RESEARCH ARTICLE

Open Access

Genomic insights into the cellular specialization of predation in raptorial protists

Zaihan Li[1], Xiao Chen[2], Fangqing Zhao[1,3] and Miao Miao[1*]

Abstract

Background Predation is a fundamental mechanism for organisms to acquire energy, and various species have evolved diverse tools to enhance their hunting abilities. Among protozoan predators, raptorial Haptorian ciliates are particularly fascinating as they possess offensive extrusomes known as toxicysts, which are rapidly discharged upon prey contact. However, our understanding of the genetic processes and specific toxins involved in toxicyst formation and discharge is still limited.

Results In this study, we investigated the predation strategies and subcellular structures of seven Haptoria ciliate species and obtained their genome sequences using single-cell sequencing technology. Comparative genomic analysis revealed distinct gene duplications related to membrane transport proteins and hydrolytic enzymes in Haptoria, which play a crucial role in the production and discharge of toxicysts. Transcriptomic analysis further confirmed the abundant expression of genes related to membrane transporters and cellular toxins in Haptoria compared to Trichostomatia. Notably, polyketide synthases (PKS) and L-amino acid oxidases (LAAO) were identified as potentially toxin genes that underwent extensive duplication events in Haptoria.

Conclusions Our results shed light on the evolutionary and genomic adaptations of Haptorian ciliates for their predation strategies in evolution and provide insights into their toxic mechanisms.

Keywords Ciliates, Haptoria, Predation, Single-cell genomics, Toxicyst

Background

Predation plays a crucial role as a fundamental mechanism by which organisms acquire the necessary energy for survival and growth. Through the active pursuit and capture of prey, animals engage in dynamic predator-prey relationships that shape ecosystems and drive the intricate web of life. In their pursuit of sustenance, organisms have evolved a remarkable assortment of tools to enhance their hunting capabilities. For instance, *Rana pipiens* employs a viscoelastic tongue to efficiently capture crickets [1], while *Naja pallida* immobilizes mice with its venomous fangs [2, 3]. These diverse tools exemplify the distinct adaptations possessed by different species to facilitate their hunting activities. Each species possesses specialized instruments uniquely suited to its ecological niche and prey preferences. Furthermore, as our knowledge of the microscopic world deepens, it has become apparent that eukaryotic microbes, often regarded as the foundation of the food chain, have also developed their own predatory weapons [4, 5]. Protists, as the deepest branches on the eukaryotic tree [6], also possess a vast genetic diversity within the eukaryotic organisms [7, 8]. These single-cell organisms serve as the ancestral forms for many functions that are essential in higher organisms. For example, micropinocytosis,

*Correspondence:
Miao Miao
miaomiao@ucas.ac.cn
[1] University of Chinese Academy of Sciences, Beijing 100049, China
[2] Laboratory of Marine Protozoan Biodiversity and Evolution, Marine College, Shandong University, Weihai 264209, China
[3] Institute of Zoology, Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China

Li *et al. BMC Biology*   (2024) 22:107

Page 2 of 11

essential for mammals, can be traced to at least the shared ancestor of Amoebozoa and Opisthokonts [9, 10]. Among the diverse groups of protists, ciliates are widely distributed in various aquatic environments, standing out as particularly intriguing [11, 12]. Ciliates exhibit remarkable functional characteristics that contribute to their enhanced survival. For example, *Spirostomum* shows an ability for responsive contraction movements through its mesh-like contractile fibrillar system [13, 14]. In contrast, plagiopylean ciliates acquire energy in anaerobic environments by forming endosymbiotic relationships with anaerobic prokaryotes like archaeal methanogens and denitrifying bacteria [15–17].

Extrusomes are cellular organelles that play a significant role in predator-prey interactions among ciliates. For ciliates that with the most in-depth study, such as *Tetrahymena*, *Paramecium*, and *Frontonia*, these specialized cellular structures serve as defense mechanisms against predator attacks [18, 19]. When a potential threat makes contact with *Paramecium*, it induces a rapid and explosive extrusion of trichocysts, enabling the ciliate to swiftly escape. Mucocysts also appear in prey ciliates such as *Tetrahymena*, functioning to respond by releasing compounds that neutralize the attack [20]. However, there is an exception to this pattern observed in Haptoria. As natural predators, Haptoria have evolved offensive extrusomes, known as toxicysts [21–24]. These toxicysts, located in their oral apparatus, are designed for predation and are rapidly discharged upon contact with prey, immobilizing them. Wessenberg and Antipa [25] discovered that when the proboscis of *Didinium* contacts with *Paramecium*, elongated extrusive organelles are instantaneously released from the oral apparatus, facilitating the capture and killing of the prey [24, 26]. The toxicysts in their resting state exhibit a highly intricate and delicate structure, consisting of an external capsule covered by a thick wall and an internal tube. It is believed that unknown toxins accumulate within the tube and are released into the prey's body through the telescopic discharge of the tubule and/or the fusion of the toxicyst membrane with the plasma membrane [27]. Toxicysts are not exclusive to haptorian ciliates. For example, *Coleps* from the class Prostomatea also possesses these toxic structures [28]. Trichostomatia, commonly found in the rumen of herbivorous mammals, has been confirmed through electron microscopy to lack the characteristic extrusome structure of Haptoria [29].

Currently, ongoing studies are dedicated to exploring the structure and discharge mechanism of toxicysts in Haptorian ciliates. This aspect has been more extensively studied in *Coleps*, where the discharge of toxicysts was found to be linked with the significant production of free fatty acids (FFAs), as revealed by mass spectrometry techniques [28]. However, our understanding of the genetic processes responsible for toxin production and the specific types of toxins contained within toxicysts remains limited. To address this gap, we utilized single-cell sequencing technology to obtain genome sequences from seven species of Haptorian ciliates collected from both marine and freshwater environments. Through comparative genomic analysis, we successfully identified mechanisms influencing the discharge of toxicysts. This research sheds light on how these haptorian protists employ tools in their predation activities.

Results

Morphology and lifestyle

To investigate the specific mechanisms of raptorial ciliates, we collected a total of seven Haptorian ciliates, including two species from seawater and five from freshwater. Despite their large size and comparatively limited maneuverability compared to their prey, these Haptorian ciliates exhibit remarkable predation abilities across diverse aquatic environments. For example, the oval-shaped *Didinium* mainly engulfs *Paramecium*, whereas *Dileptus*, possessing a slender form, consumes *Tetrahymena* (Fig. 1A, B). Despite variations in predation strategies and morphological characteristics, Haptoria species are adept at preying on medium-sized protozoa, utilizing their toxicysts as their predatory instruments.

It is believed that subcellular structures play an important role in the distinct predatory strategies of ciliates through the development of specialized organelles, such as proboscis and extrusomes [22, 24, 30]. To examine the intricate architecture of ciliates, we used a transmission electron microscope (TEM) to investigate the subcellular structures in *Litonotus*. TEM enabled us to highlight the organelles involved in the raptorial lifestyle. Notably, Fig. 1D, E presents TEM images depicting the different types of toxicysts identified in *Litonotus*, which have been classified as long toxicysts and short pexicysts based on their distinctive morphological characteristics, as observed in *Didinium nasutum* [25]. We further provided a detailed examination of the toxicyst region (Fig. 1C), which typically serves as the initial point of contact between the cell and its prey during the raptorial feeding process. Several Haptoria species possess two types of toxicysts with varying sizes and structures. Specifically, type I toxicysts commonly appear as elongated rod-like shapes and are consistently present across all Haptoria species, whereas type II toxicysts (pexicysts) are slightly shorter and may be absent in certain species [27].

Genome assembly of Haptorian ciliates

To unravel the functional role of toxicysts, we performed a comprehensive analysis of genomic data to
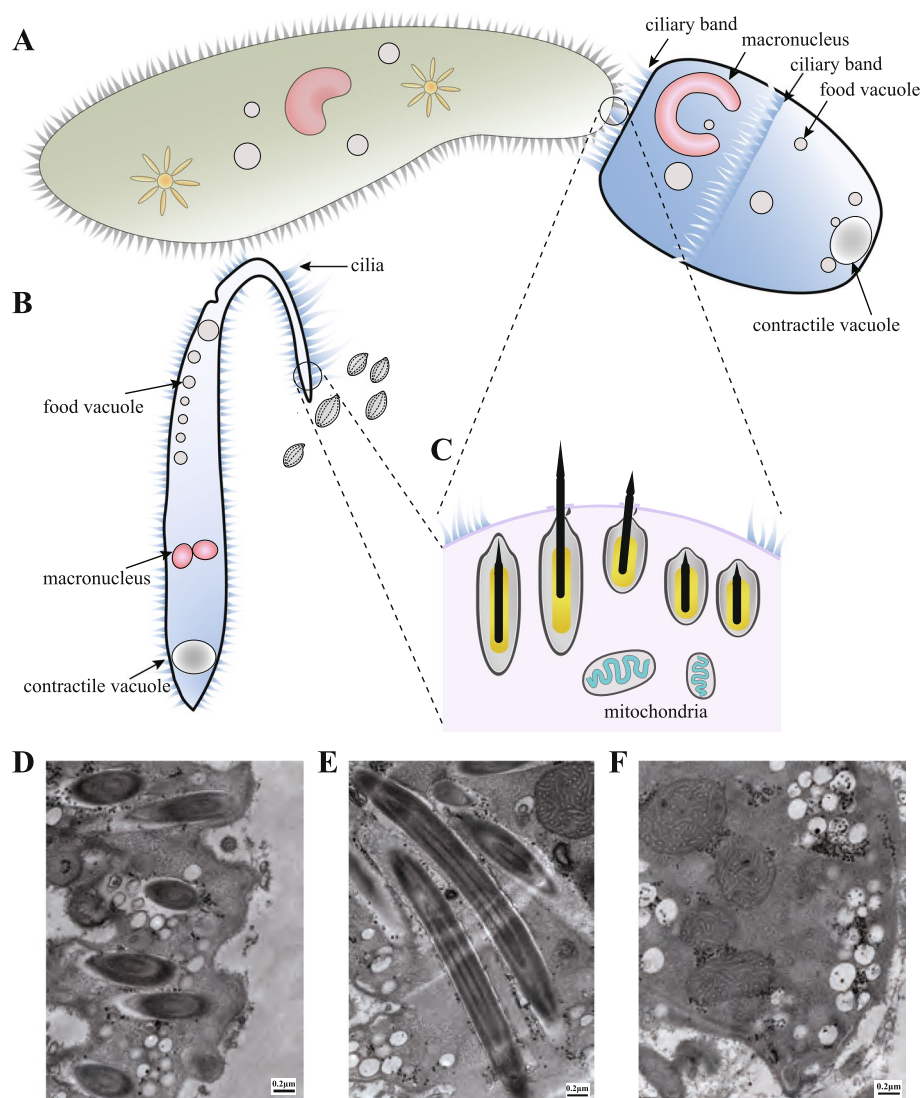
Fig. 1 Morphology of Haptoria and their subcellular structures. **A** *Didinium nasutum* (blue on the right) preying on *Paramecium caudatum* (green on the left), with *Didinium* utilizing toxicysts near its proboscis to immobilize its prey. **B** *Dileptus* sp. (blue on the left) preying on *Tetrahymena* (gray on the right). **C** Internal view of toxicysts within the surface of a Haptoria ciliate, displaying a toxicyst in the process of discharging. **D**-**F** Subcellular structures in a raptorial ciliate, *Litonotus* sp., using transmission electron microscope. **D** Short type II toxicysts. **E** Long type I toxicysts. **F** Mitochondria

identify the genes associated with their formation. After eliminating contamination originating from bacteria and algae, 7 ciliate genomes show different N50 lengths (1653 ~ 16,689 bp) and the largest contig length (38,189 ~ 287,940 bp). The average size of the seven genome assemblies was found to be 72 Mb, with GC contents ranging from 22.59 to 35.86%. Through a combination of ab initio, homology-based, and transcriptome-guided gene prediction, we finally obtained 18,691 ~ 32,428 protein-coding genes for each species (Table 1), comparable to other Litostomatea species, particularly those belonging to the Haptoria group. The

completeness of these assembled genomes was evaluated using EukCC [31], with the majority of them achieving over 85% completeness, except for *Trachelophyllidae* spp., which exhibited a completeness level of 66.7%.

Comparative genomics and phylogenomic analysis

To investigate the phylogenetic relationship of Haptoria within the Ciliophora, we conducted an analysis encompassing a total of 57 species, including 53 ciliates and 4 apicomplexans. Among the 53 ciliates, 21 species belonged to the class Litostomatea. All other genomic or transcriptomic data utilized in this study were

Table 1 Single-cell genome assemblies of Haptoria

Species	Clean reads (Gb)	Genome Size (bp)	Largest contig (bp)	N50 (bp)	GC (%)	Genome completeness (%)	No. of genes
Lacrymaria sp.1	20.5	99,803,170	137,119	2832	33.6	97.0	32,428
Lacrymaria sp.2	19.9	49,422,897	53,718	3007	29.0	90.9	22,896
Didinium sp.1	72.0	72,122,335	137,092	10,023	24.9	93.9	21,837
Trachelophyllidae spp.	40.5	47,450,859	101,362	1146	36.2	66.7	20,269
Monodinium sp.	30.6	45,171,837	38,189	1889	31.7	87.9	18,691
Litonotus sp.1	19.4	39,834,529	71,383	1993	26.2	93.9	17,390
Dileptus sp.	58.7	96,500,208	53,007	2421	25.7	90.9	26,891

obtained from the SRA or Genome database in NCBI. As shown in Fig. 2A, the class Litostomatea could be classified into 3 distinct clades. The order Cyclotrichida, consisting of a sole family *Mesodiniidae*/*Myrionecta*, appears isolated from the rest of the Litostomatea. This observation aligns well with the fact that Mesodiniidae exhibit a distinct lifestyle divergent from other organisms. Subsequently, we focused on the other two clades within the Litostomatea, which includes the 7 ciliates we sequenced. We compared the genome assemblies and gene predictions between Haptoria clade and Trichostomatia clade. We found that Haptoria exhibits a shorter average chromosome length compared to Trichostomatia (Fig. 2B). Furthermore, we observed notable differences in gene length between the 2 subgroups. Trichostomatia displayed a higher proportion of genes below 1000 bp, whereas Haptoria exhibited a higher proportion of genes exceeding 1300 bp (Fig. 2C). We also predicted genes in Mesodiniidae using the transcriptome data. The gene length distribution of Mesodiniidae is closer to the Trichostomatia clade, significantly differing from that of Haptoria (Additional file 1: Fig. S1).

To investigate the gene content within Litostomatea, we employed OrthoFinder [32] to compare the homologous genes between Haptoria and its closely related Trichostomatia, with 2 Spirotrichea species serving as outgroups. A total of 170,280 genes were clustered into 26,848 gene families across the 12 species, with 80.6% of these genes present in the analyzed orthogroups. The resulting clustered tree (Fig. 3C) generated from multiple sequence alignments shows a similar structure to the maximum likelihood (ML) tree in Fig. 2A. Notably, when examining the number of homologous genes in different species, a significant distinction between Haptoria and other species became apparent regardless of the gene count within each homologous group (Fig. 2D). From the perspective of orthogroups and gene numbers, Haptoria possesses a greater number of shared homologous genes than other members of the Litostomatea group.

In addition, we identified a specific set of 2818 orthogroups classified as core orthogroups, which were found in at least 6 species (Fig. 3A). Haptoria species exhibited a significantly higher number of core orthogroups, with a mean value of 2550, nearly double that of Trichostomatia (1490). To further assess the relationship among these predatory ciliates, we employed principal component analysis (PCA) to assess the differences between Haptoria, Trichostomatia, and Spirotrichea. As a result, more evidence suggests that the ortholog profiles in Haptoria can also be significantly distinguished from other families (Fig. 3B).

Toxic-related and transmembrane gene expansions in Haptoria

To explore the factors contributing to the genomic divergence within Litostomatea, we focused on the gene expansion events in Haptoria. To determine the divergence time of Litostomatea, we used both the BEAST and R8S [33] methods, employing Spirotrichea as a temporal reference point due to their inclusion in the fossil record [34]. The ultrametric tree generated by R8S is more similar to the phylogenetic tree, and thus, we chose the analysis results of R8S. We identified 78 expanded gene families and 120 contracted gene families in the Haptoria clade. Further examination of the expanded gene families using the PFAM database [35] allowed us to identify eight gene families: Thiolase, Heme_oxygenase, Sec10 Exocyst, LCCL-domain, ComplexI-N, Fibrinogen_C, TSP1, and COX1-like (Additional file 1: Table S1). These gene families, associated with membrane transport proteins and hydrolytic enzymes, exhibited higher abundance within the Haptoria clade, while being largely absent in both Trichostomatia clade and Spirotrichea clade (Fig. 3C).

To investigate the toxicyst's related membrane transport proteins and hydrolytic enzymes, we conducted a comprehensive analysis of the transcriptomic data
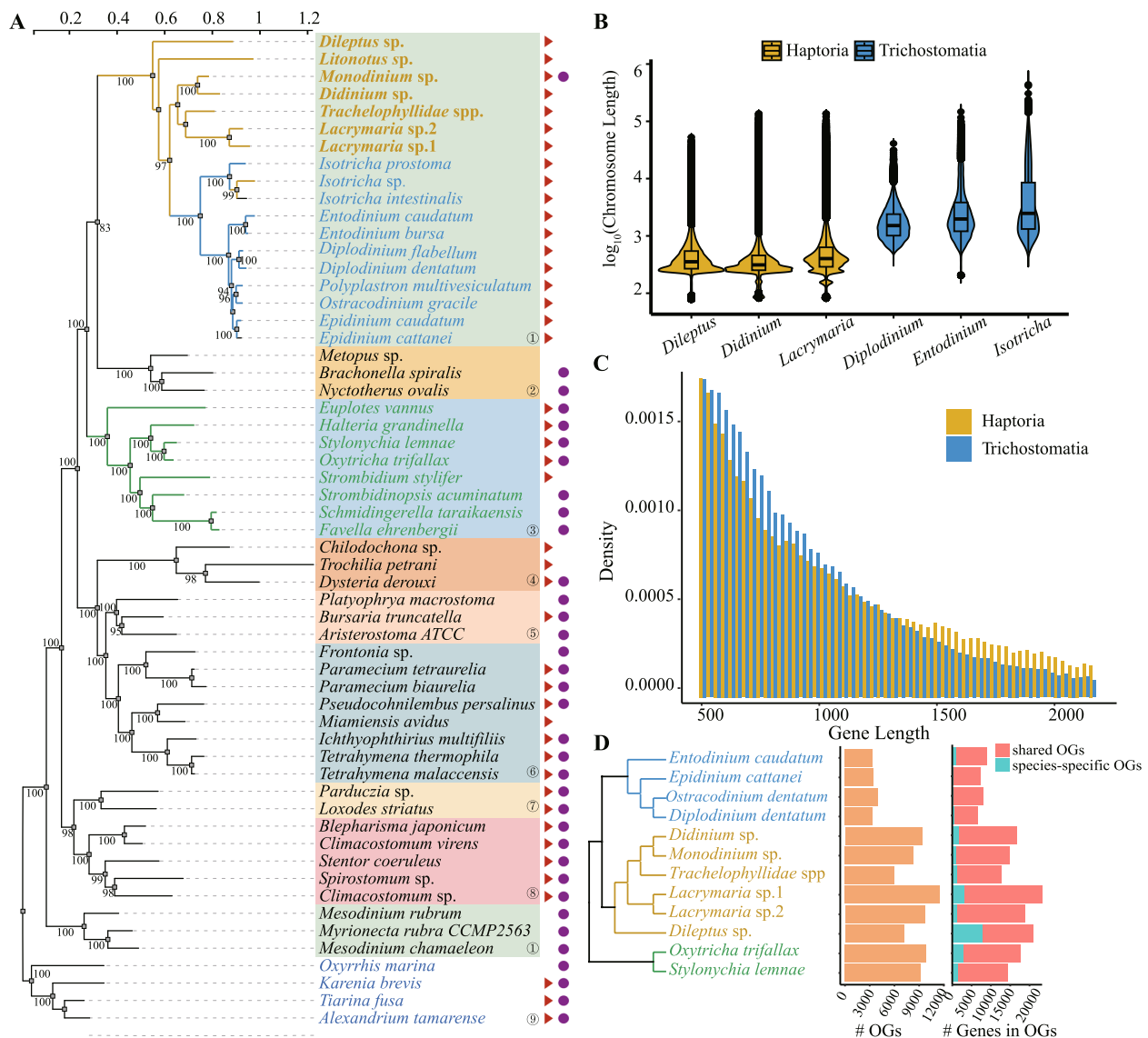
Fig. 2 Phylogenomic analyses of Litostomatea with other ciliates. **A** Phylogenomic tree of ciliates constructed using ML, depicting the evolutionary relationships among different ciliate classes. The numbered labels represent the following classes: ① Litostomatea (subclass Haptoria in yellow and subclass Trichostomatia in blue), ② Armophorea, ③ Spirotrichea, ④ Phyllopharyngea, ⑤ Colpodea, ⑥ Oligohymenophorea, ⑦ Karyorelictea, and ⑧: Heterotrichea. Species with red triangles indicate the availability of genome data, and purple dots indicate the availability of transcriptome data. **B, C** Statistics on the assembly and annotation of Haptoria and Trichostomatia genomes, showing the distribution of chromosomes and gene length. **D** Comparision of orthologs and gene counts in orthogroups among Haptoria, Trichostomatia, and Spirotrichea

from six Haptoria species and five Trichostomatia species, with *Tetrahymena thermophila*, *Paramecium bursaria*, and *Fabrea salina* as the outgroup species. We evaluated the expression of specific genes by comparing them to the TCDB [36], PHI-base [37], and VFDB [38] databases, and revealed a significantly higher abundance of membrane transport proteins and cellular toxins within Haptoria compared to both Trichostomatia and outgroup species (Fig. 3D, Additional file 1:

Table S2). We further compared the orthogroup genes between Haptoria and Trichostomatia and found that genes from Haptoria were enriched in transmembrane transporter-related gene families, while Trichostomatia exhibited enrichment in biosynthetic and non-membrane-bounded genes (Fig. 4A). Additionally, we performed gene enrichment analysis for Haptoria using WEGO2.0 [39], which revealed significant enrichment in pathways related to microtubules, transmembrane transporters, and hydrolase activity (Fig. 4B).

Li *et al. BMC Biology*   (2024) 22:107

Page 6 of 11
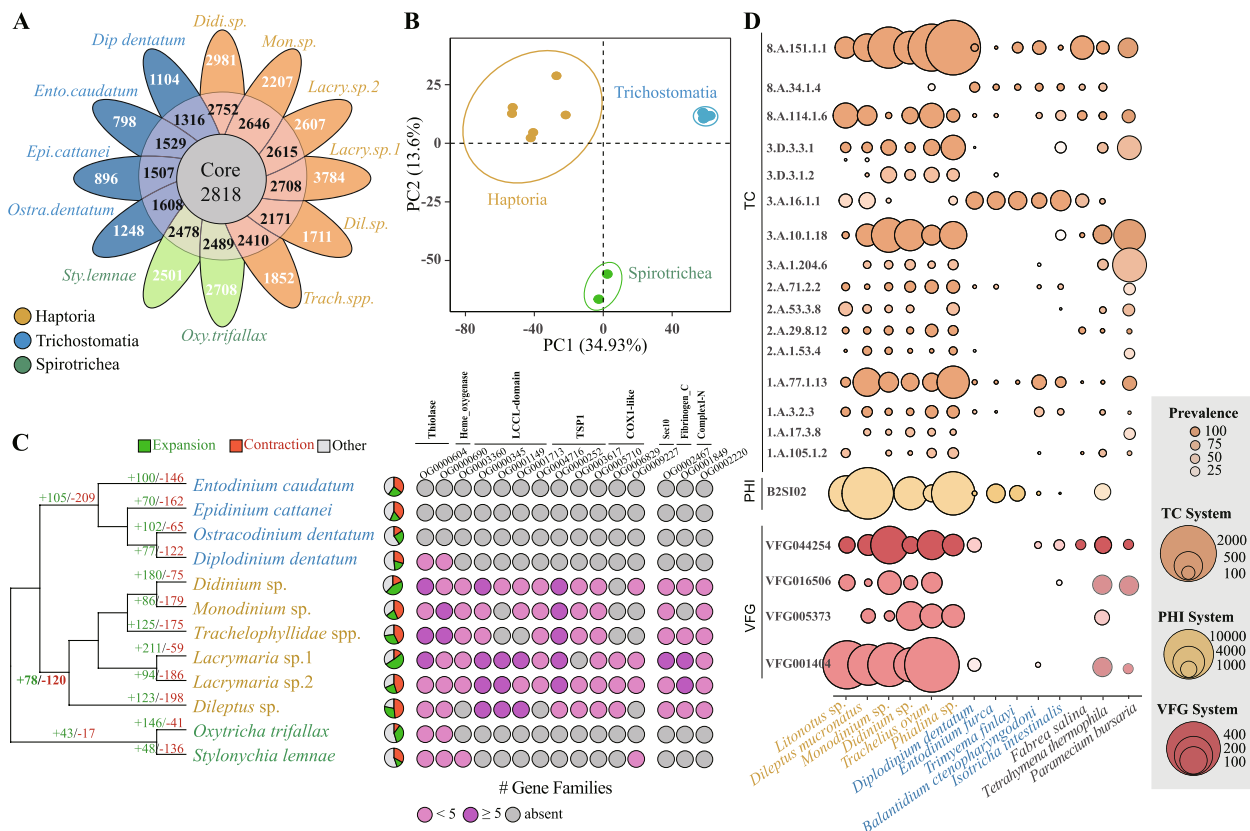
Fig. 3 Comparative genomic analysis between Haptoria and Trichostomatia. **A** The Venn diagram illustrates the number of shared orthogroups among ciliates. The value inside the circles represent the number of core orthogroups shared by at least seven species, while the values outside the circles represent the number of species-specific orthogroups. **B** Principal component analysis of ciliate orthologs shows a clear segregation between Haptoria, Trichostomatia, and Spirotrichea. **C** Gene expansion and contraction analysis in Haptoria and related species (left). The expanded gene families are shown on the right panel. **D** Gene expression and prevalence analysis of TCDB, PHI-base, and VFDB genes in Haptoria, Trichostomatia, *Fabrea salina*, *Tetrahymena thermophila*, and *Paramecium bursaria*

To identify potential toxin genes in Haptoria, we leveraged toxin proteins available in the Uniprot database. Among the discovered genes, polyketide synthases (PKS) and ʟ-amino acid oxidases (LAAO) stood out due to their extensive duplication events. PKSs play a crucial role in the biosynthesis of diverse secondary metabolites, including toxins, and exhibit eight conserved structural domains. [40] Interestingly, our analysis revealed a noteworthy disparity in the abundance of these domains between Haptoria and Trichostomatia genomes, with six domains being significantly more prevalent in Haptoria (Fig. 4C). Additionally, we explored LAAOs, enzymes responsible for the oxidative deamination of amino acids, and found that a unique structural domain exhibited a higher number of repeats in Haptoria genomes (Fig. 4D). This finding suggests a potential functional significance of LAAO in Haptoria's predatory activities.

Discussion

While the presence of toxicysts in Haptoria has long been acknowledged, the genetic makeup of this unique class has not been thoroughly investigated. One of the major difficulties in studying this group has been the inherent challenge of culturing these predators in large numbers within laboratory settings. Previously, genome sequencing efforts were primarily focused on parasitic ciliates within the Litostomatea class. However, by employing single-cell sequencing methods, we have been able to delve into the intricate genetic details using a limited number of ciliate cells. To address uncertainties regarding the termination codons of Litostomatea and the lack of protein evidence, we optimized the methods for genome assembly and gene prediction and effectively eliminated bacterial contamination. Consequently, we have successfully sequenced and assembled the single-cell macronuclear genomes of seven Haptorian ciliates.

Li *et al. BMC Biology*   (2024) 22:107

Page 7 of 11
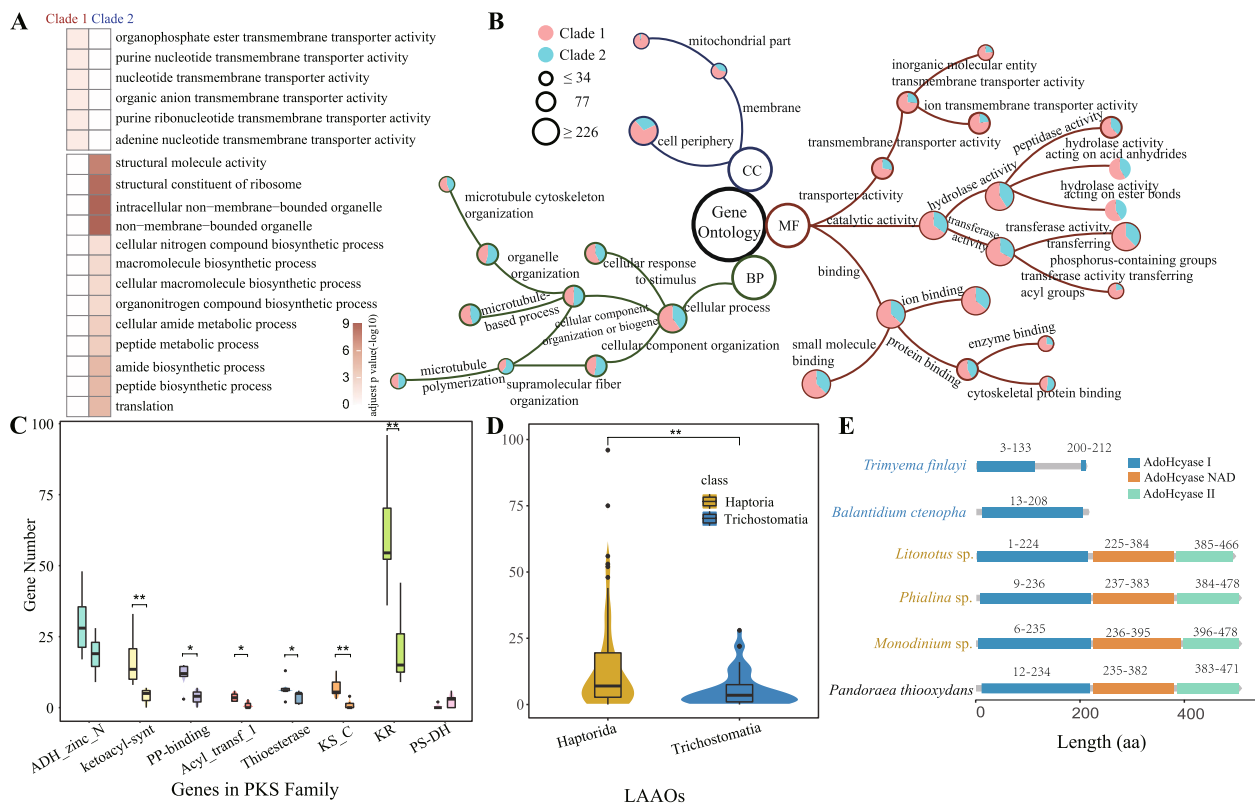
Fig. 4 Functional enrichment analysis reveals *toxin-related gene expansions in Haptoria*. **A** Gene Ontology (GO) enrichment analysis of genes with increased copy numbers in clade 1 Haptoria and clade 2 Trichostomatia. **B** Gene Ontology (GO) enrichment analysis based on genes with increased copy numbers in clade 1 Haptoria and clade 2 Trichostomatia. The enriched categories include biological process (BP), cellular component (CC), and molecular function (MF). The red and cyan circles represent the proportions of gene copies from Haptoria and Trichostomatia, respectively. The size of the circles corresponds to the total number of gene copies. **C** Comparison of the eight conserved structural domains of the PKS (polyketide synthase) family between Haptoria and Trichostomatia from the Litostomatea class. The left panel shows the genes from Haptoria, while the right panel shows the genes from Trichostomatia. **p-value < 0.01 and *p-value < 0.05 by Student's t test. **D** Comparison of the number of LAAO (L-amino acid oxidase) genes from the Litostomatea class between Haptoria and Trichostomatia. **p-value < 0.01 by Student's t test. **E** Comparison of the AdoHcyase family in different ciliates with the bacterium *Pandoraea thiooxydans*

In spite of their close evolutionary relationship, significant divergence was observed between Haptoria and other Litostomatea species. Firstly, our comparative genomics analysis provides support for the clear separation of Mesodiniidae species from other core species within the Litostomatea class. This finding indicates that the symbiotic family Mesodiniidae exhibits a significantly distinct genetic profile when compared to their predatory counterparts. Secondly, through the comparison of chromosome length and gene length, we identified notable distinctions between Haptoria and Trichostomatia. This observation is further reinforced by the analysis of orthologous genes between Haptoria and Trichostomatia. Compared to Haptoria, Trichostomatia exhibits a lower abundance of orthologous genes and genes within orthologous groups. Furthermore, the non-toxicyst clades (Trichostomatia and Spirotrichea) are characterized by a deficiency in

gene families associated with enzyme activity and ionic transmembrane functions, whereas Haptoria shows a significant increase in the number of gene duplications associated with these functions. The majority of these gene families are associated with membrane transport proteins and hydrolytic enzymes, potentially playing a role in the predation mechanism of Haptorian ciliates. During the discharge of toxicysts, the fusion between the toxicyst membrane and the cell membrane has been observed. Previous studies have shown that this process is correlated with fluctuations in Ca^{2+} concentration, leading to the rapid expulsion of toxicysts from the cell [19, 23]. Transcriptome data further strengthens these findings by showing that ciliates from the Haptorian ciliates express a higher number of genes associated with membrane transport proteins and cellular toxins compared to other ciliate groups. AdoHcyase (S-adenosylhomocysteine hydrolase) family was screened from

Li *et al. BMC Biology*   (2024) 22:107

Page 8 of 11

Haptoria transcriptomes as an essential hydrolase [41]. In Haptorian ciliates, members of this hydrolytic enzyme family exhibit generally higher gene expression levels and comparatively more intact gene structural domains. Conversely, within Trichostomatia, the same gene structure of this category appears to be absent (Fig. 4E). Furthermore, the presence of acid phosphatase, typically located in the lysosomes of animal cells, within Haptoria [26] suggests that these ciliates may employ a sophisticated set of hydrolytic enzymes to initiate prey digestion. It is worth noting that several extrusomes in protists, such as mucocysts in *Tetrahymena*, trichocysts in *Paramecium* (considered homologous organelles with mucocysts [42]), and rhoptries in Apicomplexans, are regarded as lysosome-related organelles [19, 43].

In terms of phenotype, the Litostomatea class can be divided into three distinct groups based on their feeding patterns: Mesodiniidae, known for their mixotrophic lifestyle with algae [44, 45]; Trichostomatia, which parasitize the rumen of mammals [46]; and Haptoria, characterized as free-living raptorial ciliates [47]. The diverse lifestyles within Litostomatea have resulted in significant evolutionary divergence. As parasitic species, Trichostomatia have undergone gene loss of many essential genes typically found in free-living ciliates. Instead, they have evolved CAZymes, similar to those found in gut microbes, to enhance the digestion of polysaccharides [46]. In contrast, the expansive aquatic environments in which Haptoria reside have imposed distinct evolutionary pressures, leading to the development of toxicysts as a crucial part of their predatory strategy. Compared to Trichostomatia, Haptoria exhibits a higher abundance and diversity of genes associated with transmembrane transport and cellular toxins within their genomes. This indicates that Haptoria have undergone specific adaptations, resulting in genes specialized for the discharge of toxicysts and the paralysis of prey.

In addition, our investigation revealed an increased number of gene copies in Haptoria, specifically within certain gene families. Notably, the PKS gene family, known for its gene duplications in toxin production in ciguatoxic dinoflagellates [48], exhibits a higher frequency of gene duplications in the genomes of Haptoria compared to other ciliates. Similarly, the LAAO gene, which is widely distributed in the venom glands of animals such as rabbitfish and pit vipers [49], is found in a higher number of copies in Haptoria. Previous studies have demonstrated that LAAO derived from the serum of *Siganus oramin* induces lysis and mortality in the parasitic ciliate *Ichthyophthirius multifiliis* [50]. Further studies are still needed to investigate the function of these genes on mediating the production and discharge of toxicysts in Haptorian ciliates.

Conclusions

Considering that most of raptorial protists cannot be cultured in laboratories, by leveraging single-cell genome sequencing methods, we have been able to delve into the intricate genetic details of these predators using a limited number of ciliate cells. We have successfully sequenced and assembled the genomes of seven Haptoria ciliates. Comparative genomic analysis revealed distinct gene duplications related to membrane transport proteins and hydrolytic enzymes in Haptoria, which play a crucial role in the production and discharge of toxicysts. Transcriptomic analysis further confirmed the abundant expression of genes related to membrane transporters and cellular toxins in Haptoria compared to Trichostomatia. Notably, PKS and LAAO were identified as potentially toxic genes that underwent extensive duplication events in Haptoria.

Methods
Sample collection
Two species of *Lacrymaria* were isolated from seawater in Pier-bridge, Qingdao, China (36° 06′N, 120° 31′E). *Didinium* sp.1 was collected from the freshwater pond in Baihuayuan Park, Qingdao, China (36° 04′N, 120° 22′E). Trachelophyllidae spp. and *Monodinium* sp. were collected from the freshwater of Weishan Lake, Shandong, China (34° 67′N, 117° 24′E). *Dileptus* sp. and *Litonotus* sp. were collected from the freshwater of Jiudu River, Beijing, China (40° 37′N, 116° 38′E).

Transmission electron microscope (TEM)
Litonotus was used for investigating subcellular structures in raptorial ciliates. For convenient observation, each *Litonotus* undergoes two to three washes with freshwater using glass pipettes. The samples were fixed using 2.5% glutaraldehyde (pH 7.4) for 24 h at 4 °C. Subsequently, the specimens were embedded in Eponate 812, followed by sectioning into ultrathin slices. To facilitate observation and photography using Tecnai 20 (FEI, Lausanne, Netherlands) electron microscopy, the sections were stained with uranyl acetate and lead citrate. The shape of extrusomes and spatial distribution of mitochondria is plural based on TEM.

High-throughput sequencing and genome assembly
After conducting morphological observations and identifications, we individually aspirated ciliates that had undergone 12 h of starvation using glass pipettes, placed them in freshwater, and repeated this process three to four times to minimize bacterial contamination as much

Li *et al. BMC Biology*   (2024) 22:107

Page 9 of 11

as possible, and macronuclear genomic DNA was amplified using REPLI-g Single Cell Kit (Qiagen). The Truseq Nano DNA HT Sample Preparation Kit (Illumina USA) was used for the preparation of sequencing libraries. High-throughput sequencing was performed on an Illumina HiSeq platform using 150 bp paired-end reads.

To assess the quality of the paired-end sequencing data, FastQC v0.11.9 [51] was employed. Subsequently, low-quality reads were filtered out using Trimmomatic v0.39 (HEADCROP:10 ILLUMINACLIP:2:30:10 SLIDING-WINDOW:4:19) [52]. The genome of seven species were assembled using SPAdes v3.13.0 with kmer options (-k 61, 71, 77, 81) [40]. To eliminate contamination resulting from bacteria and algae, assembled sequences were aligned to the NR databases using DIAMOND v0.9.26 (--db nr --outfmt 6 --evalue 1e-5 --max-hsps 1) [53] and MetaGeneMark v3.38 [54], and those assigned to bacteria or algae were identified and removed as contaminants.

Other 37 genomes datasets were collected from other ciliates classes retrieved from NCBI. All these genomic data were assembled using Spades v3.13.0(-k 21,33,55,77). Gene prediction was performed using AUGUSTUS v3.3.0. All the data sources are documented in Additional file 1: Table S3.

Gene prediction and annotation
To enhance the accuracy of gene prediction in the assembled genomes, the following procedures were implemented: (1) AUGUSTUS v3.3.3 [55] was used to predict genes in the seven genome datasets obtained during this study, as well as all transcript data available in the NCBI database (Additional file 1: Table S3). The model was trained using the RNA-seq data of *Monodinium* sp. (2) Blast v2.10.0 was used to search for gene fragments that corresponded to transcripts in other members of the Litostomatea class. By combining de novo approaches with transcripts from related species, it was possible to identify species-specific genes within the Litostomatea class. (3) DIAMOND v0.9.26.127 was used to annotate protein-coding genes in the seven ciliate genomes by comparing against the NR database. EukCC v0.2 was used to assess the completeness of the genomes.

Phylogenetic and gene family analysis
We collected the genomic or transcriptomic data of additional 53 ciliates and 5 apicomplexans to serve as outgroup taxa. GPSit (-b BMGE-1.12 -e 1e-5) [56] was used to align them with 157 ciliate orthologous genes. ML analysis was performed using IQ-Tree v1.6.12 [57] with the model LG+F+R6, and the results were evaluated using 1000 bootstrap replicates.

OrthoFinder v2.5.4 was applied to identify orthologous groups from the all-against-all comparisons based on DIAMOND. *Entodinium caudatum*, *Epidinium cattanei*, *Ostracodinium dentatum*, and *Diplodinium dentatum* were categorized as Trichostomatia, while *Oxytricha trifallax* and *Stylonychia lemnae* were categorized as Spirotrichea. Species divergence time was estimated using both BEAST v 1.10.4 (mcmc 1000000) and R8S v1.81 (nsites = 419706 age = 368) with the divergence time of *Entodinium* and *Dileptus* as the calibration point. Based on the gene family clustering and divergence time estimation, CAFÉ v4.2.1 [58] was used to identify the expansion and contraction events of gene families. Subsequently, to investigate the functional implications of these gene family expansions or contractions, we mapped the genes within these families to the PFAM database using InterProScan. We performed an analysis to identify significantly overrepresented Gene Ontology (GO) terms among the rapidly expanded gene families using the clusterProfiler package [59] and WEGO 2.0. This approach allows us to gain insights into the functional enrichment of these gene families in Haptorian ciliates.

Transcriptome data analysis
All transcriptome data were sourced exclusively from the NCBI Short Read Archive (SRA). *Diplodinium dentatum*, *Entodinium furca*, *Trimyema finlayi*, *Balantidium ctenopharyngodoni*, and *Isotricha intestinalis* are representatives of the Trichostomatia. Transcripts from *Litonotus* sp., *Dileptus mucronatus*, *Monodinium* sp., *Didinium* sp.2, *Trachelius ovum*, and *Phialina* sp. were chosen as representatives of the Haptoria. We also collected other 29 transcriptomes from other species in ciliates (Additional file 1: Table S3). All the RNA-seq data were assembled using Trinity 2.8.5 [60]. *Fabrea salina* served as an outgroup. We employed the databases of TCDB, PHI-base, and VFDB to identify genes related to membrane transport and toxicity genes in these transcriptome data. The significance of the differences between expression abundance values was examined by Student's *t* test with the significance threshold set at $p < 0.01$.

Abbreviations
PKS	Polyketide synthases
LAAO	L-amino acid oxidases
FFAs	Free fatty acids
TEM	Transmission electron microscope
PCA	Principal component analysis
ML	Maximum likelihood method

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12915-024-01904-2.

Additional file 1: **Table S1.** Differential Gene Statistics of Orthogroups in Haptoria, Trichostomatia and Spirotrichea. Gene counts of orthogroups in

Li *et al. BMC Biology*   (2024) 22:107

Page 10 of 11

Acknowledgements

Authors' contributions

Funding

Availability of data and materials

All the data supporting the findings of this study are available and included in this published article, its supplementary information file (Additional file 1: Table S3), and publicly available repositories. Genomic data generated in this study have been deposited to the NGDC database (https://ngdc.cncb.ac.cn/gsa/) under BioProject number PRJCA019558.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. Noel AC, Guo HY, Mandica M, Hu DL. Frogs use a viscoelastic tongue and non-Newtonian saliva to catch prey. J R Soc Interface. 2017;14(127). https://doi.org/10.1098/rsif.2016.0764.
2. Cascardi J, Young BA, Husic HD, Sherma J. Protein variation in the venom spat by the red spitting cobra, Naja pallida (Reptilia: Serpentes). Toxicon. 1999;37(9):1271–9. https://doi.org/10.1016/S0041-0101(98)00264-5.
3. Williams HF, Mellows BA, Mitchell R, Sfyri P, Layfield HJ, Salamah M, et al. Mechanisms underpinning the permanent muscle damage induced by snake venom metalloprotease. PLoS Negl Trop Dis. 2019;13(1):e0007041. https://doi.org/10.1371/journal.pntd.0007041.
4. Falkowski P. Ocean science: the power of plankton. Nature. 2012;483(7387):S17–20. https://doi.org/10.1038/483S17a.
5. Bjorbækmo MFM, Evenstad A, Røsæg LL, Krabberød AK, Logares R. The planktonic protist interactome: where do we stand after a century of research? ISME J. 2020;14(2):544–59. https://doi.org/10.1038/s41396-019-0542-5.
6. Simpson AGB, Inagaki Y, Roger AJ. Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of ""primitive"" eukaryotes. Mol Biol Evol. 2006;23(3):615–25.
7. Caron DA, Countway PD, Jones AC, Kim DY, Schnetzer A. Marine protistan diversity. Ann Rev Mar Sci. 2012;4(1):467–93. https://doi.org/10.1146/annurev-marine-120709-142802.
8. Caron DA, Hu SK. Are we overestimating protistan diversity in nature? Trends Microbiol. 2019;27(3):197–205. https://doi.org/10.1016/j.tim.2018.10.009.
9. Bloomfield G, Kay RR. Uses and abuses of macropinocytosis. J Cell Sci. 2016;129(14):2697–705. https://doi.org/10.1242/jcs.176149.
10. King JS, Kay RR. The origins and evolution of macropinocytosis. Philos Trans R Soc Lond B Biol Sci. 2019;374(1765):20180158. https://doi.org/10.1098/rstb.2018.0158.
11. Hu X. Ciliates in extreme environments. J Eukaryot Microbiol. 2014;61(4):410–8. https://doi.org/10.1111/jeu.12120.
12. Liu W, Shin MK, Yi Z, Tan Y. Progress in studies on the diversity and distribution of planktonic ciliates (Protista, Ciliophora) in the South China Sea. Mar Life Sci Technol. 2020;3(1):28–43. https://doi.org/10.1007/s42995-020-00070-y.
13. Mathijssen AJTM, Culver J, Bhamla MS, Prakash M. Collective intercellular communication through ultra-fast hydrodynamic trigger waves. Nature. 2019;571(7766):560–4. https://doi.org/10.1038/s41586-019-1387-9.
14. Zhang J, Qin W, Hu C, Gu S, Chai X, Yang M, et al. Giant proteins in a giant cell: molecular basis of ultrafast Ca2+-dependent cell contraction. Sci Adv. 2023;9(8):eadd6550. https://doi.org/10.1126/sciadv.add6550.
15. Graf JS, Schorn S, Kitzinger K, Ahmerkamp S, Woehle C, Huettel BJ, et al. Anaerobic endosymbiont generates energy for ciliate host by denitrification. Nature. 2021;591(7850):445. https://doi.org/10.1038/s41586-021-03297-6.
16. van Hoek AH, van Alen TA, Sprakel VS, Leunissen JA, Brigge T, Vogels GD, et al. Multiple acquisition of methanogenic archaeal symbionts by anaerobic ciliates. Mol Biol Evol. 2000;17(2):251–8.
17. Rotterová J, Edgcomb VP, Čepička I, Beinart R. Anaerobic ciliates as a model group for studying symbioses in oxygen-depleted environments. J Eukaryot Microbiol. 2022;69(5):e12912. https://doi.org/10.1111/jeu.12912.
18. Buonanno F, Harumoto T, Ortenzi C. The defensive function of trichocysts in *Paramecium tetraurelia* against metazoan predators compared with the chemical defense of two species of toxin-containing ciliates. Zoolog Sci. 2013;30(4):255–61. https://doi.org/10.2108/zsj.30.255.
19. Plattner H. Trichocysts-Paramecium"s projectile-like secretory organelles: reappraisal of their biogenesis, composition, intracellular transport, and possible functions. J Eukaryot Microbiol. 2017;64(1):106–33. https://doi.org/10.1111/jeu.12332.
20. Kumar S, Briguglio JS, Turkewitz AP. Secretion of polypeptide crystals from tetrahymena thermophila secretory organelles (mucocysts) depends on processing by a cysteine cathepsin, Cth4p. Eukaryot Cell. 2015;14(8):817–33. https://doi.org/10.1128/EC.00058-15.
21. Verni F. Litonotus — Euplotes (predator — prey) interaction: ciliary structure modifications of the prey caused by toxicysts of the predator (Protozoa, Ciliata). Zoomorphology. 1985;105(5):333–5. https://doi.org/10.1007/BF00312065.
22. Yanase R, Nishigami Y, Ichikawa M, Yoshihisa T, Sonobe S. The neck deformation of *Lacrymaria olor* depending upon cell states. J Protistol. 2018;51:1–6. https://doi.org/10.18980/jop.e001.
23. Iwadate Y, Katoh K, Kikuyama M, Asai H. Ca2+ triggers toxicyst discharge inDidinium nasutum. Protoplasma. 1999;206(1):20–6. https://doi.org/10.1007/BF01279249.
24. Vďačný P, Rajter Ľ, Shazib SUA, Jang SW, Shin MK. Diversification dynamics of rhynchostomatian ciliates: the impact of seven intrinsic traits on speciation and extinction in a microbial group. Sci Rep. 2017;7(1):1–10. https://doi.org/10.1038/s41598-017-09472-y.

Li *et al. BMC Biology*   (2024) 22:107

Page 11 of 11

25. Wessenberg H, Antipa G. Capture and ingestion of paramecium by Didinium nasutum. J Protozool. 1970;17(2):250–70. https://doi.org/10.1111/j.1550-7408.1970.tb02366.x.

26. Hausmann K. Food acquisition, food ingestion and food digestion by protists. Jpn J Protozool. 2002;35(2):85–95. https://doi.org/10.18980/jjprotozool.35.2_85.

27. Rosati G, Modeo L. Extrusomes in ciliates: diversification, distribution, and phylogenetic implications. J Eukaryot Microbiol. 2003;50(6):383–402. https://doi.org/10.1111/j.1550-7408.2003.tb00260.x.

28. Buonanno F, Anesi A, Guella G, Kumar S, Bharti D, La Terza A, et al. Chemical offense by means of toxicysts in the freshwater ciliate, Coleps hirtus. J Eukaryot Microbiol. 2014;61(3):293–304. https://doi.org/10.1111/jeu.12106.

29. Wright ADG, Lynn DH. Monophyly of the trichostome ciliates (phylum Ciliophora: class Litostomatea) tested using new 18S rRNA sequences from the Vestibuliferids, Isotricha intestinalis and Dasytricha ruminantium, and the Haptorian Didinium nasutum. Eur J Protistol. 1997;33(3):305–15. https://doi.org/10.1016/s0932-4739(97)80008-9.

30. Coyle SM, Flaum EM, Li H, Krishnamurthy D, Prakash M. Coupled active systems encode an emergent hunting behavior in the unicellular predator *Lacrymaria olor*. Curr Biol. 2019;29(22):3838-3850.e3833. https://doi.org/10.1016/j.cub.2019.09.034.

31. Saary P, Mitchell AL, Finn RD. Estimating the quality of eukaryotic genomes recovered from metagenomic analysis with EukCC. Genome Biol. 2020;21:244. https://doi.org/10.1186/s13059-020-02155-4.

32. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20:238. https://doi.org/10.1186/s13059-019-1832-y.

33. Sanderson MJ. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics. 2003;19(2):301–2. https://doi.org/10.1093/bioinformatics/19.2.301.

34. Vďačný P. Estimation of divergence times in litostomatean ciliates (Ciliophora: Intramacronucleata), using Bayesian relaxed clock and 18S rRNA gene. Eur J Protistol. 2015;51(4):321–34. https://doi.org/10.1016/j.ejop.2015.06.008.

35. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: the protein families database in 2021. Nucleic Acids Res. 2021;49(D1):D412–9. https://doi.org/10.1093/nar/gkaa913.

36. Saier MH, Reddy VS, Moreno-Hagelsieb G, Hendargo KJ, Zhang Y, Iddamsetty V, et al. The Transporter Classification Database (TCDB): 2021 update. Nucleic Acids Res. 2021;49(D1):D461–7. https://doi.org/10.1093/nar/gkaa1004.

37. Urban M, Cuzick A, Seager J, Wood V, Rutherford K, Venkatesh Shilpa Y, et al. PHI-base in 2022: a multi-species phenotype database for pathogen–host interactions. Nucleic Acids Res. 2022;50(D1):D837–47. https://doi.org/10.1093/nar/gkab1037.

38. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 2019;47(D1):D687–92. https://doi.org/10.1093/nar/gky1080.

39. Ye J, Zhang Y, Cui H, Liu J, Wu Y, Cheng Y, et al. WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. Nucleic Acids Res. 2018;46(W1):W71–5. https://doi.org/10.1093/nar/gky400.

40. Kohli GS, John U, Van Dolah FM, Murray SA. Evolutionary distinctiveness of fatty acid and polyketide synthesis in eukaryotes. ISME J. 2016;10(8):1877–90. https://doi.org/10.1038/ismej.2015.263.

41. Beluzic R, Cuk M, Pavkov T, Baric I, Vugrek O. S-Adenosylhomocysteine hydrolase (AdoHcyase) deficiency: enzymatic capabilities of human AdoHcyase are highly effected by changes to codon 89 and its surrounding residues. Biochem Biophys Res Commun. 2008;368(1):30–6. https://doi.org/10.1016/j.bbrc.2008.01.042.

42. Guerrier S, Plattner H, Richardson E, Dacks JB, Turkewitz AP. An evolutionary balance: conservation vs innovation in ciliate membrane trafficking. Traffic. 2017;18(1):18–28. https://doi.org/10.1111/tra.12450.

43. Aquilini E, Cova MM, Mageswaran SK, Dos Santos PN, Sparvoli D, Penarete-Vargas DM, et al. An Alveolata secretory machinery adapted to parasite host cell invasion. Nat Microbiol. 2021;6(4):425–34. https://doi.org/10.1038/s41564-020-00854-z.

44. Lasek-Nesselquist E, Johnson MD. A phylogenomic approach to clarifying the relationship of mesodinium within the ciliophora: a case study in the complexity of mixed-species transcriptome analyses. Genome Biol Evol. 2019;11(11):3218–32. https://doi.org/10.1093/gbe/evz233.

45. Altenburger A, Cai H, Li Q, Drumm K, Kim M, Zhu Y, et al. Limits to the cellular control of sequestered cryptophyte prey in the marine ciliate Mesodinium rubrum. ISME J. 2021;15(4):1056–72. https://doi.org/10.1038/s41396-020-00830-9.

46. Li Z, Wang X, Zhang Y, Yu Z, Zhang T, Dai X, et al. Genomic insights into the phylogeny and biomass-degrading enzymes of rumen ciliates. ISME J. 2022. https://doi.org/10.1038/s41396-022-01306-8.

47. Zhang Q, Simpson A, Song W. Insights into the phylogeny of systematically controversial haptorian ciliates (Ciliophora, Litostomatea) based on multigene analyses. Proc Biol Sci. 2012;279(1738):2625–35. https://doi.org/10.1098/rspb.2011.2688.

48. Van Dolah FM, Morey JS, Milne S, Ung A, Anderson PE, Chinain M. Transcriptomic analysis of polyketide synthases in a highly ciguatoxic dinoflagellate, Gambierdiscus polynesiensis and low toxicity Gambierdiscus pacificus, from French Polynesia. PLoS One. 2020;15(4):e0231400. https://doi.org/10.1371/journal.pone.0231400.

49 Costa Torres AF, Dantas RT, Toyama MH, Diz Filho E, Zara FJ, de Rodrigues Queiroz MG, et al. Antibacterial and antiparasitic effects of Bothrops marajoensis venom and its fractions: phospholipase A2 and L-amino acid oxidase. Toxicon. 2010;55(4):795–804. https://doi.org/10.1016/j.toxicon.2009.11.013.

50. Wang F, Li R, Xie M, Li A. The serum of rabbitfish (Siganus oramin) has antimicrobial activity to some pathogenic organisms and a novel serum l-amino acid oxidase is isolated. Fish Shellfish Immunol. 2011;30(4):1095–108. https://doi.org/10.1016/j.fsi.2011.02.004.

51. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

52. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170.

53. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12(1):59–60. https://doi.org/10.1038/nmeth.3176.

54. Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. Nucleic Acids Res. 2010;38(12): e132. https://doi.org/10.1093/nar/gkq275.

55 Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res. 2006;34(Web Server issue):W435–9. https://doi.org/10.1093/nar/gkl200.

56. Chen X, Wang Y, Sheng Y, Warren A, Gao S. GPSit: aAn automated method for evolutionary analysis of nonculturable ciliated microeukaryotes. Mol Ecol Resour. 2018;18(3):700–13. https://doi.org/10.1111/1755-0998.12750.

57. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74. https://doi.org/10.1093/molbev/msu300.

58. De Bie T, Cristianini N, Demuth JP, Hahn MW. CAFE: a computational tool for the study of gene family evolution. Bioinformatics. 2006;22(10):1269–71. https://doi.org/10.1093/bioinformatics/btl097.

59. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–7. https://doi.org/10.1089/omi.2011.0118.

60. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol. 2011;29(7):644–52. https://doi.org/10.1038/nbt.1883.

Publisher's Note