Molecular Ecology

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE

Linking Protistan Trophic Indicators to Phylogenetic Identity: From Single-Cell Grazing Evidence to Biogeographic Context

Luping Bi[1,2,3] | Ying Wang[1,2,3] | Fang Shi[1,2,3] | Xin Liu[4] | Ramiro Logares[5] | Nengwang Chen[2,3] | Dapeng Xu[4] | Bangqin Huang[2,4] | Ping Sun[1,2,3]

[1]Key Laboratory of the Ministry of Education for Coastal and Wetland Ecosystem, College of the Environment and Ecology, Xiamen University, Xiamen, China | [2]National Observation and Research Station for the Taiwan Strait Marine Ecosystem, Xiamen University, Zhangzhou, China | [3]Fujian Province Key Laboratory for Coastal Ecology and Environmental Studies, College of the Environment and Ecology, Xiamen University, Xiamen, China | [4]State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China | [5]Institute of Marine Sciences (ICM), CSIC, Barcelona, Catalonia, Spain

Correspondence: Dapeng Xu (dapengxu@xmu.edu.cn) | Bangqin Huang (bqhuang@xmu.edu.cn) | Ping Sun (psun@xmu.edu.cn)

Received: 2 December 2025 | **Revised:** 24 June 2026 | **Accepted:** 16 July 2026

Keywords: biogeographic occurrence | mixotrophy | phylogenetic identity | protistan grazing | single-cell PCR | trophic indicators

ABSTRACT

Protists form the foundation of aquatic food webs and drive global nutrient cycles, yet distinguishing which species photosynthesize, graze or do both remains a major challenge because most are uncultivable and community surveys seldom resolve species-level traits. We developed a field grazing-scPCR framework that integrates short-term grazing assays, single-cell microscopy and 18S rRNA sequencing to link morphology, fluorescence-based trophic indicators, ingestion evidence and phylogenetic identity in 21 individually isolated protistan cells spanning freshwater to oceanic ecosystems. Using a conservative, phylogeny-informed classification, this approach confirmed constitutive mixotrophs (*Cryptomonas curvata*, *Poterioochromonas malhamensis*), identified a candidate non-constitutive mixotroph within Katablepharidaceae, and showed that prey-derived fluorescence can overestimate mixotrophy in natural assemblages. Two *C. curvata* isolates exhibited contrasting states, an active grazer with plastid autofluorescence and a non-grazing, aflagellate cyst retaining plastid fluorescence, highlighting the limits of single-time-point assays. Linking single-cell observations to MetaPR2 and Tara Oceans exact-match records further placed trophically characterized taxa in a broader biogeographic context. This framework advances species-level resolution of protistan trophic diversity in nature while underscoring the need to interpret fluorescence and ingestion signals in phylogenetic and ecological context.

1 | Introduction

Protists play diverse and essential roles in aquatic ecosystems, functioning as autotrophs, heterotrophs and mixotrophs, and contributing fundamentally to nutrient cycling and energy flow (Sanders 1991). Through their grazing activities, heterotrophic and mixotrophic protistan grazers exert significant regulatory pressure on bacterial communities, influencing microbial abundance, community composition and metabolic activity (Jürgens and Massana 2008). They also act as important trophic intermediates, transferring carbon and nutrients from microbial prey to micro- and mesozooplankton, particularly in oligotrophic environments (Traboni et al. 2026). However, trophic mode is not strictly determined by broad taxonomic affiliation. Heterotrophy and mixotrophy occur across diverse protistan lineages (Stoecker et al. 2017), and the degree of heterotrophy or reliance on prey ingestion is often species-specific rather than group-specific (Jeong et al. 2005). This diversity reflects a long evolutionary history of endosymbiosis and functional transitions between phototrophy and phagotrophy. Eukaryotic

photosynthesis is thought to have originated through phagocytosis, in which engulfed cyanobacteria became permanent endosymbionts rather than being digested (Mansour and Anestis 2021). Over time, some lineages lost phagotrophic capacity and evolved into obligate phototrophs, while others retained both feeding and photosynthetic functions, evolving into flexible mixotrophs (Mansour and Anestis 2021). Experimental, empirical and modelling studies further show that mixotrophy can enhance trophic transfer efficiency, improve food quality for zooplankton grazers and reshape planktonic food-web structure (Ptacnik et al. 2004; Ward and Follows 2016; Traboni et al. 2021).

Despite their ecological importance, identifying protistan nutritional strategies in natural systems remains a major challenge. A commonly used method involves incubating natural protistan assemblages with fluorescently labelled bacteria or latex beads (FLBs or LBs) and assessing ingestion via epifluorescence microscopy (Beisner et al. 2019; Li et al. 2024; Hahn Martin and Höfle Manfred 1999; McManus and Okubo 1991). Pigmented protists are typically classified as phototrophs based on chloroplast autofluorescence, while those bead/bacteria-ingesting cells are considered phagotrophs (Stoecker et al. 2017). Although effective for estimating community-level grazing, this approach treats the protist community as a 'black box', lacking species-level resolution and obscuring the functional roles of individual taxa (Beisner et al. 2019). This limitation is particularly critical given that the vast diversity of marine protists, especially small flagellates (2–5 μm), are key bacterial grazers (Hahn Martin and Höfle Manfred 1999). Many of these taxa are uncultured or difficult to maintain in laboratory conditions, hindering efforts to characterize their trophic strategies and ecological functions (Massana 2015). Most research has focused on a limited set of cultured species under controlled conditions (Beisner et al. 2019), leaving the complexity of natural protistan interactions largely unexplored. To address these gaps, new approaches are needed that can directly link grazing activity with species-level phylogenetic identity at the single-cell level, enabling in situ resolution of protistan functional diversity within microbial food webs.

Single-cell polymerase chain reaction (scPCR) has emerged as a powerful tool for molecular identification at the single-cell level (Lynn and Pinheiro 2009). By isolating individual cells under a microscope and amplifying marker genes such as 18S rRNA, scPCR enables researchers to characterize previously uncultured protists without the need for cultivation (Lynn and Pinheiro 2009; Auinger Barbara et al. 2008). Cells can be preserved using fixatives like ethanol (Marín et al. 2001), glutaraldehyde (Xia et al. 2013), Lugol's solution (Auinger Barbara et al. 2008) or osmium tetroxide (OsO_4) (Takano and Horiguchi 2006), allowing recovery of both morphological and genetic information. When coupled with field-based grazing incubations, scPCR uniquely can link fluorescence-based trophic indicators, ingestion evidence, morphology and phylogenetic identity from the same individual cell, offering species-resolved insight into protist ecology in situ.

Assigning trophic mode from field-collected cells also requires caution because common microscopic indicators can be ambiguous. Pigmented cells are often treated as phototrophs, and cells showing both pigment autofluorescence and ingested prey are often interpreted as mixotrophs (Beisner et al. 2019). Protistan mixotrophy spans a continuum from constitutive mixotrophs (CMs), which possess heritable plastids, to non-constitutive mixotrophs (NCMs), which acquire photosynthetic capacity transiently through mechanisms such as retaining functional prey plastids or maintaining photosymbiotic associations with specific algal partners (Mitra et al. 2016). Red autofluorescence may also arise when heterotrophic protists ingest algal prey and temporarily contain prey-derived plastids or pigments (Flynn et al. 2019). Therefore, distinguishing intrinsic plastids, retained prey plastids, photosymbionts and recently ingested algal prey is difficult when relying solely on chlorophyll autofluorescence or observed ingestion. Integrating phylogenetic information can help distinguish red autofluorescence associated with plastid-bearing lineages from signals occurring in lineages without known intrinsic plastids. Although this approach cannot by itself distinguish retained functional prey plastids from recently ingested algal prey, it reduces ambiguity and improves the reliability of trophic assignments.

To address these challenges, we developed an integrative method, 'field grazing-scPCR', which combines short-term field grazing incubations with fluorescent latex bead, single-cell microscopy (both fluorescence and differential interference contrast [DIC]), and scPCR targeting the 18S rRNA gene. This workflow links morphology, fluorescence-based trophic indicators, ingestion evidence and phylogenetic identity from the same individual cell. Applied across freshwater, estuarine, coastal and open-ocean environments, this method provides species-level insight into protistan trophic diversity in natural assemblages. We further linked trophically characterized single-cell isolates to exact sequence matches in MetaPR2 and Tara Oceans datasets, thereby placing these molecularly defined taxa in a broader biogeographic context. This analysis was designed to trace the occurrence and whole-community relative read abundance of the taxa characterized in this study, rather than to infer global distributions of mixotrophic strategies.

2 | Materials and Methods

2.1 | Sample Collection and Short-Term Grazing Experiments

Surface water samples were systematically collected from 14 hydrologically distinct sites spanning a freshwater-to-marine gradient in the Jiulong River estuarine system (Figure 1), southern Fujian Province, China. Sampling occurred in spring (April 2023), autumn (November 2023) and winter (January 2024) (Table S1). The Jiulong River, the region's second-largest, is fed by two major tributaries, Beixi and Xixi, which converge before discharging into Xiamen Bay and ultimately the Taiwan Strait. Surface water samples from the Beixi and Xixi river sites were collected using acid-cleaned buckets, whereas samples from the Jiulong River estuary, Xiamen Bay and Taiwan Strait were collected using CTD-mounted Niskin bottles. At each site, 4 L of surface water were pre-filtered (200 μm nylon mesh) to remove mesozooplankton and divided into duplicate 2.4 L acid-washed polycarbonate bottles. Fluorescent latex beads (LBs, 0.52 μm, carboxylate-modified

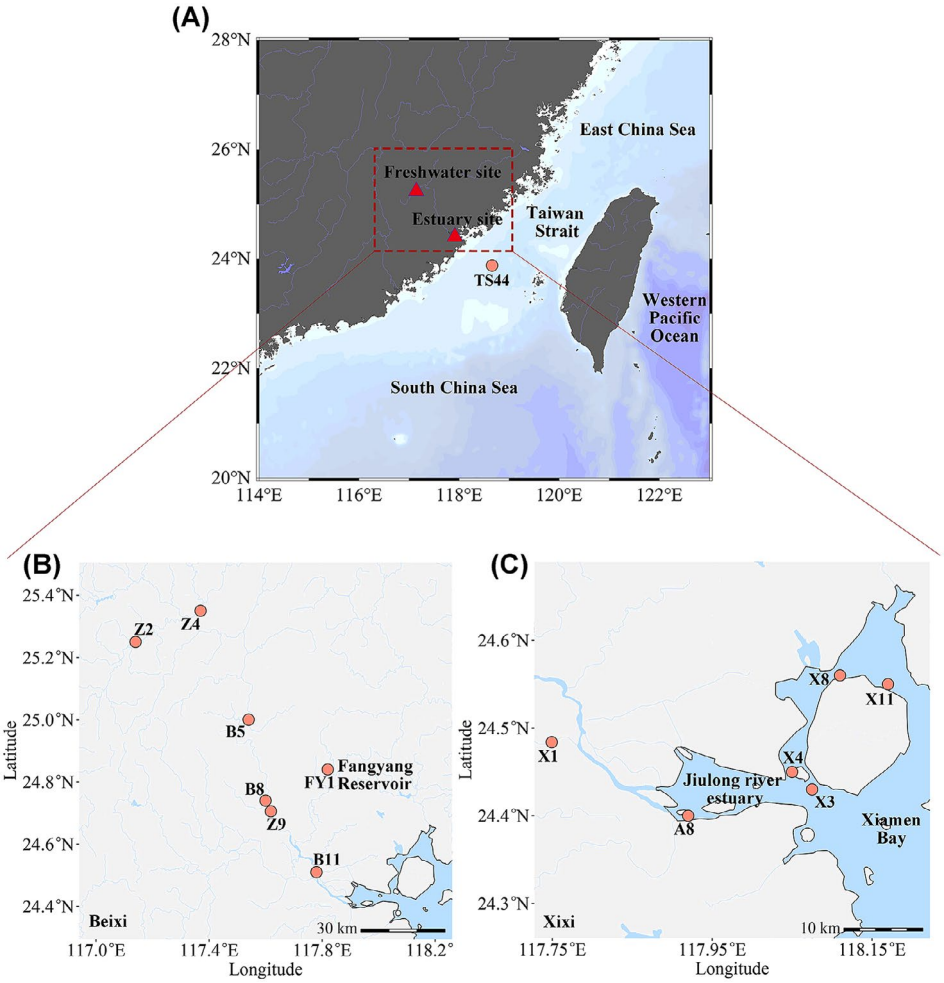

FIGURE 1 | Sampling locations across the freshwater–estuarine–coastal–marine gradient in south-eastern China and adjacent waters. (A) Regional map showing the freshwater, estuarine/coastal and Taiwan Strait sampling areas. (B) Freshwater stations in the Beixi River system and Fangyang Reservoir. (C) Estuarine and coastal stations in the Xixi River, Jiulong River estuary and Xiamen Bay. Station codes correspond to the single-cell isolates characterized in this study and listed in Table 1.

polystyrene; Sigma-Aldrich, Missouri, USA) were added as bacterial-sized prey surrogates. LBs were selected because their uniform particle size and strong fluorescence facilitate microscopic detection of particle ingestion, and because they contain no exogenous prey DNA, reducing the risk that downstream single-cell PCR amplifies DNA from labelled prey. The 0.52 μm bead size was intended to target uptake of bacterial-sized particles rather than to represent the full natural prey spectrum available to protists. The amount of LBs added was estimated from historical in situ bacterial abundance data from the Jiulong River–Xiamen Bay–Taiwan Strait long-term observation region, where bacterial abundance has been monitored seasonally in the Beixi–Xixi rivers, Jiulong River estuary and Xiamen Bay, and annually in the Taiwan Strait. LBs were generally added at tracer-level concentrations corresponding to approximately 15% of expected in situ bacterial abundance. Across these monitoring records, the estimated bacterial abundance used to guide bead additions ranged from 1.8×10^5 to 5.5×10^6 cells mL^{-1}, corresponding to final LB concentrations of 2.7×10^4 to 8.2×10^5 beads mL^{-1} in the grazing incubations. A relatively low tracer addition was used to reduce artificial enhancement of prey encounter rates and to keep prey conditions close to natural background levels. This conservative design may underestimate ingestion by slow-feeding cells, prey-selective cells or cells with low encounter probabilities during short incubations. Incubations lasted 1 h, either in situ (for estuarine, bay and strait samples) or under laboratory-simulated conditions immediately after sample transport for freshwater sites. Following incubation, 50 mL subsamples were centrifuged at 4000 rpm for 5 min to concentrate cells and remove excess water. The centrifuged samples were then fixed with ethanol to a final concentration of at least 80% v/v for subsequent microscopic and molecular analyses (Figure 2). Most single-cell molecular work was performed on ethanol-fixed samples. Ethanol fixation was selected as a practical compromise for the combined fluorescence microscopy–single-cell PCR workflow, because it retained sufficient gross cellular morphology for documenting cell size, shape and visible flagellar traits, preserved fluorescence signals needed to detect chlorophyll autofluorescence and ingested LBs, and enabled downstream DNA extraction and PCR amplification. Ethanol fixation can cause cell shrinkage and distortion; therefore, fine morphological characters were not used as the sole basis for taxonomic identification. Final taxonomic placement and trophic interpretation were based on integration of morphology, fluorescence-based

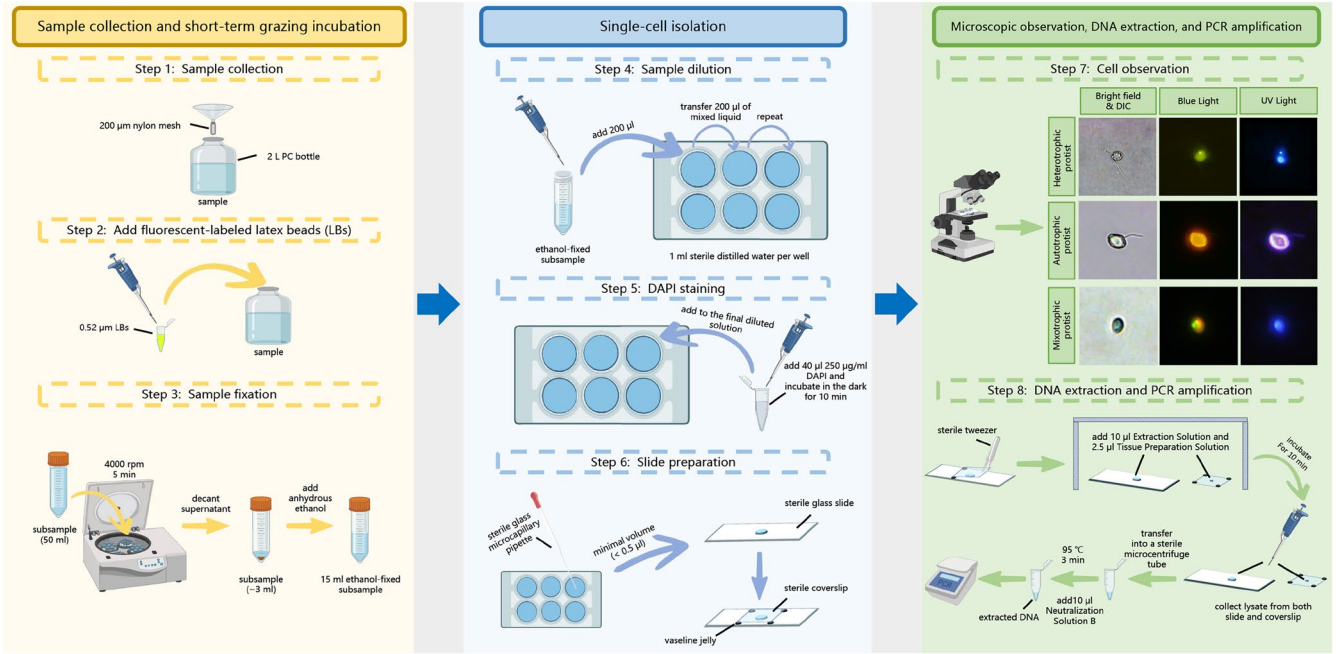

FIGURE 2 | Workflow for single-cell field grazing and molecular identification. Flowcharts depicting the procedures for sample collection, short-term grazing incubations, single-cell isolation and microscopy, and PCR amplification followed by Sanger sequencing of 18S rRNA gene.

trophic indicators, LB ingestion and 18S rRNA phylogenetic identity. Lugol's iodine was not selected because its staining and background coloration can interfere with fluorescence-based detection of chlorophyll autofluorescence and ingested LBs. Formalin was not used for the main workflow because cross-linking fixation can reduce the efficiency and consistency of downstream single-cell DNA extraction and PCR amplification. Glutaraldehyde fixation was tested in preliminary trials, but glutaraldehyde-fixed samples yielded lower and less consistent DNA extraction and PCR amplification success. Environmental parameters were measured from water samples collected at the same sites and dates as the grazing incubations. Temperature, salinity, pH and dissolved oxygen were measured in situ using either a Seabird SBE9plus CTD profiler or a YSI EXO2 multiparameter sonde equipped with the appropriate sensors. Additional water samples were collected for nutrient and chlorophyll *a* analyses, including dissolved inorganic nitrogen, dissolved inorganic phosphorus and surface chlorophyll *a*. These measurements were used to characterize the environmental context of each single-cell isolate across the freshwater–estuarine–marine gradient. Because the isolated cells were obtained through targeted single-cell isolation rather than a quantitative abundance survey, these data were not used to infer habitat preferences or environmental drivers.

2.2 | Single-Cell Isolation and Microscopic Observation

A 200 μL aliquot of each ethanol-fixed subsample was sequentially diluted with 1 mL of sterile distilled water in a multiwell plate. Dilution was repeated until a target working concentration of approximately 50 cells mL^{-1} was reached. This concentration was operationally defined as suitable for single-cell isolation because small transferred droplets generally contained zero or one candidate cell under low-magnification screening, thereby reducing the likelihood of co-isolating additional cells or visible cell fragments or debris. For reservoir samples, three sequential dilutions were typically required to reach approximately 50 cells mL^{-1} (Figure 2). Subsequently, 40 μL of 250 μg mL^{-1} 4′,6-diamidino-2-phenylindole (DAPI) was added and incubated in the dark for 10 min. A small aliquot (< 0.5 μL) of the DAPI-stained sample was carefully transferred onto an autoclaved slide using a sterile glass microcapillary pipette. A sterile coverslip was placed on top, supported by Vaseline jelly spots to prevent compression and sample evaporation. To ensure accurate single-cell isolation, we employed a two-step microscopy-based screening protocol. First, samples were pre-screened at low magnification (40×; Olympus BX51, Tokyo, Japan) to rapidly identify candidate cells. Under UV excitation, DAPI-stained nuclei appeared as discrete bright signals against a dark background, enabling rapid identification of individual nucleated cells. Concurrently, blue-light excitation revealed chloroplasts as intense red autofluorescent signals, facilitating distinction between single cells and multiple cells or debris. Any field containing multiple cells or debris was discarded, and a new preparation was made. Second, final confirmation of single-cell isolation was performed at 100× oil-immersion magnification. At this higher magnification, each candidate cell was examined under bright-field, UV and blue-light illumination to verify the presence of exactly one cell and the absence of visible additional cells, ambiguous particles or debris. Only microscopically verified single-cell preparations were advanced to downstream DNA extraction and PCR amplification. Single-cell trophic indicators were recorded from the same microscopically verified cell, including red autofluorescence and LB ingestion signal. Bright-field and differential interference contrast (DIC) microscopy were used to document morphological traits, including cell size, shape and flagella characteristics (Figure 2). Cells containing

TABLE 1 | Morphological, trophic and phylogenetic characterization of single-cell protistan isolates. Final trophic classifications were based on a conservative, phylogeny-informed framework integrating cell morphology, red autofluorescence, latex bead ingestion and 18S rRNA phylogenetic placement. HNF, ANF, CM and cNCM refer to heterotrophic nanoflagellate, autotrophic nanoflagellate, constitutive mixotroph and candidate non-constitutive mixotroph, respectively. Red autofluorescence was not treated as evidence of intrinsic chloroplasts unless supported by phylogenetic placement within plastid-bearing lineages. The number of flagella was determined from morphological observations in this study.

Number	Sample ID	Taxonomic assignment	Trophic classification	Red autofluorescence	Ingested LBs	Cell length (μm)	Cell width (μm)	Flagellum	Cell shape	Accession number
1	B5C8	*Teleaulax* sp.	ANF	Yes	None	6	3	1	Teardrop-shaped	PQ770946
2	B5C4	*Chlamydomonas* sp.	ANF	Yes	None	9	5.6	1	Ellipsoid-shaped	PQ770952
3	X11GD2	Chrysomerophyceae sp.	ANF	Yes	None	8	5	1	Kidney-shaped	PQ770943
4	FY1ET3	*Cryptomonas curvata*	Non-grazing plastid-bearing state	Yes	None	8	6.7	Undetectable	Spheroid-shaped	PQ770950
5	B5C6	*Cryptomonas curvata*	CM	Yes	1	18	10	Undetectable	Elongated ovoid	PQ770938
6	Z91	*Poterioochromonas malhamensis*	CM	Yes	2	5.1	5.0	2	Spheroid-shaped	PV931815
7	X1ET2	Katablepharidaceae sp.	cNCM	Yes (cell-filling)	5	8	8	2	Spheroid-shaped	PQ770948
8	Z4GD1	*Goniomonas* sp.	HNF	No	1	6	3	1	Egg-shaped	PQ770951
9	X11DL1	*Massisteria* sp.	HNF	No	4	4	3	2	Pear-shaped	PQ770942
10	B11C10	Thaumatomastigidae sp.	HNF	No	1	6	4	2	Teardrop-shaped	PQ770939
11	535	*Boroka* sp.	HNF	No	1	5	4.5	2	Spheroid-shaped	PQ770936
12	FY1GD1	*Spumella* sp.	HNF	No	1	4	4	2	Spheroid-shaped	PQ770955
13	X11LL1	*Pteridomonas* sp.	HNF	No	1	5	4	4	Dome-shaped	PQ770944
14	B8GD1	*Telonema* sp.	HNF; likely prey-derived red autofluorescence	Yes	None	6	5	Undetectable	Egg-shaped	PQ770954
15	A8ET3	Thaumatomastigidae sp.	HNF; likely prey-derived red autofluorescence	Yes	2	4	3	Undetectable	Egg-shaped	PQ770937

(Continues)

TABLE 1 | (Continued)

Number	Sample ID	Taxonomic assignment	Trophic classification	Red autofluorescence	Ingested LBs	Cell length (μm)	Cell width (μm)	Flagellum	Cell shape	Accession number
16	Z2ET1	Siluariidae sp.	HNF; likely prey-derived red autofluorescence	Yes	1	3	2	2	Kidney-shaped	PQ770953
17	X31	Bicosoecidae-related environmental lineage	HNF; likely prey-derived red autofluorescence	Yes	2	4	2	1	Pear-shaped	PQ770945
18	X8GD4	Stramenopiles sp. MAST-4	HNF; red autofluorescence unresolved/ prey-derived	Yes	3	3	3	2	Spheroid-shaped	PQ770957
19	X4ET5	Stramenopiles sp. MAST-12	HNF; red autofluorescence unresolved/ prey-derived	Yes	1	3	2.5	1	Pear-shaped	PQ770941
20	B11C4	Paraphysomonadaceae sp.	HNF; plastid origin unresolved	Yes	1	7	7	Undetectable	Spheroid-shaped	PQ770947
21	TS44GD1	*Cafeteria* sp.	HNF; likely prey-derived red autofluorescence	Yes	2	3	3	2	Spheroid-shaped	PQ770956

ingested LBs were treated as showing direct evidence of particle ingestion, whereas red autofluorescence was interpreted together with 18S rRNA phylogenetic placement to distinguish plastid-bearing lineages from heterotrophic lineages with possible prey-derived autofluorescence. Final trophic classifications were therefore based on an integrated assessment of morphology, fluorescence-based trophic indicators, LB ingestion and phylogenetic identity. Because LB ingestion was assessed using bacterial-sized latex beads during a short incubation, the absence of observed LB ingestion was not used as evidence against phagotrophic capacity. This framework detects particle ingestion and fluorescence-based trophic indicators, but does not assess osmotrophy, dissolved organic matter uptake or other non-engulfment nutritional pathways.

2.3 | DNA Extraction, PCR Amplification and Sequencing

Following microscopic documentation of morphology, red autofluorescence and LB ingestion, the coverslip was gently lifted using sterile tweezers. To extract DNA from single cells, 10 μL of extraction solution and 2.5 μL of tissue preparation solution from the REDExtract-N-Amp Tissue PCR Kit (Sigma-Aldrich, Missouri, USA) were added directly to the liquid drops on both the slide and coverslip. The mixture was incubated at room temperature for 10 min to allow lysis. The lysate was then carefully collected using a sterile micropipette and transferred into a sterile 1.5 mL microcentrifuge tube, heated at 95°C for 3 min and neutralized with 10 μL of Neutralization Solution B. All equipment, including slides, coverslips, tweezers and micropipettes, was autoclaved before use. The 18S rRNA gene was amplified using a two-step PCR approach. The first-step PCR was performed using primers Euk F (5′-ACCTGGTTGATCCTGCCAG-3′) and Euk R (5′-TGATCCTTCYGCAGGTTCAC-3′) (Moon-van der Staay et al. 2001), with 2 μL of single-cell DNA lysate as template. The second-step PCR was conducted using primers 63F (5′-ACGCTTGTCTCAAAGATTA-3′) and 1818R (5′-ACGGAAACCTTGTTACGA-3′) (Li et al. 2022), with 1 μL of the first-step PCR product as template. Each 25 μL PCR reaction contained 0.5 μL of each 10 μM primer, 12.5 μL of Premix Taq 2× PCR Master Mix containing Taq DNA Polymerase (Ex Taq Version 2.0; Takara Bio, Kyoto, Japan), template DNA and nuclease-free water to volume. The cycling programme consisted of an initial denaturation at 95°C for 2 min, followed by 20 cycles for the Euk F/R reaction or 30 cycles for the 63F/1818R reaction, with denaturation at 98°C for 10 s, annealing at 55°C (Euk F/R) or 50°C (63F/1818R) for 30 s and extension at 72°C for 90 s. A final extension was performed at 72°C for 10 min. PCR products were purified and directly Sanger sequenced (Sangon Biotech) to both recover the target amplicon and screen for contamination. In total, 52 microscopically selected single-cell preparations were processed for DNA extraction and PCR. Of these, 37 preparations yielded PCR or sequencing results assigned to flagellates (71.2%), whereas the remaining preparations produced no visible PCR band ($n = 9$, 17.3%) or were assigned to non-target sequences, including fungi ($n = 5$, 9.6%) and terrestrial plants ($n = 1$, 1.9%). Among the 37 flagellate-assigned preparations, 21 produced clean 18S rRNA chromatograms suitable for phylogenetic analysis (56.8%), whereas 16 showed overlapping peaks and were excluded as mixed chromatograms (43.2%). Inspection of the excluded mixed chromatograms indicated non-target contributions from terrestrial plant, diatom, or fungal sequences. The final dataset therefore consisted of 21 high-quality single-cell 18S rRNA sequences of approximately 1700 bp, which were retained for phylogenetic and trophic analyses.

2.4 | Phylogenetic Analysis

The 18S rRNA sequences from the 21 isolates were aligned with closely related SSU rDNA sequences retrieved from GenBank (accession numbers listed in Table 1 and Figures S1–S5). Initial alignments were conducted using ClustalW in BioEdit 7.2.5 (Hall 1999), followed by manual refinement in SeaView v4.0 (Gouy et al. 2010). Phylogenetic trees were constructed using maximum likelihood (ML) and Bayesian inference (BI) methods. ML analyses were performed with 1000 bootstrap replicates, and BI analyses employed Markov chain Monte Carlo (MCMC) simulations, run for 3,000,000 generations, sampling every 100 generations. The first 10,000 trees were discarded as burn-in, and posterior probabilities were estimated from a majority-rule consensus tree. The best-fit substitution model (GTR + I + G) was selected using jModelTest v2.1.7 (Posada 2008). All resulting phylogenetic trees were visualized with MEGA v.11 (Tamura et al. 2021).

2.5 | Integration of Global Sequencing Data

To place single-cell observations into a biogeographic context, rather than to infer global distributions of trophic modes, we queried two complementary resources, the metaPR2 database (https://shiny.metapr2.org/metapr2/) and the Tara Oceans global marine amplicon datasets (http://taraoceans.sb-roscoff.fr/EukDiv/), for exact sequence matches (100% identity across the entire V4 or V9 fragment) to each of the 21 isolates generated in this study. Analyses were conducted independently for the two resources to avoid potential double counting of overlapping submissions. Eight isolates met the 100% criterion. For further biogeographic analyses, we focused on three isolates that were detected in at least 80 samples within a dataset: *Cryptomonas curvata* B5C6, *Cafeteria* sp. TS44GD1, and *Boroka* sp. 535. B5C6, TS44GD1 and 535 are isolate IDs assigned to the single-cell 18S rRNA sequences generated in this study. Isolates detected in fewer than 80 samples were treated as presence-only records and are reported with per-sample raw read counts in Table S2. To account for uneven sequencing depth among libraries while avoiding interpretation of PCR-based amplicon reads as absolute abundance, we calculated the whole-community relative read abundance of each focal isolate in each sample as exact-match reads divided by total reads and expressed as a percentage. Environmental metadata (temperature, salinity, nutrients, chlorophyll *a*, picoeukaryote biomass and related variables) were retrieved from the metaPR2 portal and PANGAEA repository (doi:10.1594/PANGAEA.853810). For each focal exact-match isolate, Spearman rank correlations were used to evaluate associations between whole-community relative read abundance and environmental variables. Within each isolate, *p* values were adjusted across environmental variables using the Benjamini–Hochberg false discovery rate correction (Table S3). Geographic differences were evaluated using Wilcoxon rank-sum tests. All

analyses were performed in R 4.3.2 using the *cor.test*, *wilcox.test* and *p.adjust* functions in the R *stats* package.

3 | Results

3.1 | Integration of Morphology, Trophic Indicators and Phylogenetic Identity at the Single-Cell Level

Using the field grazing-scPCR approach, we combined field-based grazing incubations, single-cell microscopy and 18S rRNA sequencing to link morphology, fluorescence-based trophic indicators and phylogenetic identity for 21 individually isolated protistan cells from freshwater, estuarine, coastal and oceanic environments (Figure 1; Table 1 and Table S1). Each cell was isolated, imaged, screened for red autofluorescence and fluorescent latex bead (LB) ingestion, and then subjected to single-cell PCR and Sanger sequencing. Molecular identification based on BLAST searches and phylogenetic reconstruction were consistent with the morphological placement of the isolates (Figures 3–6; Figures S1–S5). Trophic assignments were made using a conservative, phylogeny-informed framework. Red autofluorescence and LB ingestion were treated as single-cell trophic indicators rather than as automatic diagnoses of trophic mode. Cells with red autofluorescence but no observed LB ingestion were classified as autotrophic/plastid-bearing candidates only when their phylogenetic placement was consistent with plastid-bearing ancestry. Cells showing both red autofluorescence and LB ingestion were classified as constitutive mixotrophs only when they belonged to lineages known to possess intrinsic plastids. In contrast, autofluorescent cells from lineages lacking evidence for intrinsic plastids were conservatively classified as heterotrophs, because their red signal could reflect prey-derived autofluorescence from recently ingested algae. Absence of observed LB ingestion was not interpreted as absence of phagotrophic capacity, because bead uptake may be affected by prey size, prey type, tracer concentration, feeding rate and cell state. These assignments refer to fluorescence and particle-ingestion indicators observed in this assay and do not exclude osmotrophy or other non-engulfment nutritional pathways. This framework yielded three autotrophs/autotrophic candidates, one non-grazing plastid-bearing *Cryptomonas curvata* stage, two confirmed constitutive mixotrophs, one candidate non-constitutive mixotroph and 14 heterotrophic nanoflagellates.

3.1.1 | Autotrophic and Non-Grazing Plastid-Bearing Cells

Three isolates showed red autofluorescence without observed LB ingestion and were assigned to plastid-bearing lineages. B5C8, affiliated with *Teleaulax* sp. within Cryptophyta, was a teardrop-shaped, uniflagellated cell (6×3 μm) that exhibited red autofluorescence but no LB ingestion (Figures 3A and 5A). B5C4, affiliated with Chlamydomonadaceae, was a larger ellipsoid cell (~9 μm) with one visible flagellum and red autofluorescence, and showed high 18S rRNA similarity to *Chlamydomonas* sp. (DQ009753; 99.65%) (Figures 3B and 5A). X11GD2, an Ochrophyta-related isolate positioned near the base of the SI clade, was kidney-shaped (8×5 μm), showed red autofluorescence, and lacked evidence of LB ingestion (Figures 3C and 6A). These three cells were therefore classified as autotrophic nanoflagellates.

FY1ET3, assigned to *Cryptomonas curvata*, also showed red autofluorescence but no observed LB ingestion (Figure 4A). Although its 18S rRNA sequence was 99.8% identical to that of the actively feeding *C. curvata* isolate B5C6 (Figure 5), FY1ET3 was morphologically distinct, appearing as a smaller spheroid cell (8×6.7 μm) lacking visible flagella. Rather than being interpreted as an actively feeding mixotroph, FY1ET3 was treated as a non-grazing autofluorescent *C. curvata* state, possibly representing a resting or cyst-like stage retaining plastid pigments.

3.1.2 | Confirmed Constitutive and Candidate Non-Constitutive Mixotrophs

Two isolates were classified as confirmed constitutive mixotrophs because they combined red chlorophyll autofluorescence, direct evidence of LB ingestion and phylogenetic placement within plastid-bearing lineages. B5C6, assigned to *Cryptomonas curvata*, was an elongated ovoid cell (18×10 μm) that showed red autofluorescence and ingestion of one LB (Figures 4B and 5). Together with its phylogenetic placement within Cryptomonadaceae, these observations support its classification as a constitutive mixotroph.

Z91 was identified as *Poterioochromonas malhamensis* within Ochromonadaceae. The cell was spheroid (5.1×5.0 μm) and biflagellated, with two unequal flagella. Under UV and blue-light excitation, it exhibited red chloroplast autofluorescence and contained two ingested fluorescent beads (Figure 4C). Phylogenetic reconstruction placed Z91 within Ochromonadaceae (Figure 6A). BLASTn analysis returned *Poterioochromonas malhamensis* (LC787607) as the top hit (100% coverage, 97.86% identity), and inspection of the top 45 non-redundant hits with ≥97% identity showed that 36 of 45 hits resolved to the genus *Poterioochromonas* using a last common ancestor approach. Combined with its diagnostic morphology and ingestion signal, Z91 was classified as a constitutive mixotroph.

X1ET2, affiliated with Katablepharidaceae, exhibited both LB ingestion and strong red autofluorescence (Figures 4D and 5A). The autofluorescent signal filled much of the cell body rather than appearing as discrete food-particle-associated signal, raising the possibility of plastid retention. The cell was spheroid (~8 μm diameter) and showed high sequence similarity to an uncultured environmental katablepharid sequence (GU068005; 99.60%). Because Katablepharidaceae are traditionally considered non-photosynthetic and because plastid functionality was not tested here, X1ET2 was therefore treated as a candidate non-constitutive mixotroph, with possible plastid retention requiring further validation.

3.1.3 | Heterotrophic Nanoflagellates With Direct Ingestion Evidence

Several isolates showed LB ingestion without red autofluorescence and were classified as heterotrophic nanoflagellates.

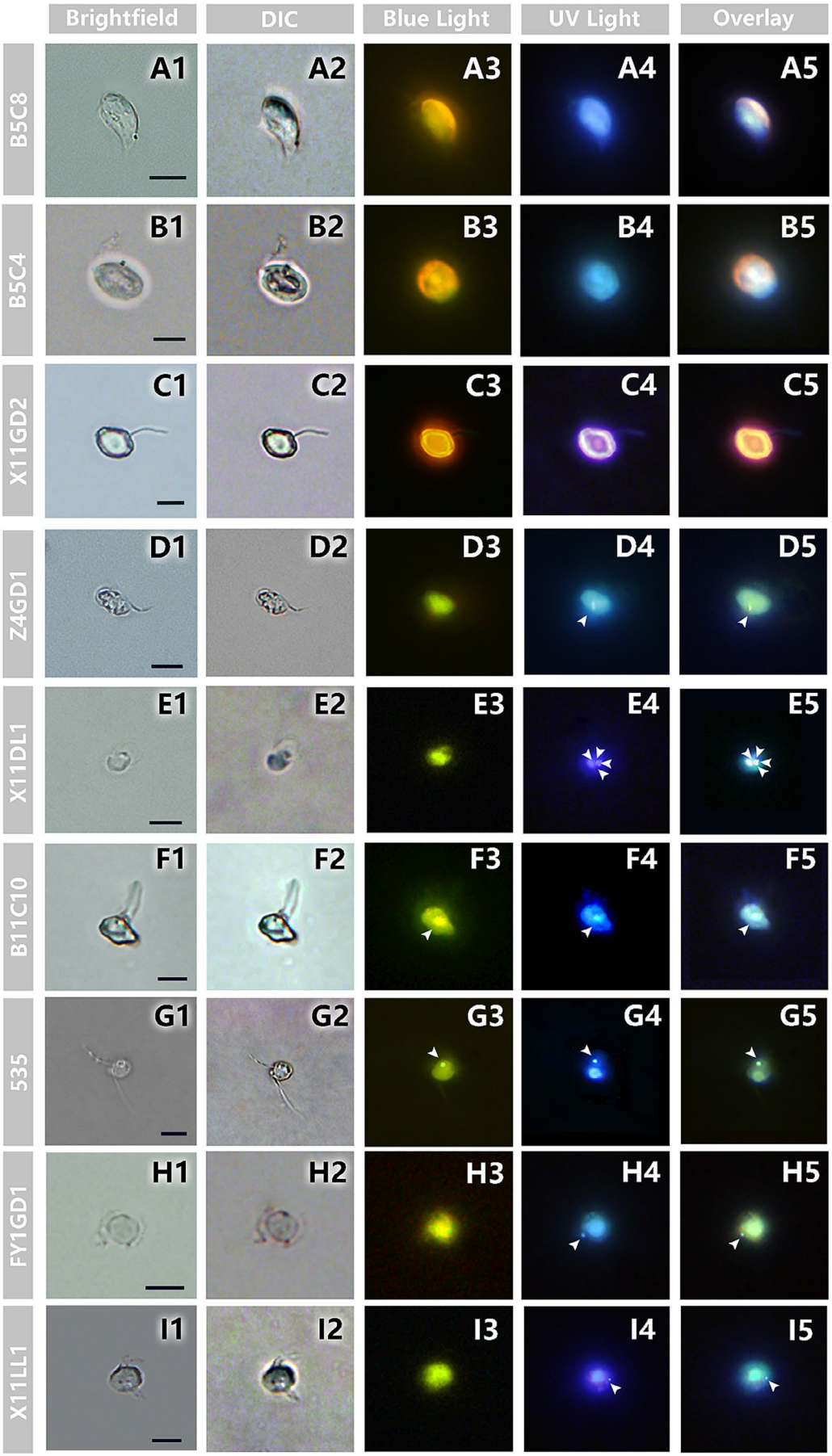

FIGURE 3 | Legend on next page.

FIGURE 3 | Microscopic identification of autotrophic and heterotrophic nanoflagellates. (A1–I1) Bright-field, (A2–I2) differential interference contrast (DIC), (A3–I3) blue-light excitation, (A4–I4) UV-light excitation and (A5–I5) combined fluorescence overlay images of (A–C) autotrophic nanoflagellates and (D–I) heterotrophic nanoflagellates, identified via field grazing-scPCR. Red autofluorescence is visible under blue-light excitation, and DAPI-stained nuclei appear blue under UV-light excitation. Arrowheads indicate ingested fluorescent beads. Scale bars are 5 μm.

Z4GD1, assigned to *Goniomonas* sp., was an egg-shaped, uniflagellated cell (6×3 μm) that ingested one LB and lacked red autofluorescence (Figures 3D and 5A). X11DL1, assigned to *Massisteria* sp., was a pear-shaped, biflagellated cell (4×3 μm) that ingested four LBs and lacked chloroplast autofluorescence (Figures 3E and 6B). Its 18S rRNA sequence grouped within Massisteriidae and showed high similarity to *Massisteria marina* (AF174369; 99.20%).

B11C10, affiliated with a Rhizaria environmental clade basal to Thaumatomastigidae, was a teardrop-shaped cell (6×4 μm) that ingested LBs but showed no red autofluorescence (Figures 3F and 6B). 535, placed within Cafeteriidae 2, was a spheroid cell (5×4.5 μm) that ingested one LB and lacked red autofluorescence (Figures 3G and 6C). Its 18S rRNA sequence showed 99.94% similarity to *Boroka* sp. (MT355113). FY1GD1, affiliated with Chromulinaceae and closely matching *Spumella* sp. (KF651119; 99.66%), was a spheroid cell (4×4 μm) that showed LB ingestion and lacked red autofluorescence, supporting its classification as heterotrophic (Figures 3H and 6A). X11LL1, a dome-shaped four-flagellated cell (5×4 μm), grouped within Actinomonadaceae and was closely related to *Pteridomonas* sp. (ON888438; 99.94%); it was also classified as heterotrophic (Figures 3I and 6A).

3.1.4 | Heterotrophic Lineages With Red Autofluorescence Likely Derived From Prey

A second set of isolates displayed red autofluorescence, in some cases together with LB ingestion, but were placed in lineages without evidence for intrinsic plastids or lacked direct evidence of chloroplast-bearing ancestry. These cells were conservatively classified as heterotrophic nanoflagellates to avoid overestimating mixotrophy from prey-derived autofluorescence.

B8GD1 showed red autofluorescence but no LB ingestion (Figure 4E). Although this signal could superficially suggest autotrophic potential, phylogenetic analysis placed B8GD1 within Telonemidae (Figure 6B), a heterotrophic lineage lacking known photosynthetic capacity. The red autofluorescence was therefore interpreted as likely prey-derived, and B8GD1 was classified as heterotrophic.

A8ET3, a Rhizaria isolate, was an egg-shaped cell (4×3 μm) that ingested two LBs and exhibited red autofluorescence (Figure 4F). Despite this combination of signals, it clustered with an uncultured Thaumatomonadida sequence (KX602143; 90.46% similarity), a group considered heterotrophic (Figure 6B). A8ET3 was therefore conservatively classified as heterotrophic, with the red signal likely reflecting recently ingested algae or prey-derived pigments.

Six stramenopile isolates also showed both red autofluorescence and LB ingestion but were conservatively classified as heterotrophs after phylogenetic evaluation. Z2ET1 was a kidney-shaped biflagellated cell (3×2 μm) that clustered within Siluariidae (Figures 4G and 6C). X31 was a pear-shaped cell (4×2 μm) with a single visible flagellum and grouped with an environmental clade sister to Bicosoecidae (Figures 4H and 6C). X8GD4 and X4ET5 were assigned to uncultured MAST-4 and MAST-12 clades, respectively (Figure 6D). X8GD4 was spheroid (3×3 μm) and biflagellated, whereas X4ET5 was pear-shaped (3×2.5 μm) with one visible flagellum (Figure 4I,J). Because MAST lineages are generally regarded as heterotrophic and stramenopiles outside photosynthetic ochrophyte lineages are not expected to possess intrinsic plastids, their red signals were interpreted as prey-derived autofluorescence, potentially reflecting recently ingested algal prey or short-term retention of prey plastids that could not be resolved with the present data.

B11C4, a spheroid Gyrista isolate (7×7 μm), clustered within an environmental clade sister to Paraphysomonadaceae in Chrysophyceae (Figures 4K and 6A). Although Chrysophyceae include both pigmented and colourless taxa, B11C4 lacked direct evidence of intrinsic chloroplasts and was closely associated with the predominantly heterotrophic lineage Paraphysomonadaceae (Figure S2). Its combined red autofluorescence and LB ingestion were therefore treated cautiously, and B11C4 was classified as heterotrophic pending future tests of plastid origin and functionality.

TS44GD1, assigned to Cafeteriidae 1, was a spheroid cell (3×3 μm) that ingested two LBs and several picoalgal cells (Figure 4L). Although red autofluorescence was observed, its placement within the well-established heterotrophic *Cafeteria* clade indicates that the signal was most likely derived from ingested phototrophic prey rather than intrinsic chloroplasts (Figure 6C). TS44GD1 was therefore classified as a heterotrophic nanoflagellate with direct evidence of both bacterial-sized particle ingestion and picoalgal prey ingestion.

3.1.5 | Environmental Context of Single-Cell Isolates Across the Freshwater–Marine Gradient

The 21 single-cell isolates were recovered from sites spanning a pronounced freshwater–estuarine–marine gradient in the Jiulong River–Xiamen Bay–Taiwan Strait system (Table S1). Salinity ranged from 0 to 33.3, temperature from 16.1°C to 25.9°C, pH from 6.84 to 9.17, dissolved oxygen from 4.78 to 10.24 mg L^{-1}, dissolved inorganic nitrogen from 4.31 to 326.5 μmol L^{-1}, dissolved inorganic phosphorus from 0.075 to 2.68 μmol L^{-1}, and surface chlorophyll *a* from 0.78 to 13.74 μg L^{-1}.

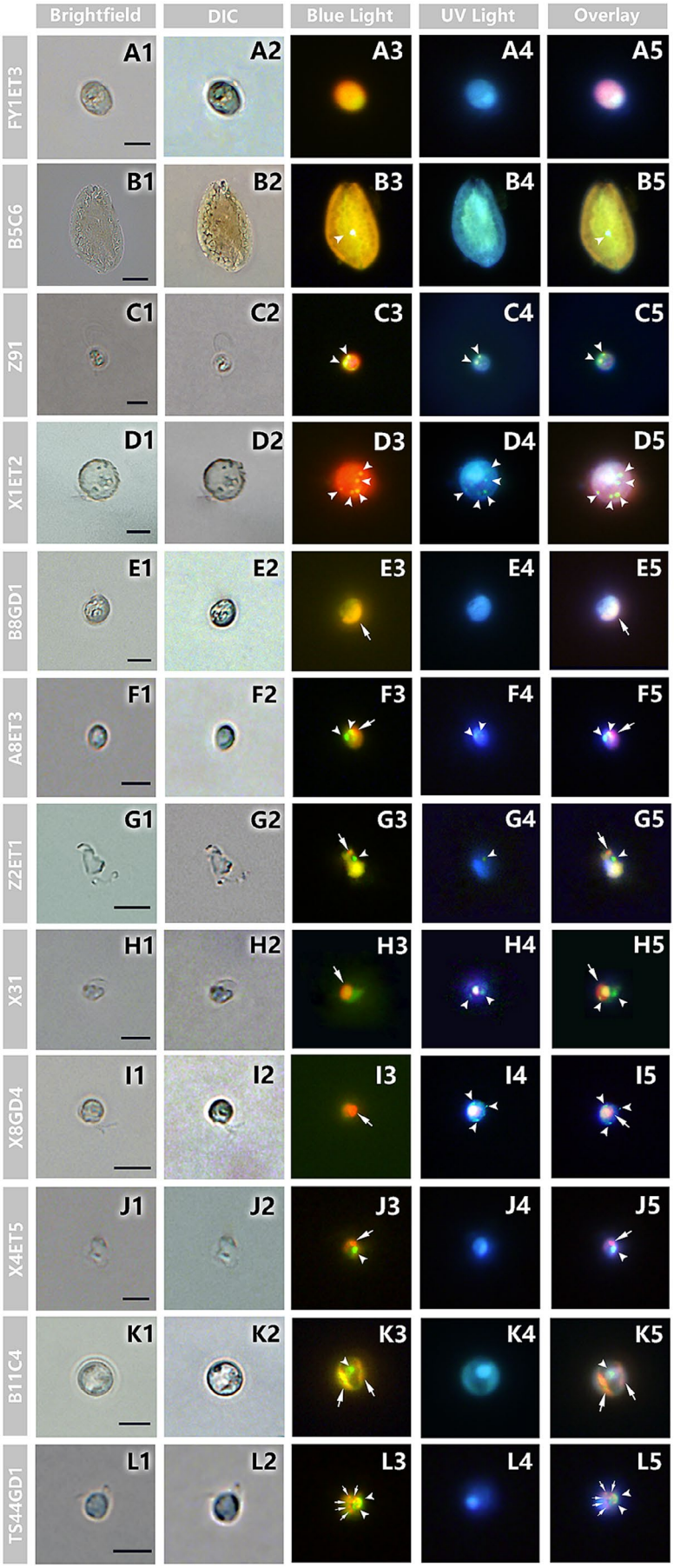

FIGURE 4 | Legend on next page.

FIGURE 4 | Microscopic identification of non-grazing autofluorescent, mixotrophic and heterotrophic nanoflagellates with unresolved or prey-derived red autofluorescence. (A1–L1) Bright-field, (A2–L2) differential interference contrast (DIC), (A3–L3) blue-light excitation, (A4–L4) UV-light excitation and (A5–L5) combined fluorescence overlay images of (A) non-grazing autofluorescent nanoflagellate, (B, C) constitutive mixotrophs, (D) a candidate non-constitutive mixotroph, and (E–L) heterotrophic nanoflagellates with unresolved or likely prey-derived red autofluorescence identified via field grazing-scPCR. Red autofluorescence is visible under blue light excitation, and DAPI-stained nuclei appear blue under UV-light excitation. Arrowheads indicate ingested fluorescent beads, and arrows indicate ingested picoalgae. Scale bars are 5 μm.

Freshwater and reservoir sites accounted for 12 isolates and included the constitutive mixotrophs *Cryptomonas curvata* B5C6 and *Poterioochromonas malhamensis* Z91, the candidate non-constitutive mixotroph X1ET2, autotrophic chlorophyte and cryptophyte cells, and several heterotrophic nanoflagellates. These freshwater isolates occurred under low-salinity conditions (0–0.6) and generally elevated dissolved inorganic nitrogen concentrations (68.2–326.5 μmol L^{-1}). The estuarine isolate A8ET3 was recovered at intermediate salinity (7.7) and relatively high nutrient concentrations, consistent with the transitional character of the Jiulong River estuary. In contrast, Xiamen Bay and Taiwan Strait isolates were recovered from high-salinity waters (28.7–33.3) and were dominated by heterotrophic lineages, including *Massisteria*, MAST, Cafeteriidae and Actinomonadaceae-related cells, together with one autotrophic Ochrophyta-related isolate.

These observations show that field grazing-scPCR can be applied across strong physicochemical gradients and can recover trophically diverse protists from freshwater, estuarine, coastal and marine waters. Because the isolates were obtained as individually selected single cells rather than standardized community counts, these data are interpreted as environmental context for the single-cell trophic assignments rather than as quantitative evidence for habitat preference.

3.2 | Whole-Community Relative Read Abundance of Exact-Match Isolates in Global Datasets

Of the 21 single-cell isolates characterized, eight had exact sequence matches at 100% identity in at least one global amplicon resource. To focus on taxa with sufficient detections for biogeographic interpretation, three prevalent isolates were retained for further analysis, namely *Cryptomonas curvata* B5C6 (312 samples), *Cafeteria* sp. TS44GD1 (166 samples) and *Boroka* sp. 535 (80 samples), which were detected in ≥ 80 samples. The remaining exact-match isolates were detected in relatively few samples (≤ 10) and are reported as presence-only records with per-sample counts in Table S2. Global patterns were evaluated using whole-community relative read abundance, calculated as exact-match reads divided by total reads in each sample and expressed as a percentage. This metric describes proportional representation in PCR-based amplicon libraries and should not be interpreted as absolute cell abundance.

Cryptomonas curvata B5C6 showed its highest mean relative read abundance in French rivers (30.40%), followed by Canadian (3.86%) and Argentinian (1.56%) sites, whereas lower values were observed in Russia (0.29%) and Arctic samples (0.07%) (Figure 7A). Its association with temperature was weak and not statistically significant (Spearman rho = 0.122, $p = 0.142$; Table S3). *Cafeteria* sp. TS44GD1 showed higher mean relative read abundance in Southern- than Northern-hemisphere ocean samples (1.37% vs. 0.47%; Wilcoxon rank-sum, $p = 0.015$; Figure 7B). Among ocean basins, mean relative read abundance was highest in the Indian Ocean (0.85%), and the South Atlantic Ocean (0.38%), followed by Mediterranean Sea (0.28%), South Pacific Ocean (0.11%), Red Sea (0.09%), North Pacific Ocean (0.06%), North Atlantic Ocean (0.04%) and Southern Ocean (0.02%) (Figure 7B). Environmental correlations were generally weak. TS44GD1 relative read abundance was weakly negatively correlated with carbon flux (Spearman rho = −0.190, $p = 0.025$), but this association was not robust after multiple-testing correction (Table S3). *Boroka* sp. 535 showed its highest mean relative read abundance in the Indian Ocean (0.032%), followed by the South Pacific Ocean and Mediterranean Sea (both 0.003%). Values were low in the South Atlantic Ocean, Red Sea and Southern Ocean (0.0005%–0.0011%) (Figure 7C). The North Atlantic record was not emphasized because only one sample was available from this region. The relative read abundance of 535 was negatively correlated with dissolved oxygen (Spearman rho = −0.393, $p < 0.001$; Table S3), and this association remained significant after false discovery rate correction (adjusted $p = 0.006$; Table S3).

4 | Discussion

4.1 | Advancing Species-Level Insights: The Power and Applicability of the Field Grazing-scPCR Approach

The field grazing-scPCR method developed in this study offers an integrative framework for linking single-cell ingestion evidence, fluorescence-based trophic indicators, morphology and phylogenetic identity. By applying this workflow to field-collected cells, the approach helps resolve trophic traits of uncultured or difficult-to-culture protists without requiring laboratory cultivation. This addresses a longstanding challenge in microbial ecology: characterizing the trophic roles of uncultured protists that dominate natural microbial communities. Capturing both trophic indicators and taxonomic identity from individual cells helps unlock the 'black box' of field-derived ecology by answering the fundamental questions: who are these cells, and what trophic roles may they play? The consistency between morphology-based placement and 18S rRNA phylogenetic identity further supports the robustness of the workflow, while the two-step microscopic screening helped reduce the likelihood of co-isolating non-target cells or debris. Its application across diverse environments, from rivers to coastal zones and open oceans, demonstrates broad applicability and positions it as a valuable tool for future investigations of species-specific interactions in microbial food webs.

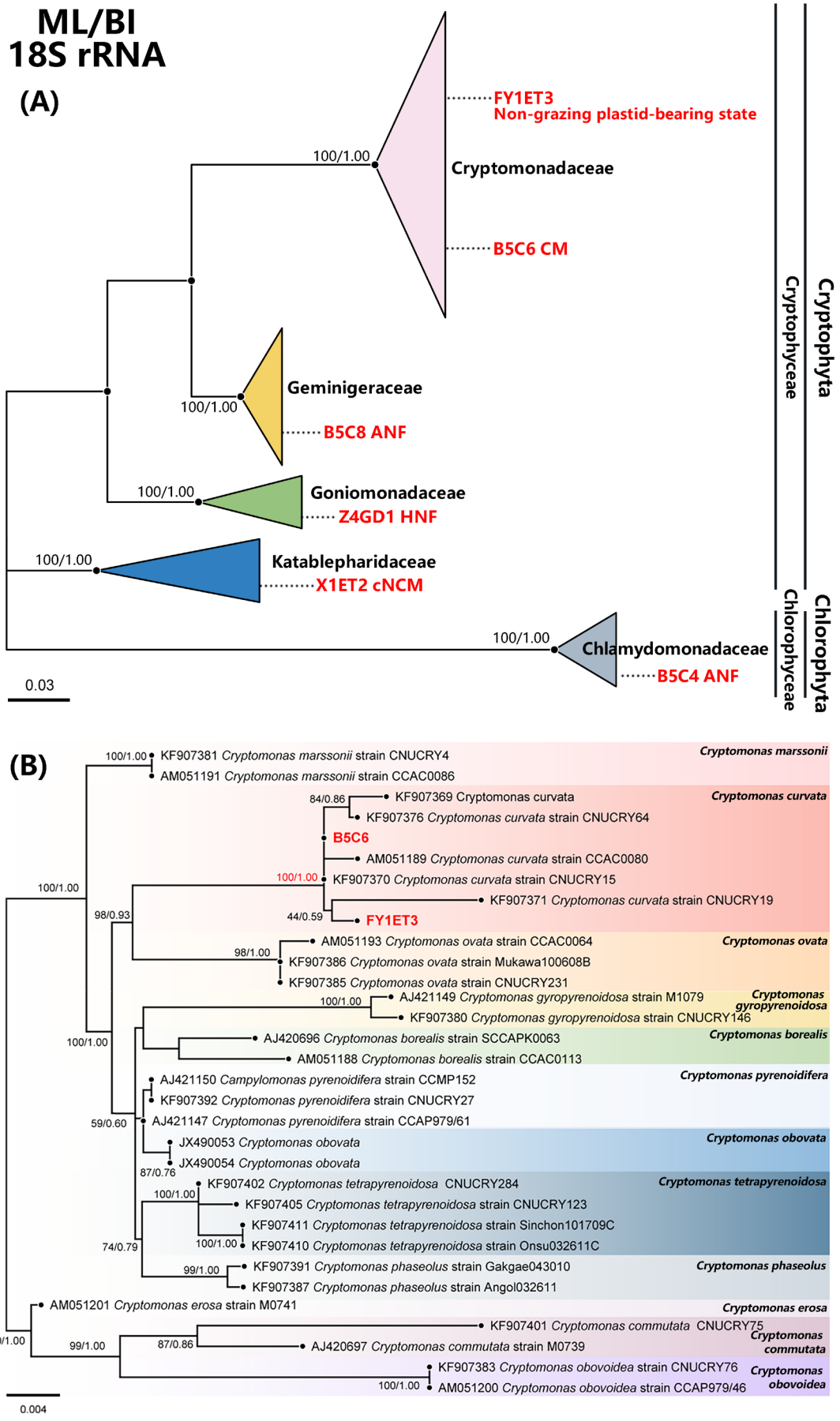

FIGURE 5 | Legend on next page.

FIGURE 5 | Phylogenetic placement of Cryptophyta, Chlorophyta and *Cryptomonas* isolates identified by Field Grazing-scPCR. Maximum likelihood (ML) tree based on 18S rRNA genes showing phylogenetic positions of species from (A) Cryptophyta and Chlorophyta, and (B) *Cryptomonas curvata* (B5C6 and FY1ET3), identified using field grazing-scPCR. Bootstrap values (ML) and posterior probabilities (Bayesian inference) are provided for each node. GenBank accession numbers are listed before species names. See Figure S1 for detailed trees. Clades distinguished within the genus *Cryptomonas* by Choi et al. (2013).

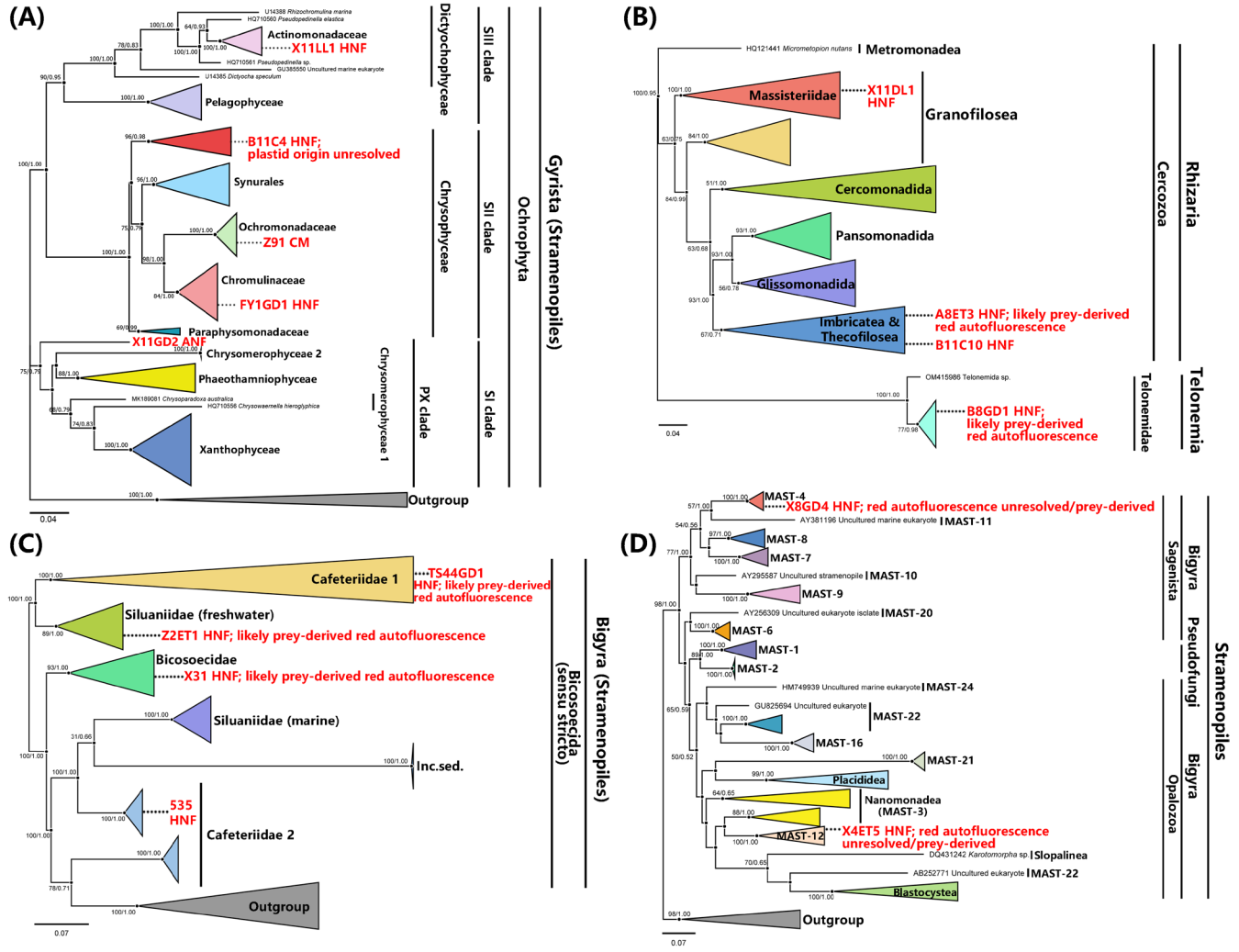

FIGURE 6 | Phylogenetic placement of Ochrophyta, Rhizaria, Telonemia, Bicosoecida and MAST isolates. Maximum likelihood (ML) tree based on 18S rRNA genes showing phylogenetic positions of species from (A) Ochrophyta, (B) Rhizaria and Telonemia, (C) Bicosoecida and (D) MAST groups identified using field grazing-scPCR. Bootstrap values (ML) and posterior probabilities (Bayesian inference) are provided for each node. GenBank accession numbers are listed before species names. See Figures S2–S5 for detailed trees.

4.2 | Insight Into Protistan Trophic Modes Inferred From the Field Grazing-scPCR Approach

4.2.1 | Interpreting Mixotrophy: Constitutive, Non-Constitutive and False Positives

Our results illustrate why fluorescence-based trophic indicators must be interpreted within a phylogenetic framework. Red autofluorescence and fluorescent-bead ingestion provided useful single-cell evidence for plastid-related signals and phagotrophic uptake, respectively, but several isolates showed that these signals cannot be interpreted as automatic diagnoses of trophic mode. We therefore applied a conservative framework that combines field grazing incubations, fluorescence microscopy and 18S rRNA phylogenetic context. Cells were classified as constitutive mixotrophs (CMs) only if they (i) displayed both red autofluorescence and fluorescent-bead ingestion, and (ii) belonged to lineages with known plastid-bearing ancestry. Autofluorescent cells from lineages lacking such ancestry were conservatively categorized as heterotrophs, unless additional morphological or molecular evidence suggested plastid retention. Using these criteria, we confirmed several CMs. Notably, isolate Z91 was identified as *Poterioochromonas malhamensis* (Ochromonadaceae, Figure 6A), a lineage well known for constitutive mixotrophy (Lie et al. 2018; Ma et al. 2018). Z91 cell ($5.1 \times 5.0\,\mu m$) exhibited both red autofluorescence and ingestion

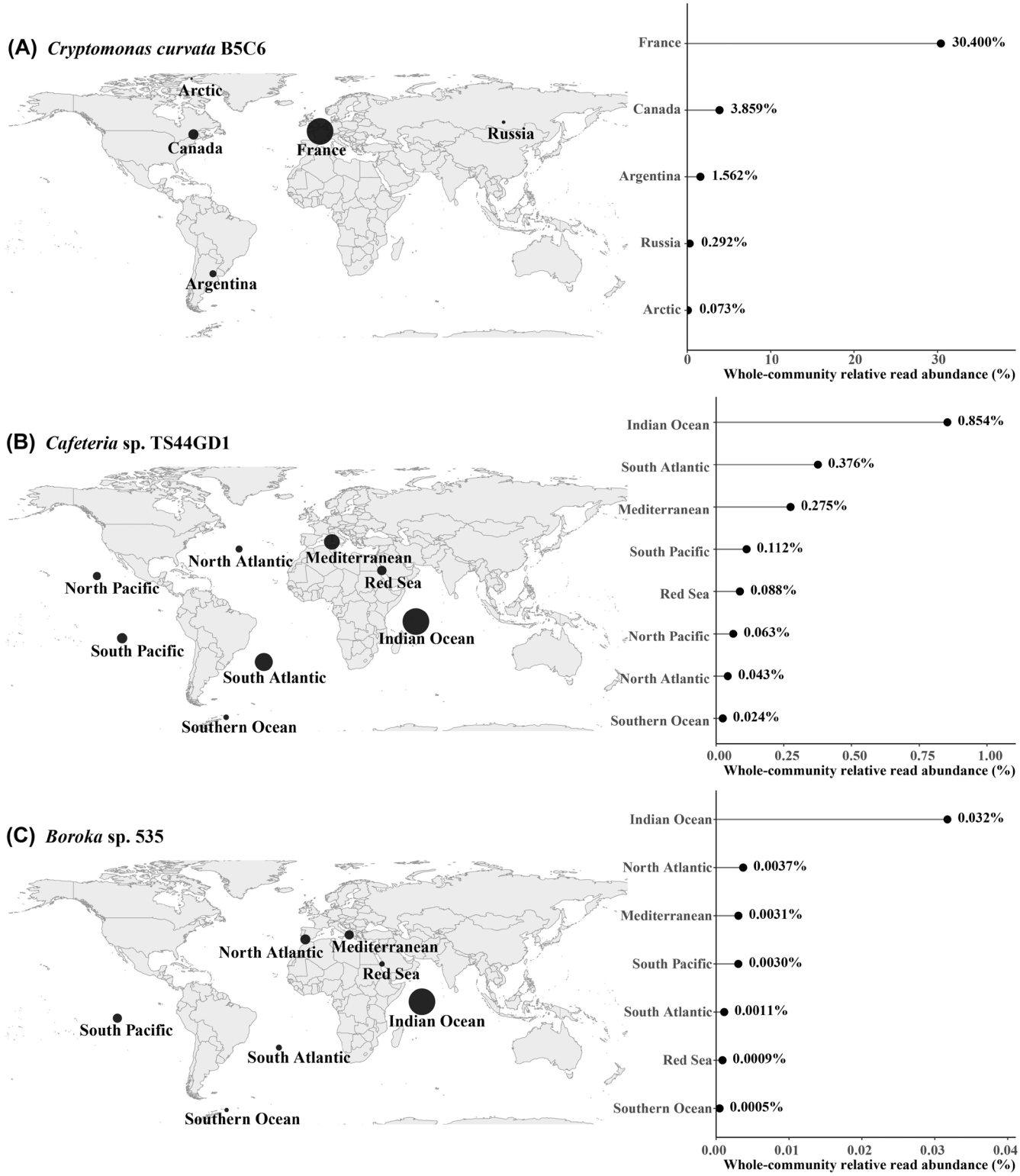

FIGURE 7 | Global relative read abundance of exact-match isolates in MetaPR2 and Tara Oceans datasets. Maps show the geographic occurrence patterns of *Cryptomonas curvata* B5C6, *Cafeteria* sp. TS44GD1 and *Boroka* sp. 535. Circle area represents the mean whole-community relative read abundance of each focal isolate within each freshwater region or ocean basin, calculated as exact-match reads divided by total reads and expressed as a percentage. Circles are scaled independently within each panel, with the largest circle corresponding to the highest regional mean for that focal isolate; circle sizes therefore support within-panel spatial comparison only. Ranked plots show the corresponding regional mean values. These values represent PCR-based relative read abundance rather than absolute cell abundance.

of two fluorescent beads (Figure 4C), corroborating previous reports of phago-mixotrophy in this species (Sanders et al. 1990; Holen 1999). Our field-based grazing-scPCR approach, employing fluorescent beads, microscopy and phylogenetic placement, thus supported the classification of Z91 as a constitutive mixotroph under natural-sample assay conditions.

However, several autofluorescent isolates fell within plastid-lacking lineages, suggesting the possibility of non-constitutive mixotrophic strategies or prey-derived fluorescence. For example, X1ET2 clustered within Katablepharidaceae, a group traditionally regarded as non-photosynthetic, yet displayed intense, uniform red autofluorescence filling the cell (Figures 4D and 5A). This pattern resembles plastid retention in *Hatena arenicola*, a katablepharid known to retain functional plastids from *Nephroselmis* prey (Okamoto and Inouye 2006). Thus, the autofluorescence observed in X1ET2 could reflect plastid retention consistent with a non-constitutive mixotrophic strategy. However, X1ET2 differed markedly in morphology (~8 μm diameter vs. 30–40 × 15–20 μm in *H. arenicola*) and sequence (118 bp divergence; > 6.84%) from described katablepharid sequences. These distinctions suggest that X1ET2 may represent an undescribed katablepharid lineage with potential plastid retention.

Isolate B11C4, which belonged to a Chrysophyceae-related environmental clade sister to Paraphysomonadaceae (Figures 4K and 6A), exhibited both red autofluorescence and bead ingestion, yet its closest described lineage, Paraphysomonadaceae, consists predominantly of colourless heterotrophic chrysophytes (Scoble and Cavalier-Smith 2014). The autofluorescence was less intense and lacked the cell-filling morphology seen in X1ET2, underscoring the need for plastid-targeted molecular assays to confirm whether the signal reflects retained plastids or recently ingested algal prey. A tractable next step to resolve the origin of the red autofluorescence would be plastid 16S rRNA (cp16S) amplicon sequencing from single-cell WGA DNA. In the absence of such data, B11C4 was conservatively classified as heterotrophic. A similar challenge arose with isolates B8GD1, A8ET3, Z2ET1 and X31, which all exhibited red autofluorescence yet clustered phylogenetically within lineages lacking known plastid ancestry. In these cases, the autofluorescence was interpreted as likely prey-derived, leading us to conservatively classify them as heterotrophic nanoflagellates.

In uncultured environmental lineages such as MAST-4 and MAST-12, our isolates X8GD4 and X4ET5 (Figure 6D) exhibited both red autofluorescence and bead ingestion, traits that might initially suggest mixotrophy (Figure 4I,J). However, stramenopiles outside photosynthetic ochrophyte lineages are not expected to possess intrinsic plastids (Derelle et al. 2016), and MAST clades are widely regarded as heterotrophic nanoflagellates (Massana et al. 2014; Lin et al. 2012). Indeed, flow cytometry-sorted MAST-4 cells from the North Pacific display red autofluorescence attributable to chlorophyll-bearing prey retained within food vacuoles (Lin et al. 2012). Importantly, the magnitude of this signal likely scales with the quantity and pigment content of recently ingested prey rather than indicating functional plastid retention. Both isolates deviated from reported MAST morphologies: X8GD4 bears two flagella instead of the single flagellum typical of MAST-4 (Massana et al. 2006), while X4ET5 has a short posterior flagellum (~5–6 μm) rather than the ~30 μm flagellum characteristic of MAST-12 (Kolodziej and Stoeck 2007), raising the possibility that they represent undescribed diversity within these clades. In the absence of plastid-targeted molecular markers and given their phylogenetic placement and consistency with existing heterotrophy data, we conservatively classified X8GD4 and X4ET5 as heterotrophic nanoflagellates.

Collectively, these cases underscore the limitations of using chlorophyll autofluorescence alone as a proxy for mixotrophy. Although our integrative framework—combining field grazing incubations, fluorescence microscopy and phylogenetic context—substantially improves trophic classification, it still cannot directly assess plastid functionality. Red autofluorescence in plastid-lacking lineages may reflect functionally retained plastids, photosymbiotic associations or recently ingested algal prey. To avoid overstatement, we conservatively classified such cases as heterotrophs unless additional evidence supported plastid retention, although this may underestimate the ecological prevalence of non-constitutive mixotrophy. Resolving these ambiguities will require plastid-targeted assays such as plastid 16S rRNA amplicon sequencing to trace plastid origin, coupled with plastid-targeted transcriptomics or gene-expression analyses to test whether retained plastids are functionally active.

4.2.2 | Morphological Plasticity and the Problem of Dormancy

Our findings highlight how morphological plasticity across life stages can complicate the identification of protistan trophic strategies. This was particularly evident in two isolates—B5C6 and FY1ET3—which clustered within the *Cryptomonas curvata* clade (Figure 5), sharing 99.8% sequence similarity. Many *Cryptomonas* species are established mixotrophs, capable of phototrophy and phagotrophy, especially under nutrient-limited conditions, where phagotrophy supplements nitrogen and phosphorus acquisition to sustain growth (Katano et al. 2008; Urabe et al. 2000; Koppelle et al. 2024). Yet phagotrophic ability appears to be lost or not manifested in other cryptophyte species (Isaksson et al. 1999; Blomqvist et al. 2001).

Florenza et al. (2024) employed a FACS-tracer approach, sorting cells with both chlorophyll autofluorescence and prey-tracer signals for single-cell 18S rRNA sequencing and taxonomic assignment using the PR2 database. One OTU (OTU_47) was annotated as *Cryptomonas curvata* based on a 100% sequence identity match to a PR2 reference, yet the authors noted that 'the cryptophyte OTUs are considered to be of uncertain feeding behaviour when comparing tracer performance', reflecting the absence of direct microscopy-verified ingestion evidence. Costa et al. (2022) further showed that phagotrophy in freshwater mixotrophic phytoflagellates, including the congeneric *Cryptomonas marssonii*, can vary with taxon identity, nutrient conditions, light availability and measurement approach. In this study, field grazing-scPCR linked ingestion evidence, plastid autofluorescence, morphology, and phylogenetic identity from the same cell. B5C6 was assigned to *C. curvata*, showed red chlorophyll autofluorescence, and contained an ingested fluorescent bead. This combination provides direct microscopy-verified evidence of bacterial-sized particle ingestion in a plastid-bearing *C. curvata* cell, supporting the classification of B5C6 as a constitutive mixotroph in the sampled natural assemblage.

Moreover, *Cryptomonas curvata*, the type species of its genus, is known to exhibit two morphological forms: cryptomorph and campylomorph (Hoef-Emden and Melkonian 2003). The B5C6 isolate exhibited typical cryptomorph traits (18 × 10 μm) and provided single-cell evidence consistent with constitutive

mixotrophy, ingesting a fluorescent bead while displaying red chlorophyll autofluorescence (Figure 4B). Interestingly, FY1ET3, despite its near-identical 18S rRNA sequence to B5C6, was smaller (8 × 6.7 μm), lacked visible flagella, did not ingest fluorescent beads and nevertheless exhibited red chlorophyll autofluorescence (Figure 4A).

Culture observations on cryptomonads indicate that cells entering encystment under nitrogen limitation can retain plastids and a substantial fraction of cellular chlorophyll *a* (~54.5% of control levels), with a characteristic absorption peak at ~678 nm, consistent with residual pigment capable of autofluorescence rather than complete loss of chlorophyll (Lichtlé 1979). Cyst autofluorescence has also been documented in related green algal resting stages, such as *Chloromonas* cysts (Procházková et al. 2020). These features suggest that FY1ET3 may represent a dormant or cyst-like autofluorescent state retaining plastid pigments (Lichtlé 1979). However, the nutrient conditions measured at the original sampling sites do not support a direct nitrogen-limitation explanation for this state. FY1ET3 was collected from Fangyang Reservoir, where dissolved inorganic nitrogen was high (DIN = 131.9 μmol L^{-1}, dominated by nitrate at 128.3 μmol L^{-1}). Similarly, B5C6 was collected from the Jiulong River Beixi B5 site, where DIN was also high (126.1 μmol L^{-1}). Therefore, although nitrogen limitation can induce encystment in cultured cryptomonads, the cyst-like morphology of FY1ET3 cannot be attributed directly to nitrogen limitation based on the available field data. Other factors, such as light history, prey availability, physiological state, hydrological conditions or microscale environmental variation, may have contributed to the observed non-grazing autofluorescent state. This case illustrates a key limitation in trophic classification: observations made at a single time point may not capture the full ecological potential of a cell, especially if it is in a non-feeding, dormant or otherwise physiologically inactive state. Protists that did not ingest fluorescent beads during the 1-h incubation may still be phagotrophic under different prey conditions, over longer incubation periods, or at different life stages. In addition, the present workflow does not evaluate osmotrophy or uptake of dissolved organic matter. Thus, cells classified as autotrophic or plastid-bearing because they showed red autofluorescence but no LB ingestion should not be interpreted as relying exclusively on photosynthesis. Some taxa may also obtain carbon or nutrients through dissolved organic matter uptake or other non-engulfment feeding pathways that were not measured here.

The interpretation of LB ingestion should also consider the choice of labelled prey. The 0.52 μm latex beads used here provide a standardized, DNA-free tracer for detecting uptake of bacterial-sized particles, but they do not reproduce the motility, surface chemistry, nutritional quality, morphology or chemical cues of natural prey. Because many protists show prey selectivity, positive LB ingestion should be interpreted as direct evidence of bacterial-sized particle uptake during the incubation, but not as evidence of the full prey spectrum of that taxon. Conversely, the absence of LB ingestion does not demonstrate absence of phagotrophic capacity. Trophic-mode assignments based on short-term grazing assays should therefore be interpreted as conservative single-cell snapshots and evaluated together with molecular phylogeny, morphology and ecological context.

4.3 | Local Environmental Context and the Freshwater–Marine Continuum

The environmental data collected at the original sampling sites provide local ecological context for interpreting the single-cell observations. The isolates were obtained across a broad freshwater–marine continuum, from nutrient-rich riverine and reservoir waters to estuarine, coastal and Taiwan Strait sites. This gradient helps contextualize the trophic diversity recovered by field grazing-scPCR: plastid-bearing and mixotrophic isolates were mainly recovered from freshwater or low-salinity sites, whereas many marine and coastal isolates belonged to heterotrophic lineages such as Cafeteriidae, MAST, *Massisteriidae* and Actinomonadaceae-related taxa. These patterns are consistent with the broad environmental heterogeneity represented by the sampling design and highlight that the method can recover trophically diverse protists across contrasting physicochemical settings.

However, the local environmental data were not used to infer formal habitat preferences or environmental drivers for individual taxa. The study was designed around targeted single-cell isolation after grazing incubations rather than replicated, quantitative community sampling. Because the local dataset did not generate replicated taxon-specific abundance measurements across sites, formal correlations between trophic type or taxon occurrence and environmental variables were not performed, as they would be statistically underpowered and potentially misleading. We therefore interpret the site-level environmental measurements as ecological context for the single-cell trophic assignments rather than as evidence for environmental control of trophic-mode distributions.

4.4 | Biogeographic Context of Trophically Characterized Isolates and Their Exact-Match Records

Integrating field grazing-scPCR with MetaPR2 and Tara Oceans exact-match searches provides a bridge between single-cell trophic observations and broader biogeographic context. This approach is useful because each global record is linked to a molecularly defined taxon whose trophic indicators were characterized at the single-cell level. At the same time, interpretation must remain conservative because public amplicon datasets are uneven in geographic and habitat coverage, and amplicon-derived relative read abundance does not represent absolute cell abundance. Moreover, because only three isolates met the prevalence threshold for formal analysis, the resulting patterns should be viewed as taxon-specific exact-match distributions rather than global patterns of trophic modes.

The freshwater constitutive mixotroph *Cryptomonas curvata* B5C6 showed its highest community-wide relative read abundance in France river samples, with lower values in Canadian, Argentinian, Russian, and Arctic samples

(Figure 7A). However, its association with temperature was weak and not statistically significant, and these data therefore do not support a temperature-driven interpretation. Among the marine heterotrophs, *Cafeteria* sp. TS44GD1 had higher relative read abundance in the Southern Hemisphere than Northern Hemisphere oceans (Wilcoxon rank-sum, $p = 0.015$; Figure 7B), with its highest values in the Indian and South Atlantic Oceans. Environmental correlations were generally weak, and no significant association with *Prochlorococcus* biomass was detected (Table S3). This distinction is important because microscopy showed that TS44GD1 can ingest picoalgal prey, but global amplicon data alone cannot demonstrate a direct predator–prey relationship. *Boroka* sp. 535 showed its highest relative read abundance in the Indian Ocean and was negatively associated with dissolved oxygen, with this association remaining significant after false discovery rate correction (Table S3). This pattern suggests that oxygen-related environmental structure may be relevant to the distribution of 535-like bicosoecids, although the underlying mechanism remains unresolved. Public amplicon datasets cannot distinguish whether this association reflects physiological tolerance, prey availability, water-mass structure or sampling heterogeneity.

Environmental associations in this analysis should therefore be interpreted as spatial correlations with available metadata rather than evidence of responses to environmental change. The datasets do not provide temporal gradients, experimental manipulations, or repeated observations of changing conditions. Thus, these correlations are best viewed as hypotheses about environmental structure associated with exact-match lineages, not as evidence that these taxa respond to environmental change.

Together, these results show that linking field grazing-scPCR to global exact-match records can extend single-cell trophic observations into broader biogeographic context. The strength of this approach lies in connecting molecularly identified cells with occurrence and relative read-abundance patterns in public datasets. Its limitation is that amplicon-based signals remain indirect, compositional and unevenly sampled. Therefore, these global analyses are best interpreted as taxon-specific hypotheses about the occurrence of identified protists rather than definitive maps of protistan trophic modes, mixotrophic strategies or functional niches.

4.5 | Future Perspectives

Further development of field grazing-scPCR could improve both throughput and functional validation. Manual microscopic screening of individual droplets is labour-intensive, and integration with microfluidics, imaging flow cytometry or fluorescence-activated cell sorting could accelerate the recovery of candidate cells while preserving microscopy-based single-cell verification. In parallel, combining field grazing-scPCR with plastid-targeted molecular assays, single-cell transcriptomics and related approaches would help distinguish transient prey signals from functional plastid retention and extend trophic characterization beyond particulate ingestion. These developments would further strengthen the ability of single-cell approaches to connect protistan identity, trophic indicators and ecological roles in natural communities.

Author Contributions

P.S. conceived and supervised the study. L.B. and F.S. performed field sampling, grazing assays, single-cell isolation and 18S rRNA sequencing. L.B. and Y.W. curated and analysed the data and prepared figures. P.S. and L.B. wrote the initial draft. P.S., R.L., X.L., D.X. and B.H. provided critical revisions and interpretation. All authors discussed the results and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China and the Natural Science Foundation of Fujian Province (grants 42276095 and 2025J01020, D.X.; grant U25A20793, X.L.; grant 42130401, B.H.; grant 2026J001061, P.S.).

Disclosure

Benefit-Sharing Statement: Sampling and research activities were conducted in accordance with relevant national regulations. No benefit-sharing obligations apply.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The 18S rRNA sequences generated in this study have been deposited in NCBI GenBank under accession numbers PQ770936–PQ770939, PQ770941–PQ770948, PQ770950–PQ770957 and PV931815. All sequences are publicly available. Additional data supporting the findings of this study are provided in the main text or the Supporting Information.

References

Auinger Barbara, M., K. Pfandl, and J. Boenigk. 2008. "Improved Methodology for Identification of Protists and Microalgae From Plankton Samples Preserved in Lugol's Iodine Solution: Combining Microscopic Analysis With Single-Cell PCR." *Applied and Environmental Microbiology* 74: 2505–2510.

Beisner, B. E., H.-P. Grossart, and J. M. Gasol. 2019. "A Guide to Methods for Estimating Phago-Mixotrophy in Nanophytoplankton." *Journal of Plankton Research* 41: 77–89.

Blomqvist, P., M. Jansson, S. Drakare, A. K. Bergström, and L. Brydsten. 2001. "Effects of Additions of DOC on Pelagic Biota in a Clearwater System: Results From a Whole Lake Experiment in Northern Sweden." *Microbial Ecology* 42: 383–394.

Choi, B., M. Son, J. I. Kim, and W. Shin. 2013. "Taxonomy and Phylogeny of the Genus *Cryptomonas* (Cryptophyceae, Cryptophyta) From Korea." *Algae* 28: 307–330.

Costa, M. R. A., H. Sarmento, V. Becker, I. L. Bagatini, and F. Unrein. 2022. "Phytoplankton Phagotrophy Across Nutrients and Light Gradients Using Different Measurement Techniques." *Journal of Plankton Research* 44: 507–520.

Derelle, R., P. López-García, H. Timpano, and D. Moreira. 2016. "A Phylogenomic Framework to Study the Diversity and Evolution of Stramenopiles (=Heterokonts)." *Molecular Biology and Evolution* 33: 2890–2898.

Florenza, J., A.-M. Divne, and S. Bertilsson. 2024. "Fluorescently Labeled Prey Surrogates in Combination With Fluorescence-Activated

Cell Sorting Successfully Discriminate Actively Feeding Mixotrophs in a Lake Water Sample." *Limnology and Oceanography* 69: 1030–1044.

Flynn, K. J., A. Mitra, K. Anestis, et al. 2019. "Mixotrophic Protists and a New Paradigm for Marine Ecology: Where Does Plankton Research Go Now?" *Journal of Plankton Research* 41: 375–391.

Gouy, M., S. Guindon, and O. Gascuel. 2010. "Seaview Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building." *Molecular Biology and Evolution* 27: 221–224.

Hahn Martin, W., and G. Höfle Manfred. 1999. "Flagellate Predation on a Bacterial Model Community: Interplay of Size-Selective Grazing, Specific Bacterial Cell Size, and Bacterial Community Composition." *Applied and Environmental Microbiology* 65: 4863–4872.

Hall, T. A. 1999. "BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT." In *Nucleic Acids Symposium Series*, vol. 41, 95–98. Oxford University Press.

Hoef-Emden, K., and M. Melkonian. 2003. "Revision of the Genus *Cryptomonas* (Cryptophyceae): A Combination of Molecular Phylogeny and Morphology Provides Insights Into a Long-Hidden Dimorphism." *Protist* 154: 371–409.

Holen, D. A. 1999. "Effects of Prey Abundance and Light Intensity on the Mixotrophic Chrysophyte *Poterioochromonas malhamensis* From a Mesotrophic Lake." *Freshwater Biology* 42: 445–455.

Isaksson, A., A.-K. Bergström, P. Blomqvist, and M. Jansson. 1999. "Bacterial Grazing by Phagotrophic Phytoflagellates in a Deep Humic Lake in Northern Sweden." *Journal of Plankton Research* 21: 247–268.

Jeong, H. J., Y. D. Yoo, J. Y. Park, et al. 2005. "Feeding by Phototrophic Red-Tide Dinoflagellates: Five Species Newly Revealed and Six Species Previously Known to Be Mixotrophic." *Aquatic Microbial Ecology* 40: 133–150.

Jürgens, K., and R. Massana. 2008. "Protistan Grazing on Marine Bacterioplankton." In *Microbial Ecology of the Oceans*, 383–441. Wiley-Blackwell.

Katano, T., S.-I. Nakano, H. Ueno, et al. 2008. "Abundance and Composition of the Summer Phytoplankton Community Along a Transect From the Barguzin River to the Central Basin of Lake Baikal." *Limnology* 9: 243–250.

Kolodziej, K., and T. Stoeck. 2007. "Cellular Identification of a Novel Uncultured Marine Stramenopile (MAST-12 Clade) Small-Subunit rRNA Gene Sequence From a Norwegian Estuary by Use of Fluorescence In Situ Hybridization-Scanning Electron Microscopy." *Applied and Environmental Microbiology* 73: 2718–2726.

Koppelle, S., M. Ivanković, M. M. Bengtsson, et al. 2024. "Contrasting Responses of Different Mixotrophic Protists to Light and Nutrient Availability." *Limnology and Oceanography* 69: 1233–1246.

Li, Q., K. Dong, Y. Wang, and K. F. Edwards. 2024. "Relative Importance of Bacterivorous Mixotrophs in an Estuary-Coast Environment." *Limnology and Oceanography Letters* 9: 81–91.

Li, Q., K. F. Edwards, C. R. Schvarcz, and G. F. Steward. 2022. "Broad Phylogenetic and Functional Diversity Among Mixotrophic Consumers of *Prochlorococcus*." *ISME Journal* 16: 1557–1569.

Lichtlé, C. 1979. "Effects of Nitrogen Deficiency and Light of High Intensity on *Cryptomonas rufescens* (Cryptophyceae)." *Protoplasma* 101: 283–299.

Lie, A. A. Y., Z. Liu, R. Terrado, A. O. Tatters, K. B. Heidelberg, and D. A. Caron. 2018. "A Tale of Two Mixotrophic Chrysophytes: Insights Into the Metabolisms of Two *Ochromonas* Species (Chrysophyceae) Through a Comparison of Gene Expression." *PLoS One* 13: e0192439.

Lin, Y.-C., T. Campbell, C.-C. Chung, G.-C. Gong, K.-P. Chiang, and Z. Worden Alexandra. 2012. "Distribution Patterns and Phylogeny of Marine Stramenopiles in the North Pacific Ocean." *Applied and Environmental Microbiology* 78: 3387–3399.

Lynn, D. H., and M. Pinheiro. 2009. "A Survey of Polymerase Chain Reaction (PCR) Amplification Studies of Unicellular Protists Using Single-Cell PCR." *Journal of Eukaryotic Microbiology* 56: 406–412.

Ma, M., Y. Gong, and Q. Hu. 2018. "Identification and Feeding Characteristics of the Mixotrophic Flagellate *Poterioochromonas malhamensis*, a Microalgal Predator Isolated From Outdoor Massive *Chlorella* Culture." *Algal Research* 29: 142–153.

Mansour, J. S., and K. Anestis. 2021. "Eco-Evolutionary Perspectives on Mixoplankton." *Frontiers in Marine Science* 8: 666160.

Marín, I., A. Aguilera, B. Reguera, and J. P. Abad. 2001. "Preparation of DNA Suitable for PCR Amplification From Fresh or Fixed Single Dinoflagellate Cells." *BioTechniques* 30: 88–93.

Massana, R. 2015. "Protistan Diversity in Environmental Eolecular Surveys." In *Marine Protists: Diversity and Dynamics*, 3–21. Springer Japan.

Massana, R., J. del Campo, M. E. Sieracki, S. Audic, and R. Logares. 2014. "Exploring the Uncultured Microeukaryote Majority in the Oceans: Reevaluation of Ribogroups Within Stramenopiles." *ISME Journal* 8: 854–866.

Massana, R., R. Terrado, I. Forn, C. Lovejoy, and C. Pedrós-Alió. 2006. "Distribution and Abundance of Uncultured Heterotrophic Flagellates in the World Oceans." *Environmental Microbiology* 8: 1515–1522.

McManus, G. B., and A. Okubo. 1991. "On the Use of Surrogate Food Particles to Measure Protistan Ingestion." *Limnology and Oceanography* 36: 613–617.

Mitra, A., K. J. Flynn, U. Tillmann, et al. 2016. "Defining Planktonic Protist Functional Groups on Mechanisms for Energy and Nutrient Acquisition: Incorporation of Diverse Mixotrophic Strategies." *Protist* 167: 106–120.

Moon-van der Staay, S. Y., R. De Wachter, and D. Vaulot. 2001. "Oceanic 18S rDNA Sequences From Picoplankton Reveal Unsuspected Eukaryotic Diversity." *Nature* 409: 607–610.

Okamoto, N., and I. Inouye. 2006. "*Hatena arenicola* gen. et sp. nov., a Katablepharid Undergoing Probable Plastid Acquisition." *Protist* 157: 401–419.

Posada, D. 2008. "jModelTest: Phylogenetic Model Averaging." *Molecular Biology and Evolution* 25: 1253–1256.

Procházková, L., D. Remias, W. Bilger, H. Křížková, T. Řezanka, and L. Nedbalová. 2020. "Cysts of the Snow Alga *Chloromonas krienitzii* (Chlorophyceae) Show Increased Tolerance to Ultraviolet Radiation and Elevated Visible Light." *Frontiers in Plant Science* 11: 617250.

Ptacnik, R., U. Sommer, T. Hansen, and V. Martens. 2004. "Effects of Microzooplankton and Mixotrophy in an Experimental Planktonic Food Web." *Limnology and Oceanography* 49: 1435–1445.

Sanders, R. W. 1991. "Mixotrophic Protists in Marine and Freshwater Ecosystems." *Journal of Protozoology* 38: 76–81.

Sanders, R. W., K. G. Porter, and D. A. Caron. 1990. "Relationship Between Phototrophy and Phagotrophy in the Mixotrophic Chrysophyte *Poterioochromonas malhamensis*." *Microbial Ecology* 19: 97–109.

Scoble, J. M., and T. Cavalier-Smith. 2014. "Scale Evolution in Paraphysomonadida (Chrysophyceae): Sequence Phylogeny and Revised Taxonomy of *Paraphysomonas*, New Genus *Clathromonas*, and 25 New Species." *European Journal of Protistology* 50: 551–592.

Stoecker, D. K., P. J. Hansen, D. A. Caron, and A. Mitra. 2017. "Mixotrophy in the Marine Plankton." *Annual Review of Marine Science* 9: 311–335.

Takano, Y., and T. Horiguchi. 2006. "Acquiring Scanning Electron Microscopical, Light Microscopical and Multiple Gene Sequence Data From a Single Dinoflagellate Cell." *Journal of Phycology* 42: 251–256.

Tamura, K., G. Stecher, and S. Kumar. 2021. "MEGA11: Molecular Evolutionary Genetics Analysis Version 11." *Molecular Biology and Evolution* 38: 3022–3027.

Traboni, C., A. Calbet, and E. Saiz. 2021. "Mixotrophy Upgrades Food Quality for Marine Calanoid Copepods." *Limnology and Oceanography* 66: 4125–4139.

Traboni, C., G. F. Cawley, K. E. Selph, M. R. Landry, and M. Décima. 2026. "Exploring the Roles of Trophic Mode and Microbial Prey Size in Grazing Pathways of Tropical Oligotrophic Waters of the Eastern Indian Ocean." *Deep Sea Research Part II: Topical Studies in Oceanography* 225: 105573.

Urabe, J., T. B. Gurung, T. Yoshida, et al. 2000. "Diel Changes in Phagotrophy by *Cryptomonas* in Lake Biwa." *Limnology and Oceanography* 45: 1558–1563.

Ward, B. A., and M. J. Follows. 2016. "Marine Mixotrophy Increases Trophic Transfer Efficiency, Mean Organism Size, and Vertical Carbon Flux." *Proceedings of the National Academy of Sciences of the United States of America* 113: 2958–2963.

Xia, S., Y. Cheng, H. Zhu, G. Liu, and Z. Hu. 2013. "Improved Methodology for Identification of Cryptomonads: Combining Light Microscopy and PCR Amplification." *Journal of Microbiology and Biotechnology* 23: 289–296.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Phylogenetic positions of Cryptophyta and Chlorophyta nanoflagellates based on 18S rRNA genes. Maximum likelihood (ML) phylogenetic tree based on 18S rRNA genes, illustrating the positions of five Cryptophyta nanoflagellates, one Chlorophyta nanoflagellate and related species, constructed using the GTR + I + G model. Node labels indicate ML bootstrap values and Bayesian inference (BI) posterior probabilities. GenBank accession numbers precede species names. Branches are drawn to scale, and only bootstrap values ≥ 50% are shown. **Figure S2:** Phylogenetic relationships of Gyrista (Stramenopiles) nanoflagellates based on 18S rRNA genes. Maximum likelihood (ML) phylogenetic tree based on 18S rRNA genes, illustrating the positions of five Gyrista (Stramenopiles) nanoflagellates and related species, constructed using the GTR + I + G model. Node labels indicate ML bootstrap values and Bayesian inference (BI) posterior probabilities. GenBank accession numbers precede species names. Branches are scaled, with bootstrap values ≥ 50% shown. **Figure S3:** Phylogenetic relationships of Telonemia and Rhizaria nanoflagellates. Maximum likelihood (ML) phylogenetic tree based on 18S rRNA genes, depicting the positions of one Telonemia nanoflagellate, three Rhizaria nanoflagellates and related species, constructed using the GTR + I + G model. Node labels indicate ML bootstrap values and Bayesian inference (BI) posterior probabilities. GenBank accession numbers precede species names. Branches are scaled, with bootstrap values ≥ 50% shown. **Figure S4:** Phylogenetic placement of Bicosoecida nanoflagellates within Stramenopiles. Maximum likelihood (ML) phylogenetic tree based on 18S rRNA genes, illustrating the positions of four Bicosoecida (Bigyra, Stramenopiles) nanoflagellates and related species, constructed using the GTR + I + G model. Node labels indicate ML bootstrap values and Bayesian inference (BI) posterior probabilities. GenBank accession numbers precede species names. Branches are scaled, with bootstrap values ≥ 50% displayed. **Figure S5:** Phylogenetic placement of MAST nanoflagellates within Stramenopiles. Maximum likelihood (ML) phylogenetic tree based on 18S rRNA genes, depicting the positions of two MAST (Stramenopiles) nanoflagellates and related species, constructed using the GTR + I + G model. Node labels indicate ML bootstrap values and Bayesian inference (BI) posterior probabilities. GenBank accession numbers precede species names. Branches are scaled, with bootstrap values ≥ 50% shown. **Table S1:** Sampling details and station-level environmental context. Sampling details and key environmental parameters for waters collected from the Jiulong River, Jiulong River Estuary, Xiamen Bay and Taiwan Strait. Environmental variables are reported for the sampling station from which each single-cell isolate was recovered. **Table S2:** Raw sequence counts for isolates with limited global occurrences in MetaPR2 datasets. Per-sample raw counts of five isolates with 100% identity matches to ≤ 10 samples in the global MetaPR2 amplicon datasets. **Table S3:** Associations between environmental variables and whole-community relative read abundance of focal exact-match isolates. Spearman rank correlations were calculated using relative read abundance defined as exact-match reads divided by total reads in each sample. FDR-adjusted *p* values were used to assess robustness after multiple-testing correction.