



UvA-DARE (Digital Academic Repository)

Expression of genes involved in phagocytosis in uncultured heterotrophic flagellates

Labarre, A.; Obiol, A.; Wilken, S.; Forn, I.; Massana, R.

DOI

[10.1002/Ino.11379](https://doi.org/10.1002/Ino.11379)

Publication date

2020

Document Version

Final published version

Published in

Limnology and Oceanography

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Labarre, A., Obiol, A., Wilken, S., Forn, I., & Massana, R. (2020). Expression of genes involved in phagocytosis in uncultured heterotrophic flagellates. *Limnology and Oceanography*, 65(S1), S149-S160. <https://doi.org/10.1002/Ino.11379>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)



Expression of genes involved in phagocytosis in uncultured heterotrophic flagellates

Aurelie Labarre ^{1*}, Aleix Obiol ¹, Susanne Wilken ^{1,2}, Irene Forn,¹ Ramon Massana ^{1*}

¹Department of Marine Biology and Oceanography, Institut de Ciències del Mar (CSIC), Barcelona, Spain

²Department of Freshwater and Marine Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands

Abstract

Environmental molecular sequencing has revealed an abundance of microorganisms that were previously unknown, mainly because most had not been cultured in the laboratory. Within this novel diversity, there are the uncultured MAST clades (MARine STRamenopiles), which are major components of marine heterotrophic flagellates (HFs) thought to be active bacterial grazers. In this study, we investigated the gene expression of natural HFs in a mixed community where bacterivory was promoted. Using fluorescence in situ hybridization and 18S rDNA derived from metatranscriptomics, we followed the taxonomic dynamics during the incubation, and confirmed the increase in relative abundance of different MAST lineages. We then used single cell genomes of several MAST species to gain an insight into their most expressed genes, with a particular focus on genes related to phagocytosis. The genomes of MAST-4A and MAST-4B were the most represented in the metatranscriptomes, and we identified highly expressed genes of these two species involved in motility and cytoskeleton remodeling, as well as many lysosomal enzymes. Particularly relevant were the cathepsins, which are characteristic digestive enzymes of the phagolysosome and the rhodopsins, perhaps used for vacuole acidification. The combination of single cell genomics and metatranscriptomics gives insights on the phagocytic capacity of uncultured and ecologically relevant HF species.

Heterotrophic flagellates (HFs) are widespread throughout the eukaryotic tree of life and may represent the most ancient eukaryote lifestyle (Cavalier-Smith 2006; Jürgens and Massana 2008; Adl et al. 2019). These colorless flagellated protists are important consumers of primary and secondary production in marine ecosystems (Arndt et al. 2000; Worden et al. 2015), and play a pivotal role in microbial food webs by ensuring the recycling of nutrients. Although they make a significant contribution, it is difficult to assess their diversity because they cannot be easily differentiated by microscopy and most remain uncultured. Moreover, HFs are often ignored in quantitative studies because they are less abundant than photosynthetic protists, and are not well represented in sequence databases, especially of sequenced genomes (del Campo et al.

2014). To study the gene expression and elucidate functional characteristics of these diverse and complex assemblages, metatranscriptomics provides a promising but also challenging opportunity.

Most HFs are considered to be bacterivorous, that is, they consume bacteria by phagocytosis. The engulfment and digestion of a foreign organism as prey, is an important nutritional process in many species of unicellular eukaryotes, and consists of three main steps: (1) motility and prey recognition, (2) internalization, formation, and maturation of the phagosome, and (3) digestion of prey within the phagolysosome (Levin et al. 2016; Uribe-Querol and Rosales 2017). In phagocytosis, prey is internalized by invagination of the plasma membrane to form an intracellular vacuole known as the phagosome (Niedergang and Grinstein 2018). Engulfment is controlled by the actin cytoskeleton and coordinated by phagocytic receptors that activate the GTPases Rac, Rho, and Cdc42 genes (Vieira et al. 2002; Niedergang and Grinstein 2018). The phagosomes then undergo a maturation process, acquire different proteins (like the Rab GTPase) (Rink et al. 2005; Fairn and Grinstein 2012), and become acidified. Mature phagosomes become functional phagolysosomes by fusing with lysosomes, which provide digestive enzymes and further acidify the environment to optimize the performance of these enzymes. Thus, a mature phagolysosome is characterized by the presence of a range of lysosomal acid hydrolases such as proteases, lysozymes,

*Correspondence: labarre@icm.csic.es (A.L.); ramonm@icm.csic.es (R.M.)

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Additional Supporting Information may be found in the online version of this article.

Special Issue: Linking Metagenomics to Aquatic Microbial Ecology and Biogeochemical Cycles. Edited by: Hans-Peter Grossart, Ramon Massana, Katherine McMahon and David A. Walsh.

and lipases (Vieira et al. 2002; Fair and Grinstein 2012). Most of our understanding of phagocytosis at the genomic level is limited to studies on a few species of eukaryotes: several mammals (Boulais et al. 2010), choanoflagellates (Dayel and King 2014), green algae (Burns et al. 2015), and amoebae (Okada et al. 2005). The molecular machinery of phagocytosis in environmentally relevant HF has not been described, which prevents us from quantitatively evaluating particular traits representative for this lifestyle.

A substantial part of marine HF assemblages are MArine STramenopiles (MASTs) (Massana et al. 2014), a set of largely uncultured clades within Stramenopiles, a taxa-rich super-group including autotrophic (Ochrophyta) and heterotrophic (Pseudofungi, Sagenista, and Opalozoa) high-rank lineages. To date, 18 MAST clades have been identified within these three heterotrophic lineages, and each one may have a distinct ecological niche (Massana et al. 2014). Some MAST lineages are widespread and highly abundant in the surface ocean (de Vargas et al. 2015; Mangot et al. 2018). Most MASTs are assumed to be bacterial feeders, as this has been demonstrated in a few of them (e.g., MAST-1 and MAST-4) by microscopic inspection of bacterial preys inside the cells (Massana et al. 2009; Lin et al. 2012). However, it is not clear that all MASTs can perform phagocytosis, nor has it been confirmed which sets of genes and proteins are used for this process. Because of the difficulty of cultivating the most relevant species of marine HF, it is not possible to perform direct physiological and gene expression studies on single species, so there has been little progress in understanding the genetic basis of phagocytosis in these organisms.

In this article, we circumvent the need for culture-based approaches using a combination of molecular tools to study a set of MAST species growing in near-natural conditions. First, we established an unamended incubation of a coastal surface sample in which active HF cells were growing by feeding on bacteria (Massana et al. 2006), and obtained metatranscriptomic data at several time points during the incubation. Second, by using the cell counts and the 18S rDNA signatures of the metatranscriptomes, we analyzed the temporal dynamics of the taxonomic groups succeeding in the incubation, tracing their positive or negative response. Third, using genomic data obtained by single cell genomics (SCG) (Mangot et al. 2017), we recruited transcripts from several MAST species, allowing the analysis of expressed genes likely contributing to the phagocytosis pathway. Thus, the combination of metatranscriptomics with single cell genomes, available for a few dominant species, allowed us for the first time to study the expression profile of uncultured HF in natural assemblages.

Materials and methods

Growth of marine HF in an unamended incubation

Approximately 100 L of surface seawater were sampled from Blanes Bay (41°40'N, 2°48'E) on 4th July 2017, prefiltered by gravity through a nylon mesh of 200 μ m, and transported

to the institute within 2 h. In the lab, 50 L of seawater were gravity-filtered through 3 μ m pore size polycarbonate filters (47 mm diameter) into a polycarbonate bottle (Nalgene). The bottle was incubated for 5 d in the dark at 24°C, the in situ temperature of the sampling site, and sampled twice a day for cell counts and once a day for molecular data. For total cell counts, aliquots were fixed with glutaraldehyde and stained with 4',6-diamidino-2-phenylindole (DAPI). Cell counts of heterotrophic bacteria, *Synechococcus*, and phototrophic and heterotrophic flagellates (2–3 μ m in size) were obtained using epifluorescence microscopy, with excitation by UV radiation (DAPI stained DNA signal) and blue light (to confirm the presence of chlorophyll) (Giner et al. 2016). For cell counts by fluorescence in situ hybridization (FISH) (Amann et al. 1995), aliquots were fixed with formaldehyde and hybridized with oligonucleotide probes specific to MAST-4, MAST-7, *Minorisa minuta*, and Prymnesiophyceae, as described previously (Cabello et al. 2016; Giner et al. 2016). Samples were then examined by epifluorescence microscopy with blue light excitation. For molecular analyses, 2 L of the incubation were filtered through 0.6 μ m pore size polycarbonate filters of 47 mm diameter, which were then flash frozen in liquid nitrogen and stored at –80°C until RNA extraction.

RNA extraction and Illumina sequencing

For RNA extraction, the filters were shattered and vortexed in a tube containing Power Soil beads (Mobio) as described previously (Alonso-Sáez et al. 2018). RNA extraction was performed using the RNeasy Mini Kit (Qiagen). About 60 μ L of the primary RNA extract were processed using Turbo DNase (Ambion, Turbo DNA-free kit) to completely remove residual DNA. The RNA extract was purified by ethanol precipitation, and the pellet resuspended in 40 μ L 10 mmol L⁻¹ TRIS. Metatranscriptomic sequencing was performed using 200–400 ng of total RNA extract. Illumina RNASeq libraries were prepared at CNAG (<https://www.cnag.crg.eu/>) using KAPA Stranded mRNA-Seq Illumina (Roche-KAPA Biosystems). The polyadenylated eukaryotic transcripts were first isolated using poly-T oligonucleotides attached to beads. Then, the mRNA was fragmented using heat and magnesium, converted to complementary DNA (cDNA) by reverse transcription, and sequencing adaptors were ligated to the ends of cDNA molecules. Ligation products were enriched by 15 polymerase chain reaction (PCR) cycles and final libraries were validated with an Agilent 2100 Bioanalyzer. Sequencing was performed at CNAG using an Illumina HiSeq 2500 (TruSeq SBS Kit v4) system, which generated 73 million paired-end reads (2 \times 100 bp) for a final sequencing depth of 15 Gbp per sample. The sequence data are available as raw reads at the NCBI BioSample database (SAMN11783926).

Taxonomic characterization of the microbial community

Focusing on the entire eukaryotic domain, we built a reference database of the hypervariable V4 region of the 18S rRNA gene using sequences from SILVA (Quast et al. 2013), and from data sets of environmental marine protists based on 454 (Massana et al.

2015), and Illumina sequencing (Giner et al. 2019). The database is available at <https://github.com/aleixop/eukaryotesV4>. These references were assigned to several “class level” taxonomic ranks using manual curation. We used local alignment by USEARCH (Edgar 2010) to retrieve reads from the metatranscriptomes related to this reference database (> 97% similarity and > 90% alignment coverage). For each individual read, we selected all hits with the same maximal score and filtered them to obtain the correct taxonomic classification: we assigned a read to a taxonomic class when all top hits belonged to that class; otherwise, the read remained unclassified. The relative abundance of a taxonomic class was obtained by dividing its number of V4 reads by the total number of V4 reads in the sample. To obtain a finer classification of reads within MAST lineages, we prepared a second reference data set using sequences classified within separate MAST lineages (Massana et al. 2014) and applied the same criteria as for read classification.

Assembling the metatranscriptome and read mapping

We performed quality/adaptor trimming of the Illumina HiSeq raw reads using Trimmomatic v0.33 (Bolger et al. 2014) with default settings. We then used SortMeRNA v2.1 (Kopylova et al. 2012) to identify and remove ribosomal RNA reads by comparing all reads against the SILVA SSU and LSU rDNA databases (Quast et al. 2013) and the PR² database (Guillou et al. 2013), using a match score e-value of < 0.01. The resulting set of rRNA free reads allowed us to perform a de novo metatranscriptomic coassembly of the six time points sampled using Trinity (Grabherr et al. 2011) with default parameters. We then used the Burrows-Wheeler aligner BWA software (Li and Durbin 2009) to map the individual cleaned paired-end reads from each sample back to the metatranscriptome assembly, allowing up to two mismatches per read, and then we estimated the expression level of each transcript in TPM units (transcripts per million) using the Salmon software (Patro et al. 2017). We computed TPM values for each isoform defined by Trinity (Grabherr et al. 2011), and for subsequent analyses we summed the signal from all isoforms from the same gene and kept the longest isoform as the representative sequence of each gene.

Extracting MAST transcripts using functionally annotated genes from single amplified genomes (SAGs)

To perform taxonomic binning of the final metatranscriptome, we used the Bowtie2 algorithm with default parameters (Langmead and Salzberg 2012), to map the assembled transcripts to predicted open reading frames of reference genomes from 10 representative MAST species obtained by SCG (Supporting Information Table S1). The single cell amplified genomes of MAST-4A and -4E are published (Mangot et al. 2017), while others have been sequenced, assembled and annotated as part of a separate study (Labarre et al. unpubl.). Gene prediction and functional annotation was performed using Augustus (Stanke et al. 2004) and InterProScan (Jones et al. 2014). Based on

the gene annotation obtained using the single cell genome, we associated each MAST-specific transcript to a gene family that has a given function. We also assigned the proteins encoded by the genomes to the eggNOG database (Huerta-Cepas et al. 2016) that provides general functional overview classified into main categories of biological metabolism. Genome assemblies are available at doi.org/10.6084/m9.figshare.c.4534562.

Results

Growth of marine HFs in an unamended incubation

We examined a mixed community of picoplanktonic microbes ($\leq 3 \mu\text{m}$) growing in a closed system, where higher trophic levels like larger predatory flagellates or ciliates had been filtered out. Using direct epifluorescence microscopy counts, we evaluated the temporal dynamics of several components of this mixed assemblage (Fig. 1a). Bacteria developed progressively during the incubation, obtaining their highest abundance at day 3 ($\sim 2 \times 10^6$ cells mL^{-1}), and appearing to decrease at day 4. During the 5 d of incubation, HFs increased continuously, multiplying by nearly 10-fold (from 10^3 to 10^4 cells mL^{-1}). Photosynthesis was inhibited by incubating the samples in the dark, such that the abundance of photosynthetic flagellates and *Synechococcus* decreased markedly from their initial abundance. As intended, the dark unamended incubation promoted the growth of HFs, which became the most important eukaryotic component of the system.

We constructed metatranscriptomic data for samples taken at six time points during the incubation. Despite the enrichment of mRNA in these transcripts (reverse transcription was based on the mRNA poly-A tail), we obtained many 18S rRNA reads (5–12% per sample), which we used to assess the taxonomic composition and dynamics of the assemblage by classifying individual reads to broad taxonomic classes (Fig. 1b). Initially, the samples were dominated by groups composed mainly of photosynthetic species, principally Prymnesiophyceae, Dictyochophyceae, and Dinoflagellates, and these decreased markedly over time. In contrast, groups with heterotrophic taxa became more abundant, namely Choanomonada and several MAST lineages (MAST-1, -7, and -3). The increase in abundance of Chlorarachniophyta, which is a generally photosynthetic class, was due to the presence of its only known heterotrophic member, *M. minuta*.

We used taxonomy profiles based on 18S rRNA data to characterize the dynamics of the 40 most abundant classes, which collectively accounted for 99% of the community (Fig. 2a). The metatranscriptomic data set showed a large reduction in relative abundance of virtually all autotrophic groups, including Archaeplastida, Cryptomonadales, Prymnesiophyceae, Diatomea, Pelagophyceae, and the three MOCH lineages. In contrast, heterotrophic groups showed a more varied response to the incubation, with some decreasing in abundance (Picozoa, MAST-11, MALV-II), others remaining stable (Telonema, Katablepharids, Cercozoa), et al increasing (Choanomonada, most MAST

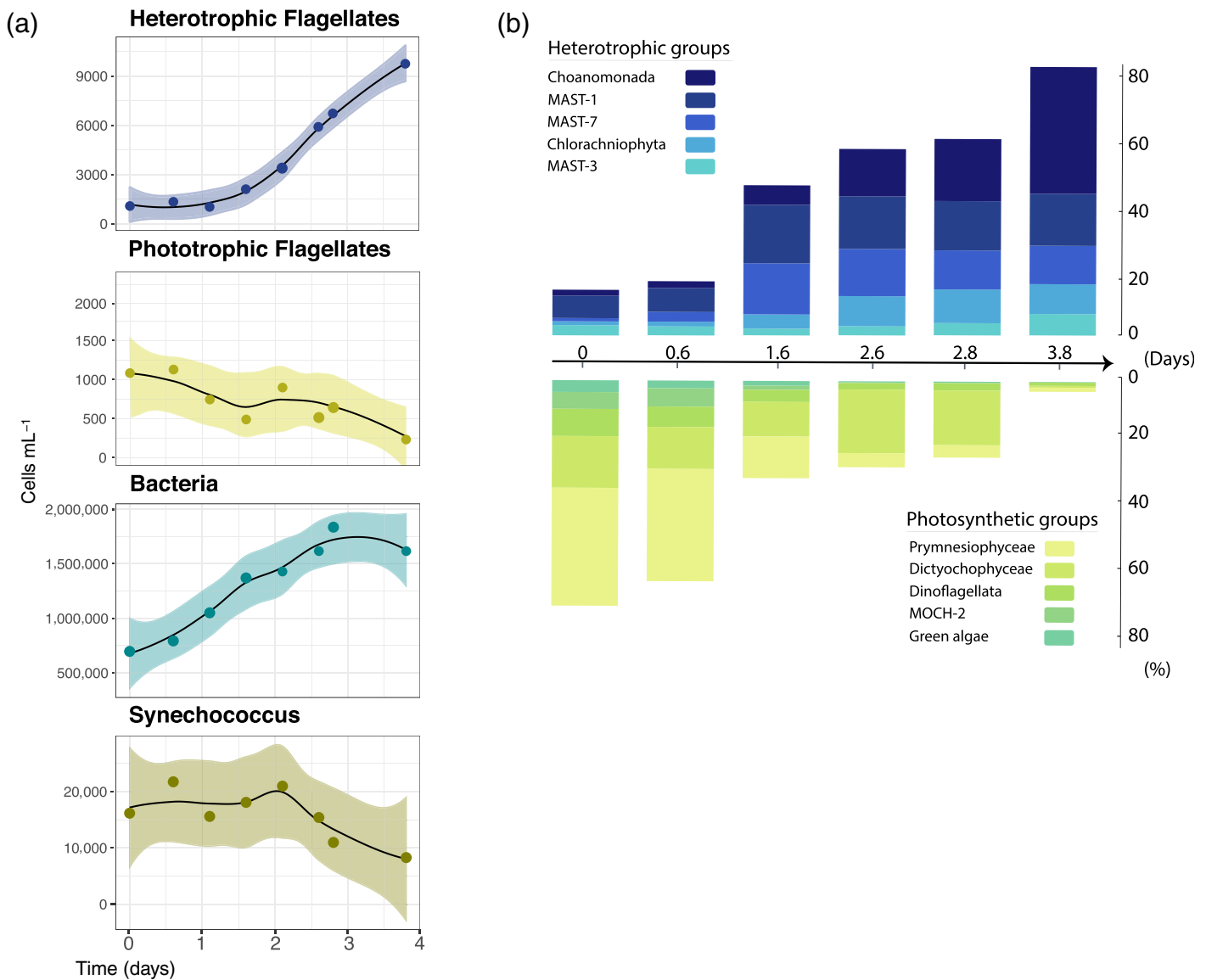


Fig. 1. Temporal dynamics of the mixed microbial assemblage. **(a)** Changes in cell abundance of HFs (2–3 μm), photosynthetic flagellates (2–3 μm), heterotrophic bacteria, and *Synechococcus* obtained by epifluorescence microscopy after DAPI staining. Dots represent actual cell counts while shades show the overall trend using the estimate of the conditional mean function with the R package ggplot using a linear model (lm). **(b)** Relative abundance of the most important taxonomic groups during the enrichment as measured by their contribution to V4 18S rDNA reads in the metatranscriptomes. The heterotrophic groups are shown in the upper graph while the photosynthetic groups are in the lower graph.

lineages, Bicosoecida and Labyrinthulomycetes). As expected, groups containing both autotrophic and heterotrophic species did not show a clear trend. To support these observations, we targeted a few groups using FISH (Fig. 2b), and observed a marked decrease in prymnesiophytes (from 400 to 3 cells mL⁻¹), and an increase in *M. minuta*, MAST-7, and MAST-4 (increasing from 65 to 1300 cells mL⁻¹, 11 to 300 cells mL⁻¹, and 47 to 200 cells mL⁻¹, respectively). Our metatranscriptomics data revealed a very complex assemblage containing a large diversity of taxonomic groups and confirmed the growth of HFs and the decrease of photosynthetic groups.

Characterization of HFs using metatranscriptomics

After excluding rRNA reads, we built a de novo metatranscriptome assembly by merging the paired-end RNA-Seq data for the six time points sampled during incubation. We obtained 3,812,907 transcripts with an average length of 414 bp. We computed the TPM value for each transcript in every sample, and summed the values for all isoforms from the same gene, resulting in 3,338,309 transcript values. The resulting gene expression profile of the mixed community is very heterogeneous, as it represents hundreds of species from distant phylogenetic groups, each one expressing a different set of specific genes.

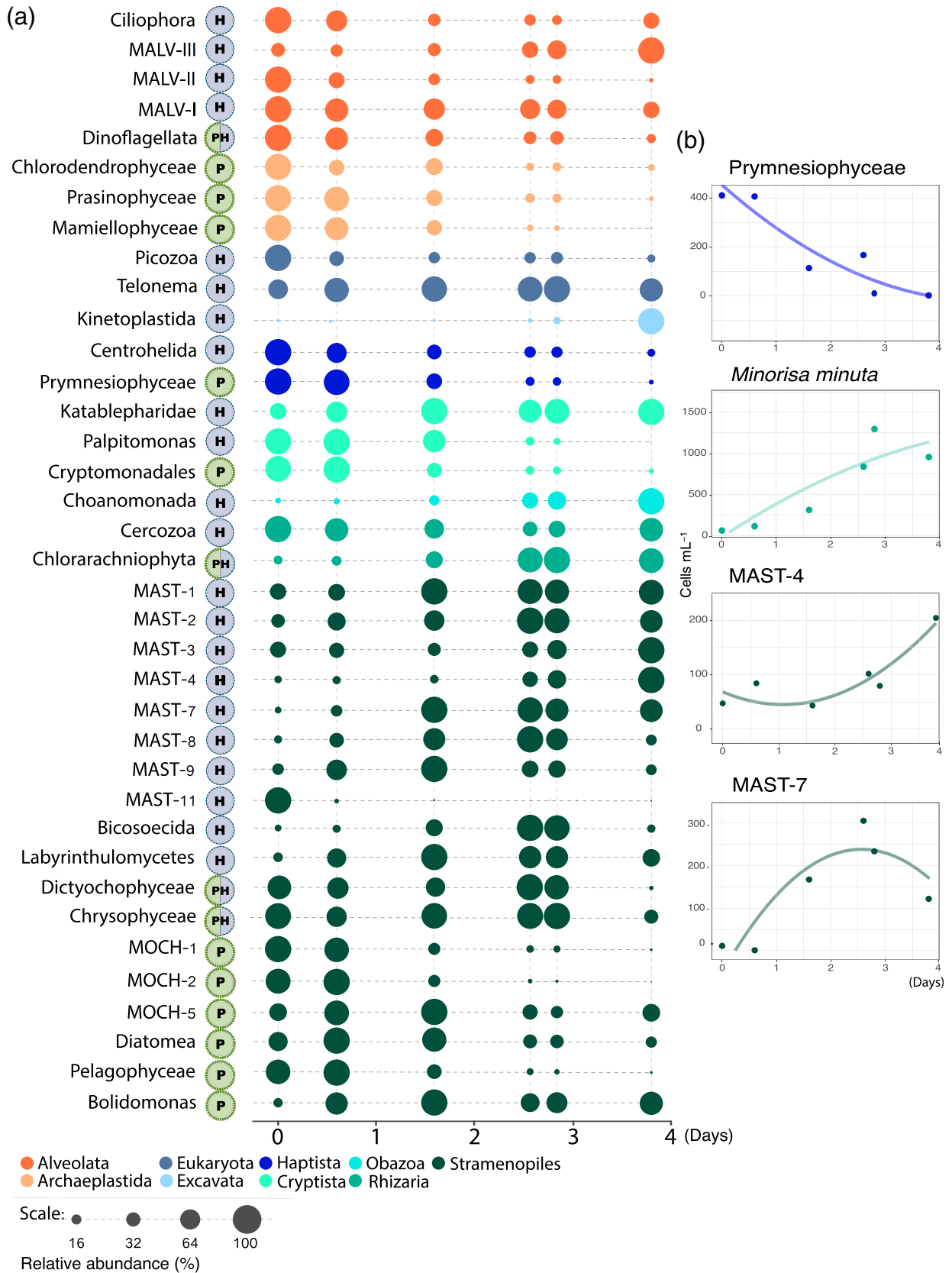


Fig. 2. Legend on next page.

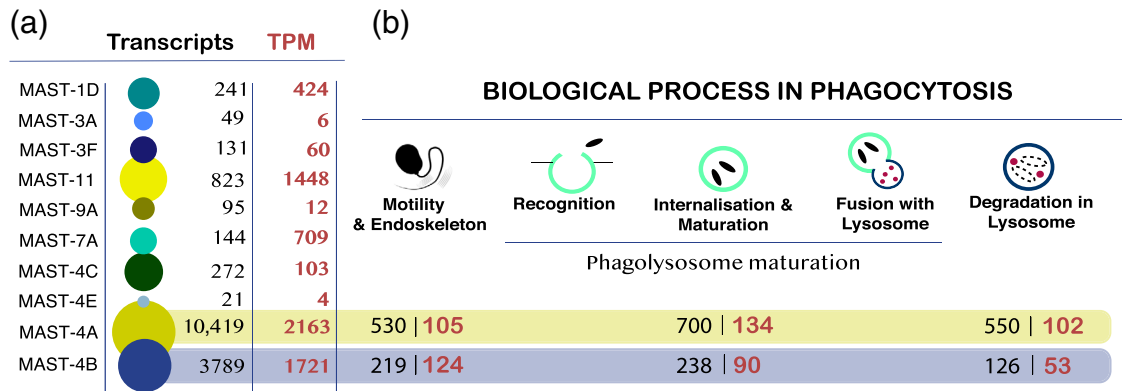


Fig. 3. Mapping of the metatranscriptome toward MAST genomes. **(a)** Number of transcripts recruited using SCG of 10 MAST species and their expression level (averaged TPM values over the six time-points) shown in red. Circles illustrate the relative count of the transcripts in every species. **(b)** A schematic representation of the main steps of phagocytosis, together with the number of transcripts (and their averaged TPM values) detected in each of them within the MAST-4A and -B species.

To assess gene expression and their putative function in individual species, we mapped the assembled transcripts to the predicted open reading frames of the SAGs of 10 MAST species. The two species that retrieved most transcripts were MAST-4A and MAST-4B (10,419 and 3789, respectively), and these had the highest expression level (Fig. 3a). Noticeably, the two closely related MAST-4C and -E retrieved very few transcripts and very little expression signal, as well as the remaining MAST genomes tested, with perhaps the exception of MAST-11 for which the expression signal was still considerable. Overall, we obtained a low recovery of transcripts mapping the SAGs, with about 0.22% of the entire metatranscriptome assigned to MAST-4A and 0.17% to MAST-4B. This low retrieval was generally consistent with the temporal changes in MAST relative abundances identified by the finest taxonomic classification of 18S rDNA reads (Supporting Information Fig. S1), which suggested little dominance of the 10 species represented by our SAGs. Thus, while MAST-7 increased in relative abundance (Figs. 1b, 2), only 144 transcripts mapped to the MAST-7A SAG (Fig. 3a), as the subclades growing in the incubation were MAST-7B, -D, and -E (Supporting Information Fig. S1). Similarly, few transcripts were fetched using the MAST-3A and -3F SAGs (the subclades growing during incubation were -3E, -I, and -J) and MAST-9A SAG (the subclade growing was -9D). The signal represented by MAST-4 was moderate and increased during the incubation both by FISH cell counts and by 18S rDNA mapping mostly to clades-A and -B. Finally, results were not consistent in two cases: MAST-11 was barely detectable by the 18S rRNA mapping but retrieved 823 transcripts with a moderate expression

profile, while MAST-1D retrieved very few transcripts but abundant 18S rDNA reads.

Gene expression of two uncultured HFIs

The transcripts associated to MAST-4A (10,419 transcripts) and to MAST-4B (3789) were annotated using the eggNOG database, which categorizes their functions into roles in metabolism, cellular and signaling processes, and information storage (Supporting Information Table S2). The large majority of transcripts were identified through the eggNOG database (only 1652 transcripts in MAST-4A and 707 in MAST-4B did not have a match) but many of them affiliated to uncharacterized proteins (3679 and 1544 transcripts). Most general functions were well represented by the expressed genes, being posttranslational modification and cytoskeleton the most represented categories.

The next step was to target expressed genes potentially involved in the phagocytosis pathway, focusing on the three main steps: prey recognition and motility, phagosome maturation, and degradation in lysosome (Table 1). Many of these genes were associated with motility and/or cytoskeleton functions (475 transcripts in MAST-4A and 219 in MAST-4B). Genes coding for actin and tubulin, which are structural components of the flagella and the cytoskeleton, were abundant and highly expressed in both species. We also detected genes associated with microtubule formation, such as myosin, dynein, and kinesin, as well as several flagella-associated proteins and intraflagellar transporters that are essential for flagellar growth. In both species, we found highly expressed genes likely involved in the phagolysosome maturation step (614 transcripts in MAST-4A and 238 in MAST-4B), although

Fig. 2. Temporal changes in the relative abundance of all taxonomic groups. **(a)** The relative abundance of a given group is calculated by dividing the number of 18S rDNA reads from the group by the total number of 18S rDNA reads in the sample; these values are normalized within each group (being the highest abundance scaled to 100). The main trophic mode of the groups is shown as cartoons next to their names: P (phototrophs), H (heterotrophs), PH (groups containing phototrophs and heterotrophs). **(b)** Actual cell abundances estimated by FISH for MAST-4, MAST-7, *M. minuta*, and Pymnesiophyceae.

Table 1. Genes involved in the phagocytosis pathway identified in MAST-4A and -4B. The number of transcripts and their averaged TPM values within each category defined within the phagocytosis pathway are displayed within the three identified steps.

	MAST-4A		MAST-4B	
	Transcripts	TPM	Transcripts	TPM
Motility and cytoskeleton				
Actin family	27	11.2	17	5.7
CAP Gly-rich domain	21	4.2	4	1.4
Cilia and flagella associated protein	16	2.5	14	4.4
Clathrin	30	4.3	6	2.5
Dynein heavy chain	221	38.7	41	13.7
Formin	12	1.7	1	0.2
Intraflagellar transport protein	3	0.4	1	1.0
Kinesin motor domain	106	17.8	65	17.4
Myosin head, motor domain	66	9.8	57	77.5
Tubulin family	28	15.0	13	5.4
Phagosome maturation				
ABC transporters	205	39.1	119	35.7
Calcium transporting	81	12.1	18	24.6
Ions transporting	182	31.8	40	15.3
Phosphatidylinositol 3/4-kinase	65	10.0	15	4.3
Proton-transporting V-type ATPase	12	3.2	8	2.3
Pyrophosphate-energized proton pump	18	6.3	0	—
Rhodopsin	8	12.1	0	—
Vacuolar protein sorting	57	8.5	18	3.9
Sec23/Sec24, helical domain	21	3.2	3	0.7
Small GTPase superfamily	38	6.2	16	14.0
SNARE coiled-coil	2	0.1	1	0.4
Syntaxin	2	0.2	0	—
WASH complex	9	1.2	0	—
Phagolysosome				
Glycoside hydrolase family	188	34.7	56	48.6
Peptidase C - S - M	362	66.1	70	29.4

some have more diverse cellular functions. These include phosphatidylinositol 3/4 kinase and several proton pumps that are potentially responsible for phagosome acidification (Table 1). For example, a rhodopsin was the third most highly expressed gene in MAST-4A, along with the vacuolar pyrophosphatase and GTPase Arf type in MAST-4B, which in turn are well known as regulators of vesicular traffic and actin remodeling. Although less expressed, the presence of the SNARE complex (Soluble N-ethylmaleimide-sensitive factor activating protein receptor) is very informative, as it is responsible for intracellular membrane fusion and trafficking steps interacting with vacuolar protein sorting. Finally, we also observed the expression of a set of genes encoding digestive enzymes (437 transcripts in MAST-4A and 126 in MAST-4B), such as the glycoside hydrolase family and peptidases, especially the lysosomal proteases cysteine cathepsins. These genes were among the most expressed in MAST-4B and were also important in MAST-4A (Table 1).

Finally, as we followed the gene expression on several time points along the incubation, we explored the possibility of changes in the expression profiles of the two species. The abundance of transcripts was relatively constant over time for MAST-4A, and increased markedly for MAST-4B (Fig. 4a). This was consistent with the 18S rDNA signal (Supporting Information Fig. S1), and suggested that the proportion of MAST-4A cells remained stable in the community, while that of MAST-4B cells increased. Then we focused on the subset of the 100 most highly expressed genes, with the expression signal normalized within each species to account for the changes of species abundance along the incubation (Fig. 4b). No clear changes seemed to occur during time, with the few temporal clusters appearing unrelated to any gene function in particular. Therefore, it seemed that the two species were transcriptionally at a similar stage throughout the incubation. Interestingly, in this subset of most highly expressed genes, we detected several that were related to the three main steps of phagocytosis described, and more specifically the

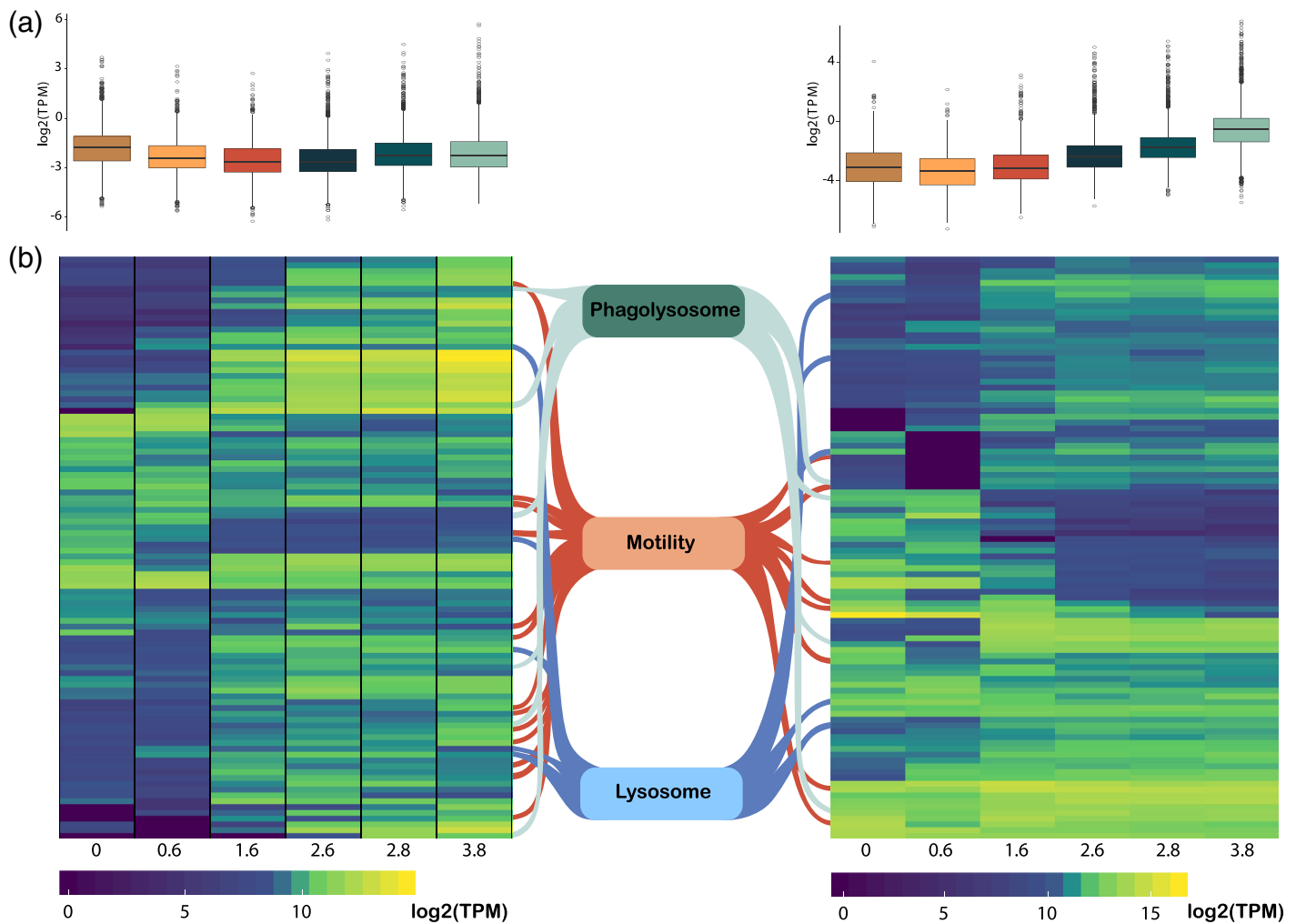


Fig. 4. Gene expression of MAST-4A (left) and MAST-4B (right) during the incubation. **(a)** Box plots showing the expression values of all transcripts as TPM data. **(b)** Expression dynamics of the 100 most expressed transcripts of MAST-4A and MAST-4B. Within each species, the TPMs for every genes are added up and this sum is scaled to 1 million, in order to get its expression profile as if it was the only member of the community. Genes involved in phagocytosis are marked and classified into three broad categories. Heatmaps were done with the R package *superheat* and the hierarchical clustering method.

presence of digestive enzymes as the cathepsins, confirming the relevance of this process for the flagellate ecophysiology.

Discussion

Using a mixed natural community sampled from surface water in Blanes Bay, we successfully promoted the growth of bacterivorous HF s in order to study their gene expression profiles. Based on both cell counts and 18S rDNA analyses, we observed a pronounced change in community composition during the incubation, most notably an increase in the abundance of HF s and a decrease in photosynthetic taxa. The switch from autotrophy to heterotrophy was expected as the experiment was conducted in the dark. This suppressed the growth of photosynthetic taxa, which in turn could be grazed by the heterotrophs. The development of the HF s was

further supported by the growth of bacteria that proliferated early in the incubation as well as the absence of higher trophic grazers, following similar dynamics to those previously observed in other incubations with different initial communities (Massana et al. 2006; Weber et al. 2012). Larger taxa (i.e., dinoflagellates and ciliates) were nevertheless detected at the beginning in our incubation, and this signal most likely derived from broken cells, explaining the modest 18S rDNA signal detected and their decrease along time (Fig. 2). By integrating taxonomic classification of transcripts with microscopical cell counts, we detected a large diversity of efficient grazers in a predator dynamics scenario.

While 18S rDNA sequences can be classified phylogenetically and are useful for ecological studies on species distribution, many lineages such as the marine stramenopiles are understudied because they remain uncultured. In particular, they generally

lack genomic information because most genomic research is biased toward a few cultured model species (Pawlowski et al. 2012; del Campo et al. 2014). This gap can now be filled by SCG, which allows retrieving the genomes of individually sorted cells without the need for culturing (Mangot et al. 2017). When combined with metatranscriptomics, which is a reliable approach for investigating metabolically active cells, SCG allowed the distinction between unicellular individual lineages present in the incubation that were so far barely recognized as most genomics focuses on a few model eukaryotes (del Campo et al. 2014). While SCG is essential to assign gene functions to uncultured species, it is also known to have some limitations, such as lack of coverage in some genomic regions (Rinke et al. 2014), and the presence of contaminant sequences that can compromise the quality of the final assembly. Thus, the absence of a particular gene in the metatranscriptomics analysis could be because it was not amplified by the multiple displacement amplification, was not assembled, or was not annotated in the final SAG used as reference. At any rate, in this study, SCG has allowed unprecedented insights into MASTs gene expression profiles, which could have not been revealed with the metatranscriptomics on its own.

Phagocytosis is a specific form of endocytosis (uptake of extracellular material) that involves engulfing large particles (Niedergang and Grinstein 2018) and that originated billions of years ago (Yutin et al. 2009). It is a complex process found in diverse eukaryotes, and that involves a variety of functional genes that are often not unique to the phagocytosis pathway but are shared with other processes (e.g., actin filament, ABC transporters). This mode of feeding has been studied in depth in macrophages in the mammalian immune system, and also in some unicellular microbial eukaryotes (Dayel and King 2014). Gotthardt et al. (2002), performed a targeted study of the proteins involved in phagocytosis (and their corresponding genes) using direct proteomic analyses of extracted phagosomes. Recently, a complementary effort using comparative genomics has attempted to identify the set of genes that are unique and representative of different trophic modes, including phagotrophy (Burns et al. 2015, 2018). These insights showed a complex process controlled by regulatory mechanisms involving numerous genes, revealing the potential for molecular detection of specific markers of phagocytosis. Linking gene expression and ecosystem function is feasible in marine bacteria, where marker genes for given biogeochemical functions in the oceans have been identified (Ferrera et al. 2015), allowing targeted studies of, for example, ammonia oxidation or phosphorous uptake (Imhoff 2016). Similar efforts toward the identification of genes indicative of a trophic strategy in microbial eukaryotes have recently been published (Alexander et al. 2015; Liu et al. 2016), including the search of specific traits driving heterotrophy (Beisser et al. 2017; Hu et al. 2018). Phagocytosis offers a unique case where genes can be related to microbial food webs. Therefore, to analyze genes that participate in phagocytosis in uncultured HFs, we used a metatranscriptomics approach on a community enrichment combined with a selection of specific

transcripts provided by single cell genomes. Our method allowed to identify genes that control multiple aspects of the phagocytosis. We have detected many known characterized genes, and even though we sampled at different times during the enrichment, we did not see noticeable transcriptional changes. Therefore, in our targeted species, the phagocytosis machinery showed similar functional signal along the sampled time.

For two uncultured HF species, MAST-4A and MAST-4B, we detected genes involved in several of the main steps of the phagocytosis pathway as have been described mostly for mammalian macrophages. Phagocytosis is an actin-dependent process that is required to initiate prey capture. We identified genes that control the nucleation of new actin filaments, namely the assembly factors of the Arp2/3 complex (May and Machesky 2001; Lai et al. 2008). Actin polymerization is also promoted by the small GTPases of the Rho family, of which we detected Cdc42 here; this is generally followed by Rac1 and Rac2 activation, but these were not observed. These GTPases interact with proteins of the WASP and Scar/WAVE family that stimulate the Arp2/3 complex (Castellano et al. 2001), and both were actively expressed in the community assemblage. Involved in phagosome maturation, the Rab-family GTPases (Vieira et al. 2003; Fairn and Grinstein 2012) participate in the formation of the phagolysosome, and here we failed to detect important genes involved in this step, such as Rab5 and Rab7. These gene absences could be due to genome incompleteness of the SAGs. In addition to GTPases, we found a putative phosphatidylinositol 3-kinase (PI3K), which is necessary for successful phagolysosome formation (Vieira et al. 2003). The food vacuole seems to be acidified by highly expressed proton and cation pumps that are organized by vacuolar ATPases (Kissing et al. 2015). In addition, we identified high expression of a rhodopsin gene in MAST-4A, and we hypothesize that the corresponding protein could act as a light-driven proton pump contributing to phagosome acidification (Slamovits et al. 2011; Kandori 2015). Microbial rhodopsins were initially found in Archaeobacteria and are now known to be widely dispersed light-driven ion transporters across all domains of life (Beja et al. 2000; Finkel et al. 2013). Finally, the strongly acidified phagolysosome becomes rich in hydrolytic enzymes that promote the degradation of the ingested microbial prey, and here we found several highly expressed digestive enzymes such as cathepsins and glycoside hydrolases. Our results indicate a set of genes related to phagocytosis that were highly expressed in environmentally relevant bacterivorous uncultured HFs. The genes defined here are often not exclusive to phagocytosis, but represent a continuum of proteins involved in different types of fusion, vesicle transport, and digestive processes. In addition to markers of phagosome acidification (V-ATPase and rhodopsins), the digestive enzymes cathepsins would be ideal candidates for detecting phagocytosis in assemblages of marine HFs. Future work will need to assess their suitability to target this trophic mode in natural communities.

Historically, the study of microorganisms, including ecogenomics and gene expression profiling, has focused on single species in pure culture. Metatranscriptomics allows us to circumvent the culture-dependent analysis of many microbial species and to perform functional studies in complex communities. In our analysis, SCG has facilitated the targeting of the transcriptional profile of uncultured HFs. This approach is a new opportunity to examine the heterogeneity of microbial communities, recover their true diversity, and better understand specific biological processes performed by particular species.

References

- Adl, S. M., and others. 2019. Revisions to the classification, nomenclature, and diversity of eukaryotes. *J. Eukaryot. Microbiol.* **66**: 4–119. doi:10.1111/jeu.12691
- Alexander, H., M. Rouco, S. T. Haley, S. T. Wilson, D. M. Karl, and S. T. Dyrman. 2015. Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. *Proc. Natl. Acad. Sci. USA* **112**: E5972–E5979. doi:10.1073/pnas.1518165112
- Alonso-Sáez, L., X. A. G. Morán, and M. R. Clokie. 2018. Low activity of lytic pelagiphages in coastal marine waters. *ISME J.* **12**: 2100–2102. doi:10.1038/s41396-018-0185-y
- Amann, R. L., W. Ludwig, and K.-H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**: 143–169.
- Arndt, H., D. Dietrich, B. Auer, E.-J. Clevén, T. Gräfenhan, M. Weitere, and A. P. Mylnikov. 2000. Functional diversity of heterotrophic flagellates in aquatic ecosystems, p. 240–268. *In* The flagellates.
- Beisser, D., and others. 2017. Comprehensive transcriptome analysis provides new insights into nutritional strategies and phylogenetic relationships of chrysophytes. *PeerJ* **5**: e2832. doi:10.7717/peerj.2832
- Beja, O., and others. 2000. Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science* **289**: 1902–1906. doi:10.1126/science.289.5486.1902
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120. doi:10.1093/bioinformatics/btu170
- Boulais, J., and others. 2010. Molecular characterization of the evolution of phagosomes. *Mol. Syst. Biol.* **6**: 423. doi:10.1038/msb.2010.80
- Burns, J. A., A. Paasch, A. Narechania, and E. Kim. 2015. Comparative genomics of a bacterivorous green alga reveals evolutionary causalities and consequences of phago-mixotrophic mode of nutrition. *Genome Biol. Evol.* **7**: 3047–3061. doi:10.1093/gbe/evv144
- Burns, J. A., A. A. Pittis, and E. Kim. 2018. Gene-based predictive models of trophic modes suggest Asgard archaea are not phagocytotic. *Nat. Ecol. Evol.* **2**: 697–704. doi:10.1038/s41559-018-0477-7
- Cabello, A. M., M. Latasa, I. Forn, X. A. G. Morán, and R. Massana. 2016. Vertical distribution of major photosynthetic picoeukaryotic groups in stratified marine waters: Photo-picoeukaryotes community structure in the DCM. *Environ. Microbiol.* **18**: 1578–1590. doi:10.1111/1462-2920.13285
- Castellano, F., C. Le Clainche, D. Patin, M.-F. Carlier, and P. Chavrier. 2001. A WASp-VASP complex regulates actin polymerization at the plasma membrane. *EMBO J.* **20**: 5603–5614. doi:10.1093/emboj/20.20.5603
- Cavalier-Smith, T. 2006. Cell evolution and Earth history: Stasis and revolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **361**: 969–910. doi:10.1098/rstb.2006.1842
- Dayel, M. J., and N. King. 2014. Prey capture and phagocytosis in the choanoflagellate *Salpingoeca rosetta*. *PLoS One* **9**: e95577. doi:10.1371/journal.pone.0095577
- de Vargas, C., and others. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**: 1261605. doi:10.1126/science.1261605
- del Campo, J., M. E. Sieracki, R. Molestina, P. Keeling, R. Massana, and I. Ruiz-Trillo. 2014. The others: Our biased perspective of eukaryotic genomes. *Trends Ecol. Evol.* **29**: 252–259. doi:10.1016/j.tree.2014.03.006
- Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461. doi:10.1093/bioinformatics/btq461
- Fairn, G. D., and S. Grinstein. 2012. How nascent phagosomes mature to become phagolysosomes. *Trends Immunol.* **33**: 397–405. doi:10.1016/j.it.2012.03.003
- Ferrera, I., M. Sebastian, S. G. Acinas, and J. M. Gasol. 2015. Prokaryotic functional gene diversity in the sunlit ocean: Stumbling in the dark. *Curr. Opin. Microbiol.* **25**: 33–39. doi:10.1016/j.mib.2015.03.007
- Finkel, O. M., O. Béjà, and S. Belkin. 2013. Global abundance of microbial rhodopsins. *ISME J.* **7**: 448–451. doi:10.1038/ismej.2012.112
- Giner, C. R., I. Forn, S. Romac, R. Logares, C. de Vargas, and R. Massana. 2016. Environmental sequencing provides reasonable estimates of the relative abundance of specific picoeukaryotes. *Appl. Environ. Microbiol.* **82**: 4757–4766. doi:10.1128/AEM.00560-16
- Giner, C. R., and others. 2019. Quantifying long-term recurrence in planktonic microbial eukaryotes. *Mol. Ecol.* **28**: 923–935. doi:10.1111/mec.14929
- Gotthardt, D., H. J. Warnatz, O. Henschel, F. Brückert, M. Schleicher, and T. Soldati. 2002. High-resolution dissection of phagosome maturation reveals distinct membrane trafficking phases. *Mol. Biol. Cell* **13**: 3508–3520. doi:10.1091/mbc.e02-04-0206
- Grabherr, M. G., and others. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**: 644–652. doi:10.1038/nbt.1883
- Guillou, L., and others. 2013. The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote small

- sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**: D597–D604. doi:[10.1093/nar/gks1160](https://doi.org/10.1093/nar/gks1160)
- Hu, S. K., Z. Liu, H. Alexander, V. Campbell, P. E. Connell, S. T. Dyhrman, K. B. Heidelberg, and D. A. Caron. 2018. Shifting metabolic priorities among key protistan taxa within and below the euphotic zone: Depth-related protistan metatranscriptomes. *Environ. Microbiol.* **20**: 2865–2879. doi:[10.1111/1462-2920.14259](https://doi.org/10.1111/1462-2920.14259)
- Huerta-Cepas, J., and others. 2016. eggNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**: D286–D293. doi:[10.1093/nar/gkv1248](https://doi.org/10.1093/nar/gkv1248)
- Imhoff, J. 2016. New dimensions in microbial ecology-functional genes in studies to unravel the biodiversity and role of functional microbial groups in the environment. *Microorganisms* **4**: 19. doi:[10.3390/microorganisms4020019](https://doi.org/10.3390/microorganisms4020019)
- Jones, P., and others. 2014. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240. doi:[10.1093/bioinformatics/btu031](https://doi.org/10.1093/bioinformatics/btu031)
- Jürgens, K., and R. Massana. 2008. Protistan grazing on marine bacterioplankton, p. 383–441. *In* D. L. Kirchman [ed.], *Microbial ecology of the oceans*. John Wiley & Sons.
- Kandori, H. 2015. Ion-pumping microbial rhodopsins. *Front. Mol. Biosci.* **2**: 52. doi:[10.3389/fmolb.2015.00052](https://doi.org/10.3389/fmolb.2015.00052)
- Kissing, S., and others. 2015. Vacuolar ATPase in phagosome-lysosome fusion. *J. Biol. Chem.* **290**: 14166–14180. doi:[10.1074/jbc.M114.628891](https://doi.org/10.1074/jbc.M114.628891)
- Kopylova, E., L. Noé, and H. Touzet. 2012. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217. doi:[10.1093/bioinformatics/bts611](https://doi.org/10.1093/bioinformatics/bts611)
- Lai, F. P., and others. 2008. Arp2/3 complex interactions and actin network turnover in lamellipodia. *EMBO J.* **27**: 982–992. doi:[10.1038/emboj.2008.34](https://doi.org/10.1038/emboj.2008.34)
- Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**: 357–359. doi:[10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923)
- Levin, R., S. Grinstein, and J. Canton. 2016. The life cycle of phagosomes: Formation, maturation, and resolution. *Immunol. Rev.* **273**: 156–179. doi:<https://doi.org/10.1111/imr.12439>
- Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760. doi:[10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324)
- Lin, Y.-C., T. Campbell, C.-C. Chung, G.-C. Gong, K.-P. Chiang, and A. Z. Worden. 2012. Distribution patterns and phylogeny of marine stramenopiles in the North Pacific Ocean. *Appl. Environ. Microbiol.* **78**: 3387–3399. doi:[10.1128/AEM.06952-11](https://doi.org/10.1128/AEM.06952-11)
- Liu, Z., V. Campbell, K. B. Heidelberg, and D. A. Caron. 2016. Gene expression characterizes different nutritional strategies among three mixotrophic protists. *FEMS Microbiol. Ecol.* **92**: fiw106. doi:[10.1093/femsec/fiw106](https://doi.org/10.1093/femsec/fiw106)
- Mangot, J.-F., and others. 2017. Accessing the genomic information of unculturable oceanic picoeukaryotes by combining multiple single cells. *Sci. Rep.* **7**: 41498. doi:[10.1038/srep41498](https://doi.org/10.1038/srep41498)
- Mangot, J.-F., I. Forn, A. Obiol, and R. Massana. 2018. Constant abundances of ubiquitous uncultured protists in the open sea assessed by automated microscopy: Automatic microscopic counts of small protists. *Environ. Microbiol.* **20**: 3876–3889. doi:[10.1111/1462-2920.14408](https://doi.org/10.1111/1462-2920.14408)
- Massana, R., L. Guillou, R. Terrado, I. Forn, and C. Pedrós-Alió. 2006. Growth of uncultured heterotrophic flagellates in unamended seawater incubations. *Aquat. Microb. Ecol.* **45**: 171–180. doi:[10.3354/ame045171](https://doi.org/10.3354/ame045171)
- Massana, R., F. Unrein, R. Rodríguez-Martínez, I. Forn, T. Lefort, J. Pinhassi, and F. Not. 2009. Grazing rates and functional diversity of uncultured heterotrophic flagellates. *ISME J.* **3**: 588–596. doi:[10.1038/ismej.2008.130](https://doi.org/10.1038/ismej.2008.130)
- 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 J.* **8**: 854–866. doi:[10.1038/ismej.2013.204](https://doi.org/10.1038/ismej.2013.204)
- Massana, R., and others. 2015. Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing: Protist diversity in European coastal areas. *Environ. Microbiol.* **17**: 4035–4049. doi:[10.1111/1462-2920.12955](https://doi.org/10.1111/1462-2920.12955)
- May, R. C., and L. M. Machesky. 2001. Phagocytosis and the actin cytoskeleton. *J. Cell Sci.* **114**: 1061–1077.
- Niedergang, F., and S. Grinstein. 2018. How to build a phagosome: New concepts for an old process. *Curr. Opin. Cell Biol.* **50**: 57–63. doi:[10.1016/j.ceb.2018.01.009](https://doi.org/10.1016/j.ceb.2018.01.009)
- Okada, M., and others. 2005. Proteomic analysis of phagocytosis in the enteric protozoan parasite *Entamoeba histolytica*. *Eukaryot. Cell* **4**: 827–831. doi:[10.1128/EC.4.4.827-831.2005](https://doi.org/10.1128/EC.4.4.827-831.2005)
- Patro, R., G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**: 417–419. doi:[10.1038/nmeth.4197](https://doi.org/10.1038/nmeth.4197)
- Pawlowski, J., and others. 2012. CBOL protist working group: Barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS Biol.* **10**: e1001419. doi:[10.1371/journal.pbio.1001419](https://doi.org/10.1371/journal.pbio.1001419)
- Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**: D590–D596. doi:[10.1093/nar/gks1219](https://doi.org/10.1093/nar/gks1219)
- Rink, J., E. Ghigo, Y. Kalaidzidis, and M. Zerial. 2005. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* **122**: 735–749. doi:[10.1016/j.cell.2005.06.043](https://doi.org/10.1016/j.cell.2005.06.043)
- Rinke, C., and others. 2014. Obtaining genomes from uncultivated environmental microorganisms using FACS-

- based single-cell genomics. *Nat. Protoc.* **9**: 1038–1048. doi:[10.1038/nprot.2014.067](https://doi.org/10.1038/nprot.2014.067)
- Slamovits, C. H., N. Okamoto, L. Burri, E. R. James, and P. J. Keeling. 2011. A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nat. Commun.* **2**: 183. doi:[10.1038/ncomms1188](https://doi.org/10.1038/ncomms1188)
- Stanke, M., R. Steinkamp, S. Waack, and B. Morgenstern. 2004. AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Res.* **32**: W309–W312. doi:[10.1093/nar/gkh379](https://doi.org/10.1093/nar/gkh379)
- Uribe-Querol, E., and C. Rosales. 2017. Control of phagocytosis by microbial pathogens. *Front. Immunol.* **8**: 1368. doi:[10.3389/fimmu.2017.01368](https://doi.org/10.3389/fimmu.2017.01368)
- Vieira, O. V., R. J. Botelho, and S. Grinstein. 2002. Phagosome maturation: Aging gracefully. *Biochem. J.* **366**: 689–704. doi:[10.1042/bj20020691](https://doi.org/10.1042/bj20020691)
- Vieira, O. V., and others. 2003. Modulation of Rab5 and Rab7 recruitment to phagosomes by phosphatidylinositol 3-kinase. *Mol. Cell. Biol.* **23**: 2501–2514. doi:[10.1128/MCB.23.7.2501-2514.2003](https://doi.org/10.1128/MCB.23.7.2501-2514.2003)
- Weber, F., J. del Campo, C. Wylezich, R. Massana, and K. Jürgens. 2012. Unveiling trophic functions of uncultured protist taxa by incubation experiments in the brackish Baltic Sea. *PLoS One* **7**: e41970. doi:[10.1371/journal.pone.0041970](https://doi.org/10.1371/journal.pone.0041970)
- Worden, A. Z., M. J. Follows, S. J. Giovannoni, S. Wilken, A. E. Zimmerman, and P. J. Keeling. 2015. Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science* **347**: 1257594–1257594. doi:[10.1126/science.1257594](https://doi.org/10.1126/science.1257594)
- Yutin, N., M. Y. Wolf, Y. I. Wolf, and E. V. Koonin. 2009. The origins of phagocytosis and eukaryogenesis. *Biol. Direct* **4**: 9. doi:[10.1186/1745-6150-4-9](https://doi.org/10.1186/1745-6150-4-9)

Acknowledgments

This project was supported by the EU project SINGEK (H2020-MSCA-ITN-2015-675752) and the Spanish project ALLFLAGS (CTM2016-75083-R, MIN-ECO). S.W. was supported by a Postdoctoral Fellowship from the European Union's Seventh Framework Program under Marie Skłodowska-Curie grant agreement 6261882. We thank Ramiro Logares and Fran Latorre for useful discussions and sharing data.

Author Contribution Statement:

A.L. and R.M. led the writing of the manuscript. A.L., S.W., A.O., and I.F. designed the experiment, developed the sample methodology, and produced the metatranscriptomic data set. Data interpretation and analysis was by A.L., S.W., A.O., and R.M. All authors reviewed the manuscript.

Conflict of Interest

None declared.

Submitted 20 February 2019

Revised 17 June 2019

Accepted 24 September 2019

Associate editor: Katherine McMahon