38.14	42.89	40.05	44.33	42.87



DISCUSSION

The Systematic Assignment of Synhymenians–Nassophoreans–Phyllopharyngeans Assemblage

The systematic position of synhymenians has long puzzled taxonomists, and their assignment to the class Nassophorea (nassulids and microthoracids) or to the Subkinetalia have been controversial. Synhymenians share ultrastructural and morphological similarities with both Nassophorea and Phyllopharyngea, suggesting that they are likely a transition group between these two classes (Eisler and Bardele, 1983; Small and Lynn, 1985; Kivimaki et al., 1997; Lynn and Small, 2002). A comprehensive comparison of morphology and ultrastructure among synhymenians, Nassophorea, and the Subkinetalia has previously addressed this issue (Gong et al., 2009). The SSU rDNA phylogeny (Figure 5) and previous studies support a close relationship between synhymenians and Subkinetalia, and thus synhymenians were suggested as a subclass (Synhymenia) within the class Phyllopharyngea (Zhang et al., 2014; Gao et al., 2016, 2017; Liao et al., 2018).

In this study we provide genomic data for a synhymenian, i.e., *Chilodontopsis depressa*, for the first time. *C. depressa* occupied

the basal position within the subclass Synhymenia (Figure 5), indicating that this species might be an ideal representative to demonstrate the ancestral candidate of the synhymenians. The phylogenomic analysis moderately supported (ML, 78) the affiliation of *C. depressa* to Subkinetalia, while the nassophoreans (represented by nassulids and microthoracids) formed another moderately supported monophyletic group (Figure 6). This is consistent with findings of previous studies based on SSU rDNA and multi-gene data which inferred that synhymenians are more closely related to Phyllopharyngea than to Nassophorea (Gong et al., 2009; Gao et al., 2016).

It is noteworthy that bootstrap support values for grouping *C. depressa* with Subkinetalia are lower in the phylogenomic analysis than those based on SSU rDNA data (ML, 78 vs. ML/BI, 93/1.00), whereas support for the monophyly of Subkinetalia are high in both analyses. These findings imply that, on a genomic scale, synhymenians and Subkinetalia are more divergent than indicated by the SSU rDNA gene (Figures 5, 6). This is consistent with the obvious ultrastructural differences between these two groups: in synhymenians, the postcilial microtubules of each kinetosome form a double row (vs. triads in cyrtophorians), and the synhymenian monokinetids lack both the subkinetal ribbons of microtubules and the dense transverse fibrils that characterize the phyllopharyngeans (Kivimaki et al., 1997). Moreover, in

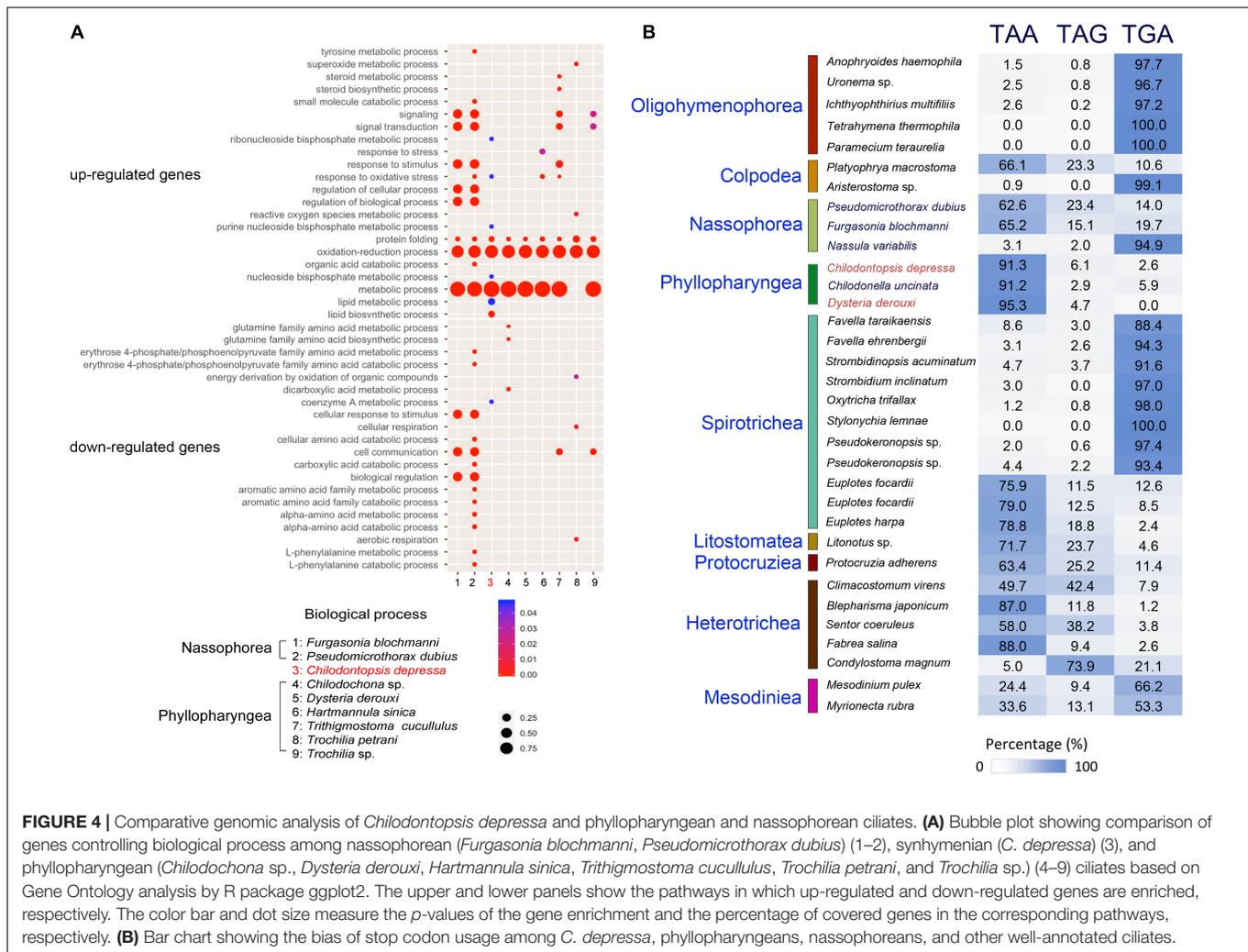


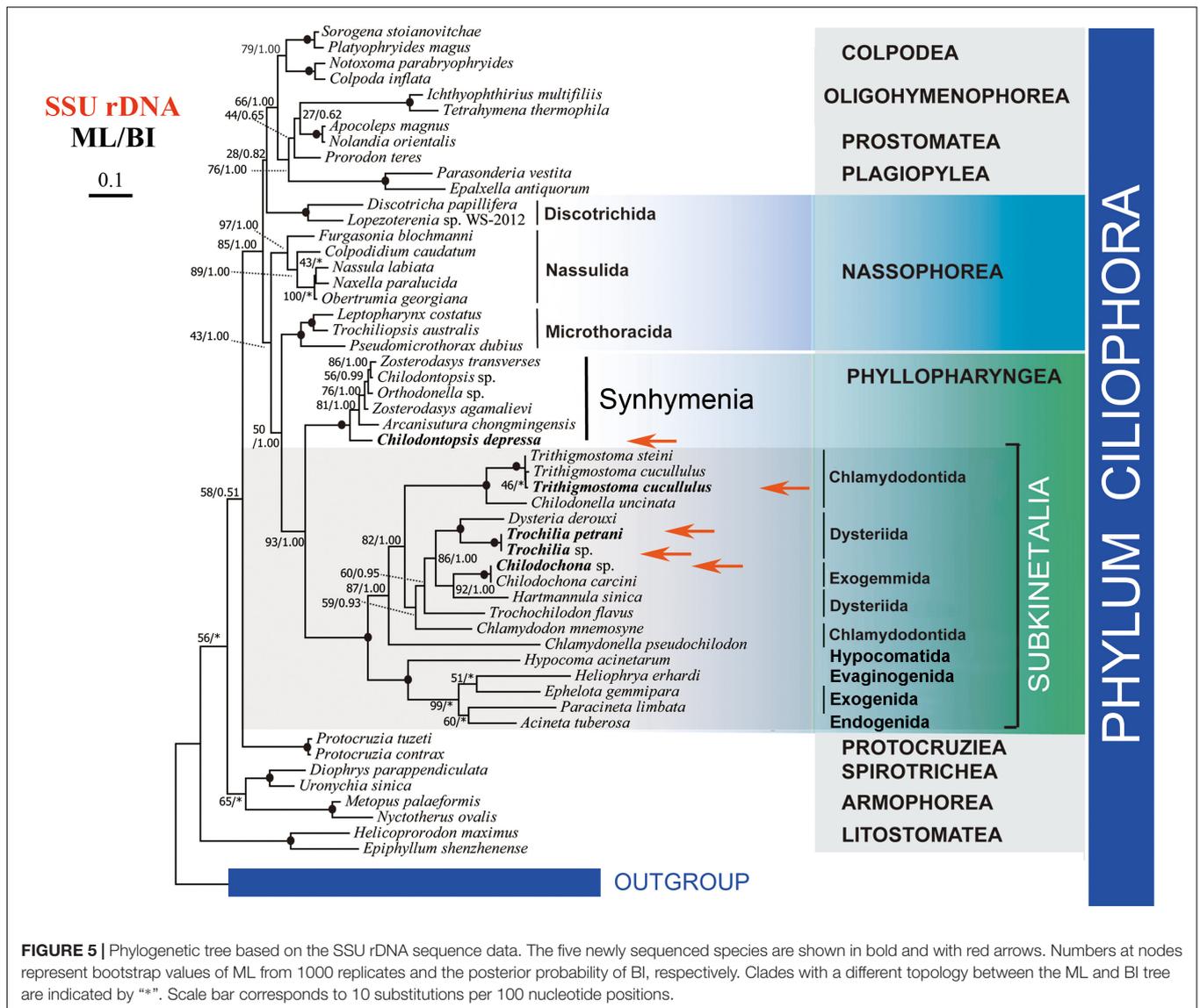
FIGURE 4 | Comparative genomic analysis of *Chilodontopsis depressa* and phyllopharyngean and nassophorean ciliates. **(A)** Bubble plot showing comparison of genes controlling biological process among nassophorean (*Furgasonia blochmanni*, *Pseudomicrothorax dubius*) (1–2), synhymenian (*C. depressa*) (3), and phyllopharyngean (*Chilodochona* sp., *Dysteria derouxi*, *Hartmannula sinica*, *Trithigmotoma cucullulus*, *Trochilia petrani*, and *Trochilia* sp.) (4–9) ciliates based on Gene Ontology analysis by R package ggplot2. The upper and lower panels show the pathways in which up-regulated and down-regulated genes are enriched, respectively. The color bar and dot size measure the p -values of the gene enrichment and the percentage of covered genes in the corresponding pathways, respectively. **(B)** Bar chart showing the bias of stop codon usage among *C. depressa*, phyllopharyngeans, nassophoreans, and other well-annotated ciliates.

Subkinetalia, there are some extremely specialized Subkinetalian groups, such as the suctorians and chonotrichians (Figure 2A), the adults of which are sessile and divide by budding. Collectively, these data indicate that there was a long history, and that a number of evolutionary events occurred, after the synhymenians and Subkinetalia diverged from their common ancestor.

The phylogenomic tree recovered only low to moderate support for the monophyly of Nassophorea (i.e., orders Nassulida and Microthoracida, with Discotrichida excluded). This is in accordance to a recent phylogenomic study of the nassulids (Lynn et al., 2018) and multi-gene analyses of the major ciliate lineages (Gao et al., 2016). The cytopharyngeal basket of nassulids is composed of one or more out of three sets of microtubular lamellae, namely cytostomal (Z), subcytostomal (Y), and nematodesmal (X) lamellae (Hausmann and Peck, 1978; de Puytorac and Njiné, 1980; Eisler, 1988). Microthoracids (e.g., *Pseudomicrothorax* and *Leptopharynx*) have only X lamella (Hausmann and Peck, 1978; de Puytorac and Njiné, 1980; Eisler, 1988), synhymenians have only Z lamella (Kivimaki et al., 1997), and cyrtophorians have only Z and Y lamellae (Kurth and Bardele, 2001). Controversial scenarios have been proposed to

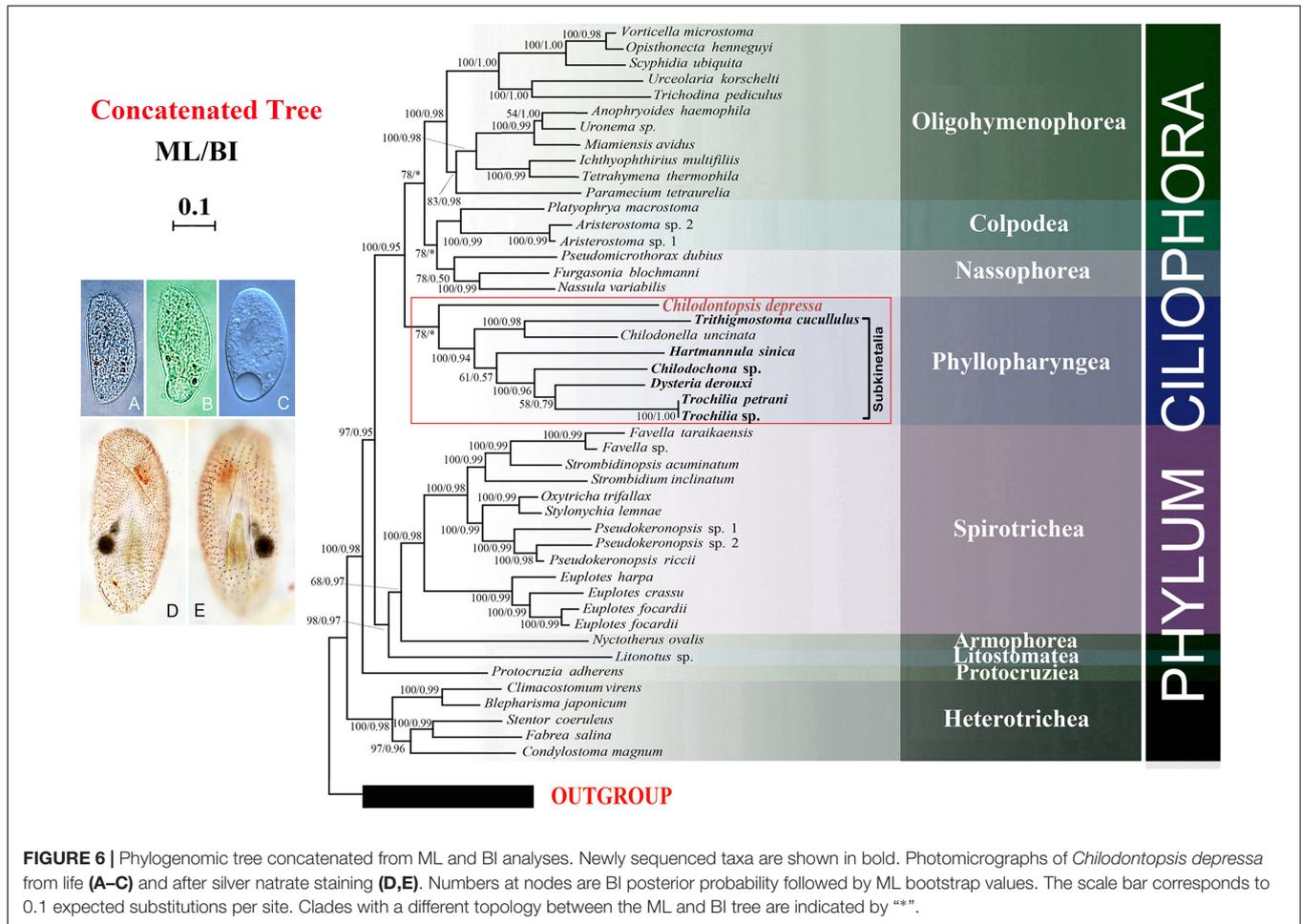
deduce the evolution of the cytopharyngeal basket apparatus based on phylogenetic (Gong et al., 2009) and phylogenomic analyses (Lynn et al., 2018). Inclusion of the new genomic data for the synhymenians in the analyses results in partial support for the previous hypothesis: synhymenians and Subkinetalia are members of the class Phyllopharyngea not Nassophorea (Gong et al., 2009; Gao et al., 2016). According to a scenario of morphological evolution deduced by Lynn et al. (2018), if the present phylogenomic tree is a probable representation of the order of divergence of these groups, i.e., a cytopharyngeal basket with microtubular nematodesmata and with only Z microtubular ribbons would probably be an ancestral feature of the CONthreeP main clade. A loss of Z lamella might have happened in the microthoracids, while the X and Y lamellae in nassophoreans and phyllopharyngeans could be apomorphies acquired later to enable the rapid ingestion of different food resources (Tucker, 1968; Hausmann and Peck, 1978; Lynn et al., 2018).

Nowadays, SSU rDNA is a powerful phylogenetic marker to resolve many evolutionary relationships, although it has its limitations (Hasegawa and Hashimoto, 1993). For example, SSU rDNA sequences that are similar in nucleotide composition



might have been placed incorrectly close together in phylogenetic trees (Woese et al., 1991). In addition, because of random variation of single genes, it is risky to infer the phylogeny from any single gene. Thus, it is suggested that single gene phylogenetic analyses should be corroborated by use of other phylogenetic markers (Ludwig and Klenk, 2001). Following the development of illumina sequencing, phylogenomic analysis is used as the gold standard to estimate evolutionary relationships on the “genome level” (Brown et al., 2001; Ciccarelli et al., 2006; Wu and Eisen, 2008). Phylogenomic data are usually expected to be well resolved without many limitations of small data sets, e.g., relatively less informative characters from independent loci and random noise (Delsuc et al., 2005; Jeffroy et al., 2006). Among the present analysis, 157 proteins sequences of 59 species were used to constructed the phylogenomic tree. Therefore, a more robust estimation of evolutionary relationships could be obtained. On the other hand, our results show

that on a genomic scale, there are more variations indicated than those on single gene phylogenetic analysis, like SSU rDNA (Figures 5, 6). Phylogenomic analysis applies much more data from whole genome/transcriptome compared to a specific one gene. Furthermore, proteins sequences exhibit more differences and variations among species when used to as a material of analysis. Some statistical support values in the tree were relatively high, but some were low or not as high as that in single-gene phylogenetic analysis, e.g., 58/0.79 versus full value at the branch of *Dysteria* and *Trochilia*. Higher support values might be achieved if the number of informative sites is inadequate or the key species are not included in the analysis. Furthermore, gene selection of phylogenomic analysis is an important process. The orthologs must to be considered in analysis, especially phylogenetic markers such as *HSP70* and tubulin (Baldauf et al., 2000; Ludwig and Klenk, 2001).



Tiny Introns Characterize Ciliate Genomes as Class Level

All eukaryotic genomes that have been sequenced harbor introns, but so far none have been identified in prokaryotes. Intron size varies in genomes of different species, e.g., from 169 bp in the model flowering plant *Arabidopsis thaliana* to 7,386 kb on average in humans (Jo and Choi, 2015; Piovesan et al., 2015). Previous studies of human genomes have suggested that intron size may play an important role in governing alternative splicing and sensing self/foreign circular RNAs (Fox-Walsh et al., 2005; Dewey et al., 2006; Kim et al., 2006; Chen et al., 2017). Furthermore, the minimum length of introns required for the splicing reaction and other advanced biological functions is thought to be 30 bp (Roy et al., 2008; Piovesan et al., 2015). However, smaller introns (15 bp) have been reported in heterotrich ciliates (e.g., *Stentor coeruleus* and *Condylostoma magnum*) (Slabodnick et al., 2017). Tiny introns have also been identified in the classes Armophorea (*Nyctotherus ovalis*: 27 bp) (McGrath et al., 2007; Ricard et al., 2008), Spirotrichea (*O. trifallax*: 38 bp) (Swart et al., 2013), and Oligohymenophorea (*T. thermophila*: 74 bp and *Paramecium* species: 26 bp) (Swart et al., 2013; Arnaiz et al., 2017). The present study reports for the first time tiny introns (36 bp) in the genome of species of the class Phyllopharyngea (Figures 2A,B).

Thus, based on the present and previous studies, tiny introns are another prevalent important feature of ciliate genomes just like the variation of rDNA copy numbers (Wang et al., 2019). Whether the small size of these introns impair their biological functions, governing alternative splicing for instance, requires further exploration.

Stop Codon Usage in Different Groups of Ciliates

Stop codon preference in the terminal region has been intensively investigated by evolutionary biologists (Alff-Steinberger and Epstein, 1994; Smith and Smith, 1996; Jungreis et al., 2011). Previous studies have reported the flexibility of the nuclear genetic code in ciliates by demonstrating that standard stop codons are reassigned to amino acids (Lozupone et al., 2001; Heaphy et al., 2016; Swart et al., 2016). One possible reason is that ambiguous genetic codes enabled the ancestors of ciliates to thrive during a certain periods in their evolutionary history (Swart et al., 2016). To determine how the stop codon rearrangement is conserved in different lineages, we systematically assessed the preference of stop codon usage in 33 ciliates based on their transcriptome data (Figure 4B). Our results show that stop codon usage varies among different

classes and, in some cases, within a class. For example, the stop codon usage of *C. depressa* (formerly classified in the class Nassophorea, order Synhymeniida) is more similar to that of *C. uncinata* (class Phyllopharyngea, order Chlamyodontida) and *D. derouxi* (class Phyllopharyngea, order Dysteriida) than to nassophoreans (e.g., *N. variabilis*, *P. dubius*, and *F. blochmanni*), suggesting Synhymeniida and other phyllopharyngeans may have a common recent ancestor. The close relationship between Synhymeniida and Phyllopharyngea is also supported by the results of phylogenomic analysis (Figure 6). We therefore posit that the preference of stop codon usage can be adopted as an extra parameter to resolve phylogenetic relationships. However, how the preference of stop codon usage is inherited remains poorly understood and should be addressed in future studies.

DATA AVAILABILITY STATEMENT

All Illumina sequencing datasets are deposited in the NCBI under the Accession Number PRJNA546036.

AUTHOR CONTRIBUTIONS

MM, BP, and XC conceived the study. BP and XC analyzed the data and interpreted the results using bioinformatics methods. LH, QZ, and ZQ performed the phylogenetics analysis and morphology identification and description. AW revised the writing of the manuscript. All authors contributed to the manuscript and agreed on the manuscript before review.

FUNDING

This work was supported by the Marine S&T Fund of Shandong Province for Pilot National Laboratory for Marine Science and

Technology (Qingdao) (2018SDKJ0406-2), the Natural Science Foundation of China (31672279, 31672251), the Natural Science Foundation of Shandong Province (JQ201706), the University of Chinese Academy of Sciences (Y8540XX1W2), and the China Scholarship Council.

ACKNOWLEDGMENTS

Our thanks are due to: Prof. Weibo Song (Ocean University of China, China) for his kind help and advice on preparing the manuscript; Dr. Denis H. Lynn (University of British Columbia, Canada) for supplying genomic data for *Chilodochona* sp.; and Prof. Fangqing Zhao (Chinese Academy of Sciences, China) for institutional support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.02819/full#supplementary-material>

FIGURE S1 | Molecular information of *Chilodontopsis depressa* and phyllopharyngean ciliates. Distribution of species with best hits by BLAST. The X-axis represents percentage of contigs or genes. (A) *Hartmannula sinica*, (B) *Chilodochona* sp., (C) *Dysteria derouxi*, (D) *Trithigmostoma cucullulus*, (E) *Trochilia petrani*, (F) *Trochilia* sp., and (G) *Chilodontopsis depressa*.

FIGURE S2 | Comparative genomic analysis of *Chilodontopsis depressa* and phyllopharyngean and nassophorean ciliates. Bubble plot showing comparison of genes controlling cell component (A) and molecular function (B) among nassophorean (*Furgasonia blochmanni*, *Pseudomicrothorax dubius*) (1,2), synhymenian (*Chilodontopsis depressa*) (3), and phyllopharyngean (*Chilodochona* sp., *Dysteria derouxi*, *Hartmannula sinica*, *Trithigmostoma cucullulus*, *Trochilia petrani* and *Trochilia* sp.) (4–9) ciliates based on Gene Ontology analysis by R package ggplot2.

REFERENCES

- Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., et al. (2019). Revisions to the classification, nomenclature, and diversity of eukaryotes. *J. Eukaryot. Microbiol.* 66, 4–119. doi: 10.1111/jeu.12691
- Alff-Steinberger, C., and Epstein, R. (1994). Codon preference in the terminal region of *E. coli* genes and evolution of stop codon usage. *J. Theor. Biol.* 168, 461–463. doi: 10.1006/jtbi.1994.1124
- Arnaiz, O., Van Dijk, E., Bétermier, M., Lhuillier-Akakpo, M., de Vanssay, A., Duharcourt, S., et al. (2017). Improved methods and resources for *Paramecium* genomics: transcription units, gene annotation and gene expression. *BMC Genomics* 18:483. doi: 10.1186/s12864-017-3887-z
- Baldauf, S. L., Roger, A., Wenk-Siefert, I., and Doolittle, W. F. (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* 290, 972–977. doi: 10.1126/science.290.5493.972
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Brown, J. R., Douady, C. J., Italia, M. J., Marshall, W. E., and Stanhope, M. J. (2001). Universal trees based on large combined protein sequence data sets. *Nat. Genet.* 28, 281–285. doi: 10.1038/90129
- Chaaban, S., and Brouhard, G. J. (2017). A microtubule bestiary: structural diversity in tubulin polymers. *Mol. Biol. Cell* 28, 2924–2931. doi: 10.1091/mbc.E16-05-0271
- Chen, H., and Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 12:35. doi: 10.1186/1471-2105-12-35
- Chen, X., Li, L., Al-Farraj, S. A., Ma, H., and Pan, H. (2018a). Taxonomic studies on *Aegyria apoliva* sp. nov. and *Trithigmostoma cucullulus* (Müller, 1786) Jankowski, 1967 (Ciliophora, Cyrtophoria) with phylogenetic analyses. *Eur. J. Protistol.* 62, 122–134. doi: 10.1016/j.ejop.2017.12.005
- Chen, X., Wang, Y., Sheng, Y., Warren, A., and Gao, S. (2018b). GPSit: an automated method for evolutionary analysis of nonculturable ciliated microeukaryotes. *Mol. Ecol. Resour.* 18, 700–713. doi: 10.1111/1755-0998.12750
- Chen, X., Pan, H., Huang, J., Warren, A., Al-Farraj, S. A., and Gao, S. (2016). New considerations on the phylogeny of cyrtophorian ciliates (Protozoa, Ciliophora): expanded sampling to understand their evolutionary relationships. *Zool. Scr.* 45, 334–348. doi: 10.1111/zsc.12150

- Chen, X., Zhao, X., Liu, X., Warren, A., Zhao, F., and Miao, M. (2015). Phylogenomics of non-model ciliates based on transcriptomic analyses. *Protein Cell* 6, 373–385. doi: 10.1007/s13238-015-0147-143
- Chen, Y. G., Kim, M. V., Chen, X., Batista, P. J., Aoyama, S., Wilusz, J. E., et al. (2017). Sensing self and foreign circular RNAs by intron identity. *Mol. Cell* 67, 228–238. doi: 10.1016/j.molcel.2017.05.022
- Ciccarelli, F. D., Doerks, T., Von Mering, C., Creevey, C. J., Snel, B., and Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science* 311, 1283–1287. doi: 10.1126/science.1123061
- Corliss, J. (1979). *The Ciliated Protozoa. Characterization, Classification and Guide to the Literature*. Oxford: Pergamon Press.
- Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- de Puytorac, P. (1994). "Phylum ciliophora doflein, 1901," in *Traité de Zoologie, Tome II, Infusoires Ciliés, Fasc. 2, Systématique*. Ed. P. de Puytorac., (Paris: Masson) 1–15.
- de Puytorac, P., and Njiné, T. (1980). A propos des ultrastructures corticales et buccales du cilié hypostome *Nassula tumida* Maskell, 1887. *Protistologica* 16, 315–327.
- Delsuc, F., Brinkmann, H., and Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361–75. doi: 10.1038/nrg1603
- Dewey, C. N., Rogozin, I. B., and Koonin, E. V. (2006). Compensatory relationship between splice sites and exonic splicing signals depending on the length of vertebrate introns. *BMC Genomics* 7:311. doi: 10.1186/1471-2164-7-311
- Dobrzaprnska-Kaczanowska, J. (1963). Comparaison de la morphogenese des ciliés: *Chilodonella uncinata* (Ehrbg.), *Allosphaerium paraconvexa* sp. n. et *Heliochona scheuteni* (Stein). *Acta Protozoologica* 1, 353–394.
- Eddy, S. R. (2009). A new generation of homology search tools based on probabilistic inference. *Genome Inform.* 23, 205–211.
- Eisler, K. (1988). Electron microscopical observations on the ciliate *Furgasonia blochmanni* Fauré-Fremiet, 1967: part I: an update on morphology. *Eur. J. Protistol.* 24, 75–93. doi: 10.1016/s0932-4739(88)80012-80019
- Eisler, K., and Bardele, C. (1983). The alveolocysts of the Nassulida: ultrastructure and some phylogenetic considerations. *Protistologica* 19, 95–102.
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2018). The Pfam protein families database in 2019. *Nucleic Acids Res.* 47, D427–D432. doi: 10.1093/nar/gky995
- Fahrni, J. F. (1982). Morphologie et ultrastructure de *Spirochona gemmipara* Stein, 1852 (Ciliophora, Chonotrichida). I. Structures corticales et buccales de l'adulte 1. *J. Protozool.* 29, 170–184. doi: 10.1111/j.1550-7408.1982.tb04009.x
- Fox-Walsh, K. L., Dou, Y., Lam, B. J., Hung, S.-P., Baldi, P. F., and Hertel, K. J. (2005). The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *PNAS* 102, 16176–16181. doi: 10.1073/pnas.0508489102
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Gao, F., Huang, J., Zhao, Y., Li, L., Liu, W., Miao, M., et al. (2017). Systematic studies on ciliates (Alveolata, Ciliophora) in China: progress and achievements based on molecular information. *Eur. J. Protistol.* 61, 409–423. doi: 10.1016/j.ejop.2017.04.009
- Gao, F., Song, W., and Katz, L. A. (2014). Genome structure drives patterns of gene family evolution in ciliates, a case study using *Chilodonella uncinata* (Protista, Ciliophora, Phyllopharyngea). *Evolution* 68, 2287–2295. doi: 10.1111/evo.12430
- Gao, F., Warren, A., Zhang, Q., Gong, J., Miao, M., Sun, P., et al. (2016). The all-data-based evolutionary hypothesis of ciliated protists with a revised classification of the phylum Ciliophora (Eukaryota, Alveolata). *Sci. Rep.* 6:24874. doi: 10.1038/srep24874
- Gao, S., Huang, J., Li, J., and Song, W. (2012). Molecular phylogeny of the cyrtophorid ciliates (Protozoa, Ciliophora, Phyllopharyngea). *PLoS One* 7:e33198. doi: 10.1371/journal.pone.0033198
- Gentekaki, E., Kolisko, M., Gong, Y., and Lynn, D. (2017). Phylogenomics solves a long-standing evolutionary puzzle in the ciliate world: the subclass Peritrichia is monophyletic. *Mol. Phylog. Evol.* 106, 1–5. doi: 10.1016/j.ympev.2016.09.016
- Gong, J., Stoeck, T., Yi, Z., Miao, M., Zhang, Q., Roberts, D. M., et al. (2009). Small subunit rRNA phylogenies show that the class Nassophorea is not monophyletic (Phylum Ciliophora). *J. Eukaryot. Microbiol.* 56, 339–347. doi: 10.1111/j.1550-7408.2009.00413.x
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Grain, J., and Batisse, A. (1974). Étude ultrastructurale du cilié chonotriche *Chilodochona quennerstedti* Wallengren, 1895. I. Cortex et structures buccales. *J. Protozool.* 21, 95–111. doi: 10.1111/j.1550-7408.1974.tb03621.x
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Hasegawa, M., and Hashimoto, T. (1993). Ribosomal RNA trees misleading? *Nature* 361:23. doi: 10.1038/361023b0
- Hausmann, K., and Peck, R. K. (1978). Microtubules and microfilaments as major components of a phagocytic apparatus: the cytopharyngeal basket of the ciliate *Pseudomicrothorax dubius*. *Differentiation* 11, 157–167. doi: 10.1111/j.1432-0436.1978.tb00979.x
- Heaphy, S. M., Mariotti, M., Gladyshev, V. N., Atkins, J. F., and Baranov, P. V. (2016). Novel ciliate genetic code variants including the reassignment of all three stop codons to sense codons in *Condylostoma magnum*. *Mol. Biol. Evol.* 33, 2885–2889. doi: 10.1093/molbev/msw166
- Huson, D. H., and Bryant, D. (2005). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030
- Jankowski, A. W. (1973). Taxonomic revision of subphylum Ciliophora Doflein, 1901. *Zool. Zh.* 52, 165–175.
- Jeffroy, O., Brinkmann, H., Delsuc, F., and Philippe, H. (2006). Phylogenomics: the beginning of incongruence? *Trends Genet.* 22, 225–231. doi: 10.1016/j.tig.2006.02.003
- Jo, B.-S., and Choi, S. S. (2015). Introns: the functional benefits of introns in genomes. *Genomics Inform.* 13, 112–118. doi: 10.5808/gi.2015.13.4.112
- Jungreis, I., Lin, M. F., Spokony, R., Chan, C. S., Negre, N., Victorsen, A., et al. (2011). Evidence of abundant stop codon readthrough in *Drosophila* and other metazoa. *Genome Res.* 21, 2096–2113. doi: 10.1101/gr.119974.110
- Keller, O., Kollmar, M., Stanke, M., and Waack, S. (2011). A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* 27, 757–763. doi: 10.1093/bioinformatics/btr010
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Kim, E., Magen, A., and Ast, G. (2006). Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res.* 35, 125–131. doi: 10.1093/nar/gkl924
- Kivimäki, K. L., Bowditch, B. M., Riordan, G. P., and Lipscomb, D. L. (2009). Phylogeny and systematic position of *Zosterodasys* (Ciliophora, Synhymeniida): a combined analysis of ciliate relationships using morphological and molecular data. *J. Eukaryot. Microbiol.* 56, 323–338. doi: 10.1111/j.1550-7408.2009.00403.x
- Kivimäki, K. L., Riordan, G. P., and Lipscomb, D. (1997). The ultrastructure of *Zosterodasys agamaliievi* (Ciliophora: Synhymeniida). *J. Eukaryot. Microbiol.* 44, 226–236. doi: 10.1111/j.1550-7408.1997.tb05705.x
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Kurth, T., and Bardele, C. F. (2001). Fine structure of the cyrtophorid ciliate *Chlamydonon mnemosyne* Ehrenberg, 1837. *Acta Protozool.* 40, 33–47.
- Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25, 2286–2288. doi: 10.1093/bioinformatics/btp368
- Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. doi: 10.1093/molbev/msn067
- Liao, W., Fan, X., Zhang, Q., Xu, Y., and Gu, F. (2018). Morphology and phylogeny of two novel ciliates, *Arcanisutura chongmingensis* n. gen., n. sp. and *Naxella paralucida* n. sp. from Shanghai, China. *J. Eukaryot. Microbiol.* 65, 48–60. doi: 10.1111/jeu.12431
- Lozupone, C. A., Knight, R. D., and Landweber, L. F. (2001). The molecular basis of nuclear genetic code change in ciliates. *Curr. Biol.* 11, 65–74. doi: 10.1016/s0960-9822(01)00028-28

- Lu, S., Zong, C., Fan, W., Yang, M., Li, J., Chapman, A. R., et al. (2012). Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science* 338, 1627–1630. doi: 10.1126/science.1229112
- Ludwig, W., and Klenk, H. (2001). “A phylogenetic backbone and taxonomic framework for prokaryotic systematics.” In *The Archaea and the Deeply Branching and Phototrophic Bacteria*. eds DR. Boone., and RW. Castenholz, (New York, NY: Springer-Verlag).
- Lynn, D. H. (2008). *The Ciliated Protozoa: Characterization, Classification, and Guide to the Literature*. Berlin: Springer.
- Lynn, D. H. (2016). The small subunit rRNA gene sequence of the chonotrich *Chilodochona carcini* Jankowski, 1973 confirms chonotrichs as a dysteriid-derived clade (Phyllopharyngea, Ciliophora). *Int. J. Syst. Evol. Microbiol.* 66, 2959–2964. doi: 10.1099/ijsem.0.001127
- Lynn, D. H., Kolisko, M., and Bourland, W. (2018). Phylogenomic analysis of *Nassula variabilis* n. sp., *Furgasonia blochmanni*, and *Pseudomicrothorax dubius* confirms a nassophorean clade. *Protist* 169, 180–189. doi: 10.1016/j.protis.2018.02.002
- Lynn, D. H., and Small, E. (1997). A revised classification of the phylum Ciliophora Doflein, 1901. *Rev. Soc. Mex. Hist. Nat.* 47, 65–78.
- Lynn, D. H., and Small, E. B. (2002). “An Illustrated Guide to the Protozoa,” in *Society of Protozoologists, Phylum Ciliophora Doflein, 1901*, 2nd Edn, eds J. J. Lee., G. F. Leedale, and P. Bradbury, (Lawrence: Allen Press), 371–656.
- Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21, 3448–3449. doi: 10.1093/bioinformatics/bti551
- McGrath, C. L., Zufall, R. A., and Katz, L. A. (2007). Variation in macronuclear genome content of three ciliates with extensive chromosomal fragmentation: a preliminary analysis. *J. Eukaryot. Microbiol.* 54, 242–246. doi: 10.1111/j.1550-7408.2007.00257.x
- Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). “Creating the CIPRES Science Gateway for inference of large phylogenetic trees,” in *2010 Gateway Computing Environments Workshop (GCE)*, (Piscataway, NY: IEEE), 1–8.
- Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., et al. (2018). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* 47, D351–D360. doi: 10.1093/nar/gky1100
- Nurk, S., Bankevich, A., Antipov, D., Gurevich, A. A., Korobeynikov, A., Lapidus, A., et al. (2013). Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J. Comput. Biol.* 20, 714–737. doi: 10.1089/cmb.2013.0084
- Nylander, J. (2004). *MrModeltest v2. Program Distributed by the Author*. Uppsala: Uppsala University.
- Pan, H., Wang, L., Jiang, J., and Stoeck, T. (2016). Morphology of four cyrtophorian ciliates (Protozoa, Ciliophora) from Yangtze Delta, China, with notes on the phylogeny of the genus *Phascalodon*. *Eur. J. Protistol.* 56, 134–146. doi: 10.1016/j.ejop.2016.08.005
- Penn, O., Privman, E., Ashkenazy, H., Landan, G., Graur, D., and Pupko, T. (2010). GUIDANCE: a web server for assessing alignment confidence scores. *Nucleic Acids Res.* 38(Suppl_2), W23–W28. doi: 10.1093/nar/gkq443
- Perte, M., Kim, D., Perte, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095
- Piovesan, A., Caracausi, M., Ricci, M., Strippoli, P., Vitale, L., and Pelleri, M. C. (2015). Identification of minimal eukaryotic introns through GeneBase, a user-friendly tool for parsing the NCBI Gene databank. *DNA Res.* 22, 495–503. doi: 10.1093/dnares/dsv028
- Ricard, G., de Graaf, R. M., Dutilh, B. E., Duarte, I., van Alen, T. A., van Hoek, A. H., et al. (2008). Macronuclear genome structure of the ciliate *Nyctotherus ovalis*: single-gene chromosomes and tiny introns. *BMC Genomics* 9:587. doi: 10.1186/1471-2164-9-587
- Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029
- Roy, M., Kim, N., Xing, Y., and Lee, C. (2008). The effect of intron length on exon creation ratios during the evolution of mammalian genomes. *RNA* 14, 2261–2273. doi: 10.1261/rna.1024908
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Slabodnick, M. M., Ruby, J. G., Reiff, S. B., Swart, E. C., Gosai, S., Prabakaran, S., et al. (2017). The macronuclear genome of *Stentor coeruleus* reveals tiny introns in a giant cell. *Curr. Biol.* 27, 569–575. doi: 10.1016/j.cub.2016.12.057
- Small, E.B., and Lynn, D.H. (1985). “Phylum Ciliophora, Doflein, 1901” In *An Illustrated Guide to the Protozoa*. Eds J.J. Lee., S. H. Hutner., & E. C. Bovee, (Kansas: Society of Protozoologists) 393–575.
- Smit, A. F. (1996). The origin of interspersed repeats in the human genome. *Curr. Opin. Genet. Dev.* 6, 743–748. doi: 10.1016/s0959-437x(96)80030-x
- Smith, J. M., and Smith, N. (1996). Site-specific codon bias in bacteria. *Genetics* 142, 1037–1043.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Swart, E. C., Bracht, J. R., Magrini, V., Minx, P., Chen, X., Zhou, Y., et al. (2013). The *Oxytricha trifallax* macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. *PLoS Biol.* 11:e1001473. doi: 10.1371/journal.pbio.1001473
- Swart, E. C., Serra, V., Petroni, G., and Nowacki, M. (2016). Genetic codes with no dedicated stop codon: context-dependent translation termination. *Cell* 166, 691–702. doi: 10.1016/j.cell.2016.06.020
- Tucker, J. B. (1968). Fine structure and function of the cytopharyngeal basket in the ciliate *Nassula*. *J. Cell Sci.* 3, 493–514. doi: 10.1016/j.jwt.2014.09.005
- UniProt, C. (2018). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515. doi: 10.1093/nar/gky1049
- Wang, Y., Wang, C., Jiang, Y., Katz, L. A., Gao, F., and Yan, Y. (2019). Further analyses of variation of ribosome DNA copy number and polymorphism in ciliates provide insights relevant to studies of both molecular ecology and phylogeny. *Sci. China Life Sci.* 62, 203–214. doi: 10.1007/s11427-018-9422-9425
- Weisenberg, R. C., Broisy, G. G., and Taylor, E. W. (1968). Colchicine-binding protein of mammalian brain and its relation to microtubules. *Biochemistry* 7, 4466–4479. doi: 10.1021/bi00852a043
- Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data Analysis*. Berlin: Springer.
- Woese, C., Achenbach, L., Rouviere, P., and Mandelco, L. (1991). Archaeal phylogeny: reexamination of the phylogenetic position of *Archaeoglobus fulgidus* in light of certain composition-induced artifacts. *Syst. Appl. Microbiol.* 14, 364–371. doi: 10.1016/s0723-2020(11)80311-80315
- Wu, M., and Eisen, J. A. (2008). A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* 9:R151. doi: 10.1186/gb-2008-9-10-r151
- Zhang, Q., Yi, Z., Fan, X., Warren, A., Gong, J., and Song, W. (2014). Further insights into the phylogeny of two ciliate classes Nassophorea and Prostomatea (Protista, Ciliophora). *Mol. Phylogen. Evol.* 70, 162–170. doi: 10.1016/j.ympev.2013.09.015
- Zhang, T., Wang, C., Katz, L. A., and Gao, F. (2018). A paradox: rapid evolution rates of germline-limited sequences are associated with conserved patterns of rearrangements in cryptic species of *Chilodonella uncinata* (Protista, Ciliophora). *Sci. China Life Sci.* 61, 1071–1078. doi: 10.1007/s11427-018-9333-9331
- Zong, C., Lu, S., Chapman, A. R., and Xie, X. S. (2012). Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* 338, 1622–1626. doi: 10.1126/science.1229164

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Pan, Chen, Hou, Zhang, Qu, Warren and Miao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.