Chytrid infecting the bloom-forming marine diatom *Skeletonema* sp.: morphology, phylogeny and distribution of a novel species within the Rhizophydiales

Andrea Garvetto¹, Yacine Badis¹, Marie-Mathilde Perrineau¹, Cecilia Rad-Menéndez², Eileen Bresnan³, Claire M. M. Gachon³

¹ Scottish Association for Marine Science, Scottish Marine Institute, Oban, UK.
² Culture Collection of Algae and Protozoa, Scottish Association for Marine Science, Scottish Marine Institute, Oban, UK
³ Marine Scotland Science, Marine Laboratory, Aberdeen, UK

Abstract

Chytrids have long been recognised as important parasites of microalgae in freshwater systems, able to shape the dynamics of blooms, the genetic structure of their host and phytoplankton succession. In the sea however, where the presence of these organisms is erratic and ephemeral, studies concerning chytrids are sparse and confined to metabarcoding surveys or microscopy observations. Despite the scarcity of data, chytrid epidemics are supposed to play an important role in marine biogeochemical cycles, being one of the drivers of phytoplankton dynamics. Here we combine microscopy observations and *in silico* mining of a single-cell whole genome to molecularly and morphologically characterise a novel chytrid parasite of the dominant diatom genus *Skeletonema*. Morphological observations highlight features of the thallus and ascertain the parasitic nature of the interaction whilst the genetic markers obtained allows for a phylogenetic reconstruction, placing the new species in the order Rhizophydiales. Thanks to the molecular data obtained we are also able to provide a first investigation of the global distribution of this organism by screening the Ocean Sampling Day (OSD) dataset, highlighting a northern transatlantic dissemination.
Keywords

Marine microbial diversity, marine fungi, plankton parasites, single-cell analysis, NGS

1. Introduction

Zoosporic true fungi belonging to the early diverged phylum Chytridiomycota (Hibbett et al., 2007) are widespread microorganisms known to dominate the fungal community in freshwater ecosystems (Monchý et al., 2011; Jobard et al., 2012) where they play an important role as parasites and saprobes (Gleason et al., 2008; Frenken et al., 2017). Parasitic associations of chytrids with microalgae have long been known in lentic systems (Canter and Lund, 1953; Sparrow, 1960), where chytrid propagules act as an important career-carrier of fixed carbon from inedible algae to grazing zooplankters by means of the mycoloop (Kagami et al., 2014); and are suspected to be one of the main drivers maintaining the genetic diversity of their algal hosts through highly host-specific associations (Ibelings et al., 2004; De Bruin et al., 2008; Gsell et al., 2013). Despite the prevalence of studies targeting lakes, chytrids have been recorded as a major component of the microbial community in various ecosystems by means of molecular surveys, especially in extreme habitats such as alpine snow (Naff et al., 2013), sea-ice (Hassett et al., 2017), high elevation soils (Freeman et al., 2009) and marine hydrothermal vents (Le Calvez et al., 2009). Whenever these habitats allow for the survival of photosynthetic organisms, the co-occurrence (Freeman et al., 2009; Naff et al., 2013; Brown et al., 2015) or parasitic association (Hassett and Gradinger, 2016; Hassett et al., 2017) between phototrophs and chytrids has been reported. In the sea, recent metabarcoding surveys unveiled an unexpectedly large and undescribed biodiversity within the Chytridiomycota (Comeau et al., 2016; Hassett and Gradinger, 2016; Hassett et al., 2017). Despite not being the yearlong dominant fungal taxon (Picard, 2017), a high abundance of chytrids in the plankton was often recorded in short time windows matching coastal blooms of phytoplankters such as diatoms (Taylor and Cunliffe, 2016). In less productive marine ecosystems, such as open ocean oligotrophic gyres, chytrids have been found parasitising dominant photosynthetic picoeukaryotes within the families
Chrysophyceae and Prymnesiophyceae (Lepère et al., 2016). Even if culturing remains a paramount step to fully appreciate and understand the biology of these organisms, establishing and maintaining a dual culture of the parasite and its host (i.e. pathosystem) in the laboratory is challenging (Rad-Menéndez et al., 2018). The isolation of chytrids is further complicated by the erratic and ephemeral presence of these “dazzling and elusive creatures” in the field (Sparrow, 1958).

Therefore, most of the information on chytrid ecological role and diversity is underpinned by either sparse microscopic observations that are informative regarding the parasitic interaction but often lack taxonomic resolution; or by broad molecular surveys; which provide sequencing data and an overview of chytrid phylogenetic diversity, but can only be suggestive in terms of ecological function. For these reasons, a big part of chytrid diversity is assigned to the undescribed, uncultivated and functionally uncharacterised dark matter fungi (Grossart et al., 2016), hindering the appreciation of their role in ecosystem dynamics and global biogeochemical cycles.

Diatoms within the genus *Skeletonema* are among the main contributors to planktonic primary productivity in coastal temperate areas worldwide, often representing the dominant diatom during the spring bloom (Borkman and Smayda, 2009; Canesi and Rynearson, 2016). This diatom genus often constitutes the basis of food webs (Bergkvist et al., 2012) and can account for up to 99% of the microphytoplankton community during seasonal blooms (Canesi and Rynearson, 2016). Whilst the contribution and effects of zooplankton grazing has been extensively studied (Bergkvist et al., 2012; Amato et al., 2018), only one study reports chytrids infecting *Skeletonema*, addressing their role in bloom dynamics and phytoplankton species succession off the Chilean coast (Gutiérrez et al., 2016).

Here, using a simple method that combines microscopy and single-cell whole genome amplification, we identify a novel chytrid species infecting *Skeletonema* sp. from the west coast of Scotland. We provide a morphological description, molecular identification and a phylogenetic placement for this chytrid within the Rhizophydiales. Finally, thanks to the obtained molecular data, we use the OSD worldwide metabarcoding database (Kopf et al., 2015) to investigate the global distribution of this new species.
2. Methods

2.1 Isolation of the biological material

Plankton samples were collected on the 25th of April 2016 off the pontoon of the Scottish Association for Marine Science (56°27'12.6"N 5°26'12.2"W, Oban, UK) and screened in plastic Petri dishes (145 mm diameter, Greiner Bio One) with a Zeiss IDO3 microscope (32X magnification). Skeletonema chains carrying chytrid sporangia were isolated by mouth pipetting. Infected chains were transferred into droplets (~ 5 μL) of fresh f/2+Si medium into 50 mm glass bottom Petri dishes for microscopy observations. Cultivation of the chytrid parasitoid was attempted by inoculating baiting the chytrid parasitoid from environmental samples with Skeletonema spp. strains CCAP 1077/1B, CCAP 1077/1C, CCAP 1077/3, CCAP 1077/4, CCAP 1077/5 and CCAP 1077/7 (maintained in f/2+Si) with infected environmental samples. None resulted in a successful infection.

2.2 Microscopy

Subsamples containing infected cells were treated with the stains Calcofluor White (CW, Fluka), Wheat Germ Agglutinin Fluorescein isothiocyanate (WGA-FITC, Vector Labs) and Nile Red (NR, Sigma Aldrich). Subsamples of 10 mL were incubated in 100X diluted stain stock solutions (final concentrations: 5 μg mL⁻¹ for WGA-FITC or CW; 1 μg mL⁻¹ for NR) for 15-30 minutes at room temperature in the dark. After incubation samples were either screened directly in glass-bottom Petri dishes or transferred onto microscopy slides. Observations were carried out at a 100X magnification with a Zeiss LSM 510 microscope (equipped with an AxioCam HRC) in bright field and UV-epifluorescence, using an excitation wavelength of either ~ 488 nm (WGA-FITC and NR) or ~ 360 nm (CW). Unstained infected chains were observed in bright field on glass bottom Petri dishes before being isolated singularly for downstream molecular analysis.

2.3 DNA amplification and sequencing

Single infected chains were transferred by mouth pipetting into 1.5 mL sterile Eppendorf tubes. Care was taken to sterilise the glass tip of the mouth pipette with boiling freshwater before each
Isolation. Isolated single chains were frozen at -20 °C for at least 12 hours before being thawed and subjected to multiple displacement whole genome amplification (MDA; Lasken, 2007) with the Qiagen REPLI-g® Single Cell Kit, as per the manufacturer’s instruction. Amplified material was stored at -80 °C and aliquots were diluted 100 times in autoclaved Milli-Q water before downstream PCR amplifications. PCRs were performed in a total reaction volume of 50 μL containing 3 μL of diluted genomic DNA, 1 μL (0.2 ng/μl) of each primer, 25 μL of master mix solution (Taq PCR Mastermix, Qiagen) and 20 μL of autoclaved Milli-Q water. Partial 18S, 28S sequences and the complete ITS region of the chytrid parasite were obtained with the following primers: EaF3 (Pröschold et al., 2001), MH2 (Vandenkoornhuyse and Leyval, 1998), NS1, NS4, ITS4 (White et al., 1990) and D3Ca (Lenaers et al., 1989). The PCR parameters for each primer couple are specified in Table 1.

Note that the decreasing annealing temperature gradient for the primer couple MH2-NS4 aimed at increasing the specificity of the reaction in the first steps whilst favouring the DNA yield towards the end of the cycle. PCR products were separated by agar gel electrophoresis (35 min at 100 V on 1.5% agarose in TBE buffer) and multiple bands isolated by band excision with GeneJET Gel Extraction (Thermo Scientific) and purified by DNA clean-up Micro Kit (Thermo Scientific), followed by Sanger sequencing (GATC Biotech, Konstanz, Germany). Partial rDNA operon sequences were identified for two single cell isolates, SkChyt5 and SkChyt3. The longest sequence was obtained for isolate SkChyt5, for which 1464 bp within the 18S were obtained by merging the products of primers EaF3, NS1, MH2 (forward) and NS4 (reverse). The products of the reverse primers ITS4 (757 bp in the ITS region) and D3Ca (573 bp in the 28S), despite a good per base quality, were too short to be assembled with the rest of the rDNA operon. The partial 18S of a second single cell isolate, i.e. SkChyt3, was amplified via PCR using the primer couple MH2-NS4 and sequenced following the above described procedure, resulting in a 933 bp marker gene matching with 100% identity SkChyt5 18S. Therefore a 20 μL aliquot of the original MDA product from isolate SkChyt5 was diluted 1:4 (V:V) in autoclaved Milli-Q water and sent for genome library construction (NEBNext Ultra DNA
Library Prep Kit) and sequenced via Illumina HiSeq 3000 150 bp paired end sequencing (University of Leeds Sequencing Facility, Leeds, United Kingdom).

2.4 Genome querying for genetic barcodes

FastQC (Andrews, 2010) was used to assess the quality of SkChyt5 HiSeq data (200,984,460 paired end reads) before trimming with Trimmomatic (Bolger et al., 2014, parameters detailed in Supplementary Materials 1). The resulting SkChyt5 196,165,072 trimmed reads were queried via BLAST (Altschul et al., 1990) using in-house scripts (Supplementary Materials 1) designed to extract subsets of reads as per their identity to a given query sequence. The best BLAST hit of SkChyt5 18S amongst type material (i.e. Kappamyces laurelensis AFTOL-ID 690) was used to generate a concatenated rDNA query sequence (GenBank accession numbers: DQ536478, DQ536494 and DQ273824) to extract chytrid rDNA reads from the single cell genomic reads pool. The obtained subset of reads (1453 merged reads) was de novo assembled into 13 contigs using Geneious 6.1.8 (Kearse et al., 2012), of which 6 were retained on the basis of BLAST annotation (i.e. fungal rDNA) and coverage (i.e. >10 reads; with 5 contigs out of 6 having a coverage >100 reads). The 6 selected contigs were further merged to the SANGER sequences into two supercontigs, spanning the nearly complete rDNA, but not overlapping.

In a complementary attempt to retrieve the complete rDNA operon, we generated a de novo assembly of SkChyt5 in CLC Genomics Work Bench 8.5.1 (https://www.qiagenbioinformatics.com/) and queried the resulting 299,415 contigs using the 18S sequence obtained by PCR on SkChyt5. This approach resulted in the selection of 2 de novo assembled contigs of 1289 bp and 7688 bp in length, both annotated as chytrids rDNA as per their best BLAST hit amongst cultured strains, i.e. Kappamyces laurelensis isolate AFTOL-ID 690 18S (DQ536478.1) and Boothiomycetes macroporosum CBS 122107 28S (NG_027566.1) respectively. These two contigs were, when overlapping, 100 % identical to all the sequences previously obtained by PCR and in silico. The 18S sequence obtained by PCR filled the gap between the two contigs, allowing for the assembly of the complete rDNA operon,
which was retained for following phylogenetic analysis. The different methods used to reconstruct the full rDNA operon are summarized in Supplementary Figure 1. The rDNA operon was manually annotated in Geneious 6.1.8, by BLAST against GeneBank. Introns were analysed with the same software in order to test whether they contained coding DNA sequences (CDS) and further annotated via UniProt. The entirety of the rDNA operon for the isolate SkChyt5 was deposited in GenBank under the accession number MH643793.

2.5 Phylogenetic reconstructions

A custom dataset of the 18S, 5.8S and 28S rDNA sequences of chytrids was built paying particular attention to sample diversity within the Rhizophydiales (Letcher et al., 2006), since preliminary analysis pointed towards an inclusion of isolate SkChyt5 into this order. Sequences for each gene were aligned separately using the MAFFT (Katoh et al., 2002) algorithm implemented in Geneious 6.1.8 and manually curated to remove the non-informative introns ITS1 and ITS2. Substitution models were assessed for each gene separately in IQ-TREE 1.5.5 (Nguyen et al., 2015) through ModelFinder (Kalyaanamoorthy et al., 2017), resulting in TIM2+R5 (18S and 28S) and TPM2u+R4 (5.8S). A manual concatenation of the three alignments was then analysed in IQ-TREE 1.5.5 with a partitioned model (Chernomor et al., 2016), allowing each gene to evolve at its own pace (–spp option). Phylogenetic reconstructions were computed with a Maximum Likelihood method using a 1000 replicates Ultrafast bootstrap approximation (UFBoot) test of phylogeny (Minh et al., 2013).

Aiming at including a better representation of uncultured environmental diversity, a phylogenetic reconstruction based on Rhizophydiales 18S was computed with the same above described procedure under the GTR+R4 model of molecular evolution. Finally, for consistency with previous work on Alphamycetaceae and Rhizophydiales phylogeny reconstructions (Letcher et al., 2012; Seto and Degawa, 2018) we used a concatenation of 5.8S (substitution model SYM+I+G4) and 28S (substitution model TIM3+G4) to further confirm the phylogenetic position of our isolate.

2.6 Assessment of global distribution
The V4 hypervariable region of the 18S rDNA of isolate SkChy5 was used to query the raw metabarcode datasets of the Ocean Sampling Day through the Sequence Read Archive (SRA) (PRJEB8682, Kopf et al., 2015) using an in-house script referred to as MOULINETTE which implements EDirect and SRA Toolkit utilities (Badis et al., subm.). Briefly, MOULINETTE extracts from SRA deposited metabarcoding datasets all reads matching a given query sequence above given identity, coverage and e-value thresholds. Here a rather stringent approach was used, i.e. reads were only retained when at least 99% identical over 80% of their length to the query sequence (refer to Suppl. Material 1 for details on the parameters used). Paired reads were filtered (expected error over 1.0) and aligned to the V4 hypervariable region of the 18S of isolate SkChyt5 in Geneious 6.1.8, resulting in two clusters divided by a shared synapomorphy (OTU1 and OTU2). Identity matrices were generated to investigate the percentage of identity between the reads and SkChyt5-V4-18S-rDNA with the same software. Sampling stations were deemed positive for the presence of the investigated organism, and therefore retained, when >10 merged paired reads were detected. For these stations GPS coordinates were extracted and maps generated using the open source online platform GPSVisualizer (http://www.gpsvisualizer.com/, Adam Schneider).

3. Results

3.1 Morphology

Infected Skeletonema cells were distinguished from healthy ones (Fig. 1A) thanks to the presence of nearly spherical sporangia growing epiphytically on diatom colonies (Fig. 1B). Infected cells showed signs of plasmolysis, chloroplast collapse (Fig. 1B) and decreased (if any) chlorophyll fluorescence in UV light (Fig. 1C). The majority of the observed sporangia had a diameter of ~ 5 μm, but reached 7.5 μm across in fully mature specimens. Sporangia were usually attached to the girdle band region of the infected cell. In a few instances however, sporangia could also be observed growing on the valvar surface within the cage of intercalary fultoportula processes connecting two adjacent Skeletonema cells (Fig. 1B, inlet). Examination of CW-stained sporangia under UV light in epifluorescence highlighted the presence of a cell wall with distinct zones of wall thickening in
mature sporangia (Fig. 1C arrowheads). A reduced, unbranched rhizoidal system (Fig. 1C arrow) developed from the germ tube observed in encysted zoospores (Fig. 1C, arrow in the inlet). The presence of chitin was detected in sporangia cell walls, as shown by the positive reaction to the N-acetylglucosamine specific stain WGA-FITC (Fig. 1D, arrowhead). Sporangia contain refractive lipid globules variable in size and number (Fig. 1E), whilst encysted zoospores contain a single lipid droplet (Fig. 1C, arrowhead). Zoospores were sub-spherical, ~2.5 μm across, bearing a single whiplash flagellum (Fig. 1F-G, arrows) and a single refractive lipid droplet (Fig. 1F, arrowhead). Zoospores release or zoospore locomotion could never be directly observed. In one occasion a putative resting spore, with a thick cell wall and bearing a single eccentric refractive globule was observed (Fig. 1H). CW/NR staining highlighted the presence of a zoospore encysted on this structure, possibly suggestive of sexual reproduction (Fig. 1I).

3.3 Phylogenetic placement of isolate SkChyt5

Our maximum likelihood phylogenetic reconstruction on the concatenated rRNA encoding genes 18S, 5.8S and 28S (Fig. 2), retrieves the class Chytridiomycetes (sensu Powell & Letcher, 2014) and all the major chytrid orders, most of them with a high (> 98% UFboot) support. The core order Rhizophydiales is retrieved with 100% support to the exclusion of the genera Batrachochytrium, Entophlyctis and Homoloaphlyctis. The isolate SkChyt5 appears to cluster in a highly supported clade containing the family Kappamycetaceae and Alphamycetaceae; it falls within the latter family although with very weak support (43% UFboot). In a second reconstruction, consistently with previous Rhizophydiales phylogenies (Letcher et al., 2012; Seto and Degawa, 2018), we assessed the phylogeny of the order Rhizophydiales using a concatenation of 5.8S and 28S (Fig. 3). Even if the clade composed by Kappamycetaceae and Alphamycetaceae was still retrieved with high support (99.2% UFboot), the overall result slightly changed, by including isolate SkChyt5 as a sister clade to Gammamyces ourimbahiensis and Alphamyces chaetifer, with moderate support (91% UFboot). Furthermore the family Alphamycetaceae appeared to be paraphyletic, with the genus Betamyces (93.5% UFboot) clustering as sister clade to the Kappamycetaceae. Finally, in order to be able to...
assess the relationships between our isolate and the environmental diversity of Rhizophydiales, we computed an 18S-based phylogeny including environmental molecular data (Fig. 4). In this phylogenetic reconstruction, isolate SkChyt5 was found to be sister (96% UFboot) to Betamyces sp. isolate Barr-316 and three environmental sequences. The clade composed of Alphamycetaceae and Kappamycetaceae was once more strongly supported (100% UFboot) and encompassed a rich environmental diversity from lakes and high elevation soils from around the world. Within the 18S of SkChyt5 we identified two intronic sequences, one of which contained a CDS encoding a His-Cys box homing endonuclease as per its UniProt annotation (Q8TGE3).

3.4 Global distribution

Screening of the Ocean Sampling Day (Kopf et al., 2015) metabarcoding dataset revealed the presence of organisms closely related to isolate SkChyt5 in 6 sampling stations across the North Atlantic (Fig. 5). The OSD project is a collaborative and global sequencing campaign aiming at analysing the marine microbial community annually on the summer solstice (21st of June). Here we screened the databased generated by the OSD 2014 campaign (PRJEB8682), during which seawater samples collected from 191 sites within exclusive economic zones (EEZs) around the world (see annexed map Suppl. Fig. 3) were filtered and processed for bulk DNA extraction and amplification of the 18S V4-hypervariable region with universal primers. Paired reads were from 96.5% to 100% identical to SkChyt5-V4-18S-rDNA and clustered in two different OTUs. Reads below an identity threshold of 98.4 % to SkChyt5-V4-18S-rDNA clustered in OTU2 (blue in Fig. 5) and shared a synapomorphy, which was not detected in any of the reads belonging to OUT1 (> 98.4 %, red in Fig. 5) nor in SkChyt5 18S sequence (Suppl. Figure 2). OTU1 was detected on both sides of the North Atlantic, in particular in Raunefjorden (Norway, 60°09'40.4"N, 5°06'54.1"E) and Booth Bay (Maine, USA, 43°50'39.8"N, 69°38'27.2"W), the former with a higher number of retrieved reads (n=67) as compared to the latter (n=34). Reads belonging to OTU2 were retrieved from the European Atlantic coasts only, with the highest read abundance (n=67) offshore Pasaia (Basque Country, Spain, 43°20'00.0"N, 1°55'30.0"W). Weaker read abundances associated to OTU2 have been detected in
Discussion

The overall thallus morphology of the parasitic organism infecting *Skeletonema* during the 2016 spring bloom already suggested its inclusion within the Chytridiomycota. Zoospores with a single whiplash flagellum and a lipid globule (Fig. 1C inlet, F, G), later developing in an endogenously generated monocentric thallus (Fig. 1B inlet), are in line with extant descriptions of chytrid zoospores and with the definition of the *Chytridium*-type development given by Sparrow (1960). WGA-FITC highlighted the presence of chitin in the sporangia cell wall (Fig 1, D), a characteristic of true fungi useful in distinguishing chytrids from other unicellular parasites (Sparrow, 1960). As already observed in previous studies (Rasconi *et al.*, 2009), a better labelling of the rather inconspicuous intracellular rhizoidal system was obtained with the less selective stain Calcofluor White (Fig. 1C arrow). Such a reduced rhizoidal system has already been reported for other parasitic Rhizophydiales interacting with their algal host via a “peg-like” rhizoid (Van den Wyngaert *et al.*, 2017). Calcofluor White staining also highlighted the presence of cell wall thickenings on the surface of mature sporangia (Fig. 1C), reminiscent of the papillae, later developing into discharge pores, reported for many Rhizophydiales (Longcore, 2004; Letcher *et al.*, 2015) and other chytrid orders (Simmons *et al.*, 2009; Davis *et al.*, 2016). Unfortunately, spore discharge could not be directly observed in our samples; therefore the fate of the sporangia wall thickenings remains hypothetical. The rarely observed thick-walled structures, bearing a single eccentric lipid globule, clearly remind typical chytridiaceous resting spores (Longcore, 2004; Letcher and Powell, 2005; Letcher *et al.*, 2006). In Fig. 1 (H-I) epifluorescence microscopy and CW staining highlighted the presence of an encysted spore on one of these structures, consistent with one of the modality of sexual reproduction in the order Rhizophydiales, where a contributing spore (male gamete) encysts on a receptive thallus (female gamete), to which it transfers its cytoplasmic contents resulting in the
formation of the thick-walled resting spore (Sparrow, 1960; Van den Wyngaert et al., 2017). In line with morphological evidence, the nearly complete rDNA operon obtained in silico for isolate SkChyt5 confirmed that this epibiotic parasite of Skeletonema belongs to the order Rhizophydiales in the phylum Chytridiomycota. Within this order, our isolate is firmly placed within a well-supported group composed of the families Alphamycetaceae and Kappamycetaceae (Letcher and Powell, 2005; Letcher et al., 2012). However the Skeletonema-infecting chytrid does not clearly belong to any of these two families in our concatenated phylogenetic reconstruction (Fig. 2); instead, isolate SkChyt5 clusters sister to the Alphamycetaceae with a very weak ultrafast bootstrap support (43%). The family Alphamycetaceae is divided in two main groups; the first composed of the genus Betamyces alone and the second composed of the genera Alphamyces and Gammamyces. Those two main sub-clades have previously been found not to be supported by strong bootstraps (Letcher et al., 2012; Seto and Degawa, 2018), a feature consistent with our 5.8S-28S phylogenetic reconstruction where the Alphamycetaceae appear paraphyletic (Fig. 3). Even if in this latter phylogenetic reconstruction the parasitoid of Skeletonema clusters closer to the genera Alphamyces and Gammamyces, any conclusion on the affinity of SkChyt5 to extant taxa would be assumptive. In the absence of taxonomically informative information data on zoospore ultrastructure (James et al., 2000), and of a broader sequenced diversity within the Rhizophydiales as comparison, we refrain from any taxonomic treatment of the investigated Skeletonema parasitoid and we adopt the provisional name SkChyt, following published guidelines for temporary nomenclature for protists (Berney et al., 2017).

We used phylogenetic analyses of molecular data from 18S-based metabarcoding surveys to assess whether chytrids related to SkChyt shared similar habitats or ecological strategies. High bootstrap support (96%, Fig. 4) places SkChyt5 sister to the saprobe Betamyces sp. Barr-316 (Smith et al., 2014) and to a well-supported clade of isolates from freshwater Antarctic cyanobacterial mats (Jungblut et al., 2012), a glacier forefield in Tibet (Khan and Kong, unpublished) and Japanese lakes (Ishida et al., 2015). No marine isolate clusters close to SkChyt5 and only three isolates come from marine/brackish water in our phylogenetic reconstruction. This result is likely an artefact due to the
undersampling of the marine environment, potentially coupled with the marked seasonality shown by chytrids in the plankton (Picard, 2017); the latter can be exacerbated in the case of parasites, whose ephemeral presence and short lasting dynamics remain often undetected by molecular surveys (Garvetto et al., 2018). The global distribution of the SkChyt highlighted its presence in coastal areas around the north Atlantic (Fig. 5). Since the Ocean Sampling Day 2014 campaign is limited to coastal areas sampled at one time point (21st of June; Kopf et al., 2015), we cannot exclude that SkChyt may also occur in the open sea. Likewise, we cannot rule out the hypothesis that a strong seasonality for the studied chytrid (possibly driven by its diatom host) might have prevented its detection in the southern hemisphere, where Skeletonema is virtually absent during the Austral winter. Indeed, microscopic evidence of a chytrid parasite infecting Skeletonema sp. in the Humboldt Current System off central Chile was recently reported, but unfortunately no matching molecular data were available (Gutiérrez et al., 2016). This chytrid was hypothesised to infect both Skeletonema sp. and Thalassiosira sp. during the austral spring and late summer, therefore driving the phytoplankton succession by allowing for the development of blooms of the immune diatom Chaetoceros sp. On the other hand, one operational taxonomic unit (OTU14) closely related to Betamyces americaemeridionalis ARG063, and therefore to SkChyt, was detected for three years (2008, 2009 and 2011) in co-occurrence with Chaetoceros sp. dominated blooms in the English Channel (Taylor and Cunliffe, 2016). Although no direct observation of parasitism exists, these results strongly suggest that Chaetoceros might be a host of OTU14. Our observations do not allow speculations on the infection of hosts other than Skeletonema, since both Thalassiosira and Chaetoceros were present only in very low abundances in our samples, and none was found to bear sporangia. The peculiar detached sporangia described in Gutiérrez et al. (2016) were not observed in the samples investigated here, where all chytrid thalli grew epiphytically on Skeletonema colonies.

Our screening of metabarcoding datasets highlighted the presence of two OTUs clustering around SkChyt5 with an identity varying from 100 to 98.4% (OTU1) and from 98.4 to 96.5% (OTU2). Each of them shared synapomorphies on the V4 hypervariable region of the 18S rDNA gene (detailed in
Supplementary Figure 2), suggesting that these two OTUs correspond to genuine, distinct taxa. Whether or not the two OTUs retrieved in this study belong to different yet closely related species, the possibility that other chytrid parasites infecting marine diatoms fall within the clade Alphamycetaceae/Kappamycetaceae should also be considered.

Thanks to the combination of PCR and Hiseq sequencing, we identified an intronic sequence one base pair upstream the primer binding site used for the amplification of the V4 hypervariable region; a second intron contains a coding sequence for a homing endonuclease enzyme, indicating that it is active and spreading within SkChyt population (Chevalier and Stoddard, 2001). The presence of intronic sequences, a well-known feature for the 18S ribosomal RNA of chytrids and other fungi (Karpov et al., 2017), may therefore interfere with the amplification of metabarcodes. It should be envisaged that metabarcoding studies, despite their proven power in unveiling hidden planktonic microbial diversity (de Vargas et al., 2015), may suffer from such biases, especially towards certain taxa. We argue that the caveats described here might have been causing a strong underappreciation of the role of fungal parasites in controlling and affecting Skeletonema and other diatoms in the global ocean, possibly causing biases in modelling biogeochemical cycling. To date, predation is believed to be the largest source of mortality of phytoplankton (Calbet and Landry, 2004) causing changes in prey physiology and behaviour (Amato et al., 2018) with effects potentially cascading throughout the food web; but little, if any, information on parasitism is included in marine food web models. Examples from freshwater systems have been pointing towards the involvement of chytrids in driving the succession of dominant phytoplankters (Van Donk and Ringelberg, 1983) and results from (Gutiérrez et al., 2016) seem to confirm these observations in the marine habitat. Molecular markers obtained by SC analyses may constitute a useful tool to target taxa that remain undetected in metabarcoding surveys, with the added benefit of potentially linking barcodes to ecological functions, thus empowering the ecological interpretation of long term plankton time series. We hope that the molecular data provided here will help triggering further investigation on the presence
and function of marine chytrids, which have the potential to be important drivers of phytoplankton dynamics and associated biogeochemical cycles in the ocean.

Acknowledgements

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 642575 (ALFF), the FP7 Marie Curie Actions PEOPLE-2007-2-2. ERG No. 230865, the grant agreement No. 727892 (GENIALG) and from the United Kingdom NERC under the grants GlobalSeaweed (NE/L013223/1) and MultiMARCAPP (NE/L013029/1).

Figure Captions

Table 1 Details of the parameters used for the PCR reactions carried out in this study.

Fig. 1 Microscopy observations of the chytrid parasite of Skeletonema. (A) Uninfected Skeletonema sp. chain; (B) Infected Skeletonema chain showing two spherical sporangia filled with refractive globular structures. Inset: Spherical sporangium filled with refractive globules, growing on the valvar surface of the host frustule, among fultoportula processes. (C) Calcofluor White staining (in blue) of the chytrid sporangia shown in (B). Note the presence of an infection plug (arrow) and various cell wall thickenings (arrowheads). Red chlorophyll autofluorescence of diatom phaeoplasts highlights the suffering of the sporangia-bearing cells. Inset: Encysted zoospore showing a germination tube (arrow) and a small Nile Red-positive lipid globule (arrowhead). (D) WGA-FITC staining (in green) of a chytrid sporangium in DIC (left) and epifluorescence (right). (E) Epifluorescence picture of a Nile Red-stained sporangium, highlighting the lipidic nature of the refractive globules observed in the sporangium. (F-G) Zoospores settled on the valvar surface of Skeletonema, showing the refractive lipid globule (arrowhead) and flagellum (arrows). (H) Putative resting spore, with thick cell wall and eccentric lipid globule in DIC. The blue colouring observed in DIC stems from the Evans Blue.
counterstaining of Calcofluor White. (I) Epifluorescence observation of an attached zoospore to the putative resting spore shown in (H). Scale bars = 5 μm.

**Fig. 2** Maximum Likelihood reconstruction (1000 Ultra-Fast Bootstraps) of chytrid phylogeny based on three concatenated rRNA encoding gene sequences (18S, 5.8S and 28S). Symbols near the species name indicate the presence (*) or absence (-) of genes encoding 18S, 5.8S and 28S respectively in the alignment. The chytrid parasitoid of *Skeletonema* sp. SkChyt5 is highlighted in bold.

**Fig. 3** Maximum Likelihood phylogenetic reconstruction (1000 Ultra-Fast Bootstraps) of the order Rhizophydiales based on the concatenated rDNA genes 5.8S and 28S. The chytrid parasitoid of *Skeletonema* sp. is in bold.

**Fig. 4** Maximum Likelihood phylogenetic reconstruction (1000 Ultra-Fast Bootstraps) of the order Rhizophydiales based on the 18S rDNA, taking into account environmental sequences from metabarcoding surveys. The coloured dots besides the entries highlight ecological annotations: freshwater (light blue), soil (brown), high elevation soil/snow/ice (grey), marine (dark blue), brackish water (light and dark blue). A red “P” indicates entries known to be parasites. A white dot with “?” indicates lack of information. Sequences of known taxonomic affiliation are in bold. The chytrid parasitoid of *Skeletonema* SkChyt5 is highlighted by a larger font.

**Fig. 5** Map showing the distribution of OTUs related to the parasitoid of *Skeletonema* SkChyt5, as reconstructed from the Ocean Sampling Day database. OTU1 (red) comprises reads matching SkChyt5 rDNA-18S-V4 with an identity above 98.4 %, whilst OTU2 (blue) comprises reads identical to the reference from 98.4 to 96.5 %. The diameter size of the points on the map is proportional to the number of reads retrieved in the related sampling station, expressed by the digits within the circle.
The red star indicates the isolation point of SkChyt5 and the other biological samples used in this study.


Collimycetaceae

Paranamyces uniporus PL157
Paranamyces uniporus WJD150

Ulkenomyces aestuarii ATCC26190
Halomyces littoreus ATCC36100
Paludomyces mangrovei ATCC26191

Protrudomyces sp. JEL134
Protrudomyces sp. JEL400
Protrudomyces lateralis ARG072

Aquamyces sp. ARG009
Aquamyces chlorogonii ARG020
Aquamyces chlorogonii ARG002

Angulomycesargentensis ARG008
Angulomycesargentensis ARG015

Coralloidiomyces digitatus PL163

Operculomyces laminatus JEL223

Dinomyces arenysensis P233

Dinomyces sp. P242

Uebelmesseromyces harderi ATCC24053

Urceomyces sphaerocarpus ARG038
Urceomyces sphaerocarpus ARG048
Globomycespollinis-pini JEL150
Globomycespollinis-pini ARG068
Globomycespollinis-pini Barr003

Staurastromyces oculus STAU-CHY3

Pateramyces corrientinensis ARG031
Pateramyces corrientinensis ARG027
Pateramyces corrientinensis ARG046

Rhizophydi um sp. ARG051
Rhizophydi um sp. PLUSR002
Rhizophydi um globosum JEL222
Rhizophydi um brooksianum JEL136
Rhizophydi um sp. PL127

Betamyces americaemeridionalis ARG063
Betamyces sp. Barr-316
Betamyces sp. PL173

Kappamyces sp. JEL001
Kappamyces laurelensis PL098
Kappamyces sp. PL104
Kappamyces sp. JEL356

Gammamyces ourimbahensis PL116
Alphamyceschaetifer ARG025
Alphamyceschaetifer ARG058

Skele tonea parasi toid SkChyt5
Gammamycesourimbahensis PL116
Alphamyceschaetifer ARG025

Boothiomyces macroporosum PLAU021
Boothiomyces macroporosum PLAU021
Boothiomyces sp. JEL055
Terramyces subangulosum PL003
Terramyces subangulosus PL152

Betamyces americaemeridionalis ARG063
Betamyces sp. Barr-316
Betamyces sp. PL173

Kappamyces sp. JEL001
Kappamyces laurelensis PL098
Kappamyces sp. PL104
Kappamyces sp. JEL356

Gammamycesourimbahensis PL116
Alphamyceschaetifer ARG025
Alphamyceschaetifer ARG058

Skele tonea parasi toid SkChyt5
Gammamycesourimbahensis PL116
Alphamyceschaetifer ARG025

Boothiomyces macroporosum PLAU021
Boothiomyces macroporosum PLAU021
Boothiomyces sp. JEL055
Terramyces subangulosum PL003
Terramyces subangulosus PL152

Outgroup

Gorgonomyces haynaldii ARG036
Gorgonomyces haynaldii Barr100

Gorgonymycetaceae

Collimycetaceae

Paranamyces uniporus PL157
Paranamyces uniporus WJD150

Ulkenomyces aestuarii ATCC26190
Halomyces littoreus ATCC36100
Paludomyces mangrovei ATCC26191

Protrudomyces sp. JEL134
Protrudomyces sp. JEL400
Protrudomyces lateralis ARG072

Aquamyces sp. ARG009
Aquamyces chlorogonii ARG020
Aquamyces chlorogonii ARG002

Angulomycesargentensis ARG008
Angulomycesargentensis ARG015

Coralloidiomyces digitatus PL163

Operculomyces laminatus JEL223

Dinomyces arenysensis P233

Dinomyces sp. P242

Uebelmesseromyces harderi ATCC24053

Urceomyces sphaerocarpus ARG038
Urceomyces sphaerocarpus ARG048
Globomycespollinis-pini JEL150
Globomycespollinis-pini ARG068
Globomycespollinis-pini Barr003

Staurastromyces oculus STAU-CHY3

Pateramyces corrientinensis ARG031
Pateramyces corrientinensis ARG027
Pateramyces corrientinensis ARG046

Rhizophydi um sp. ARG051
Rhizophydi um sp. PLUSR002
Rhizophydi um globosum JEL222
Rhizophydi um brooksianum JEL136
Rhizophydi um sp. PL127

Betamyces americaemeridionalis ARG063
Betamyces sp. Barr-316
Betamyces sp. PL173

Kappamyces sp. JEL001
Kappamyces laurelensis PL098
Kappamyces sp. PL104
Kappamyces sp. JEL356

Gammamycesourimbahensis PL116
Alphamyceschaetifer ARG025
Alphamyceschaetifer ARG058

Skele tonea parasi toid SkChyt5
Gammamycesourimbahensis PL116
Alphamyceschaetifer ARG025

Boothiomyces macroporosum PLAU021
Boothiomyces macroporosum PLAU021
Boothiomyces sp. JEL055
Terramyces subangulosum PL003
Terramyces subangulosus PL152

Outgroup

Gorgonomyces haynaldii ARG036
Gorgonomyces haynaldii Barr100

Gorgonymycetaceae
OTU1 (98.4-100%)
OTU2 (96.5-98.4%)

Maine Booth Bay (USA)
Oostende (Belgium)
Zeebrugge (Belgium)
Oban (UK)
Raunefjorden (Norway)
Scalloway (UK)
Pasaia (Spain)
<table>
<thead>
<tr>
<th>Primer Couple</th>
<th>1st Step (Denaturation)</th>
<th>2nd Step (Amplification)</th>
<th>3rd Step (Elongation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Denaturation</td>
<td>Annealing</td>
<td>Elongation</td>
</tr>
<tr>
<td>MH2-NS4</td>
<td>94°C x 3 min</td>
<td>94°C x 30 s</td>
<td>58°C x 1 min</td>
</tr>
<tr>
<td></td>
<td>94°C x 30 s</td>
<td>55°C x 1 min</td>
<td>72°C x 1 min</td>
</tr>
<tr>
<td></td>
<td>94°C x 30 s</td>
<td>52°C x 1 min</td>
<td>72°C x 1 min</td>
</tr>
<tr>
<td></td>
<td>94°C x 30 s</td>
<td>48°C x 1 min</td>
<td>72°C x 1 min</td>
</tr>
<tr>
<td>NS1-ITS4</td>
<td>94°C x 30 s</td>
<td>94°C x 15 s</td>
<td>48°C x 1 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EaF3-D3Ca</td>
<td>94°C x 30 s</td>
<td>94°C x 15 s</td>
<td>58°C x 1 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Figure 1: PCR and *in silico* reconstruction of *SkChyt5* rDNA

**Legend:**

**PCR**: primers detailed in the figure on each PCR product.

**Guided assembly approach**: contigs retrieved by the customised pipeline (Supplementary material 1, ContigBlaster_V8.sh) using *Kappamyces laurelensis* AFTOL-ID 690 (DQ536478, DQ536494 and DQ273824) as query gene on the dataset of paired and reads. The resulting subset (1453 reads) was afterwards assembled in Geneious (KL_CONTIGs in figure below).

**De novo approach**: contigs retrieved by the customised pipeline (Supplementary material 1, ContigBlaster_V8.sh) using SkChyt5 18S retrieved by PCR (MH2-NS4) as query gene on the contig dataset generated by unsupervised assembly of paired end reads in CLC Genomics Work Bench 8 (DeNovo_CONTIG in figure below).

**PCR products and contigs retrieved by both methods match with 100% identity**
Supplementary Figure 2: Synapomorphies detected in OTU 1 and OTU 2

98.4% identity to SkChyt5 18S V4
Suppl. Figure 3: The Ocean Sampling Day Campaign of 2014. Sampling stations are indicated by the grey dots on the map.
Supplementary materials 1: Bioinformatics details

**Trimmomatic**: parameters used.

```java
java -jar trimmomatic.jar PE -threads 32 -phred33 -trimlog SkChyt5_trimming SkChyt5_S2_L002_R1_001.fastq SkChyt5_S2_L002_R2_001.fastq SkChyt5_paired1 SkChyt5_unpaired1 SkChyt5_paired2 SkChyt5_unpaired2
ILLUMINACLIP:FastQC_detected_IlluminAdapters.fasta:2:30:10 SLIDINGWINDOW:4:15 LEADING:3 TRAILING:3 MINLEN:36
```

Adapters detected by FastQC:

```
>TruSeq Adapter, Index 4 (100% over 50bp)
GATCGGAAGAGCAGACGTCTGAACTCCAGAGTGTAGATCTCGTATGC

>Illumina Single End PCR Primer 1 (100% over 50bp)
GATCGGAAGAGCGGTGTAGGAAAAGAGGTAGATCTCGTGGTGCCCG
```
Customised script to retrieve target genes from transcriptome/genome datasets

#!/bin/bash

#input: positional arguments: 1) assembled transcriptome or genome (.fasta), 2) seq(s) of gene(s) (.fasta) of interest to be searched in the genome/transcriptome

#usage example: ContigBlaster_V8.sh genome_transcriptome.fas query_seqs.fas

#step 1: Transform the genome in a BLAST database

for i in $1; do
    BLASTDB=$(makeblastdb -in $i -dbtype nucl -title "$1-BLASTdb" -out "$1"bdb -parse_seqids)
    done

#step 2: search the gene(s) of interest in the genome and extract the names of the matching contigs and use them to retrieve their sequences.

#linearizes and shortens the name to produce a fasta as below:

#>CONTIG1_NAME
#sequence1tagtgtaggtgtgctgcgagtcatgtatagct

#>CONTIG2_NAME
#sequence2atgctgtatcgactgtagctgcatatatagct

#remove useless file and store BLASTdb in a single folder

#message: how many Hits your gene(s) have in the genome, i.e. how many contigs were retrieved by the gene.

for i in $2; do
    SEARCHDB=$(blastn -db "$1"bdb -query $i -outfmt "10" > "$i"_bestHits.csv)
    NAMEXTRACT=$(awk -F ""*,""*"" '{print $2}' "$i"_bestHits.csv > "$i"_besthits_Ids)
    SEQSEXTRACT=$(blastx -db "$1"bdb -dbtype nucl -entry_batch "$2"_besthits_Ids -outfmt %f -out "$2"_seqs.fasta)
    LINEARIZE=$(sed -e 's/\([^>-]*\)/\1\n' "$2"_seqs.fasta | tr -d "\r" | tr -d "\n" | sed -e 's/$/#/' | tr "#" "\n" | sed -e '/^$/d' > "$2"_blastres.fasta)
    NAMESHORTENING=$(sed 's/len.*//g' "$2"_blastres.fasta > "$2"_contigs.fasta)
    CLEANING=$(rm "$i"_bestHits.csv | rm "$i"_besthits_Ids | rm "$2"_seqs.fasta | rm "$2"_blastres.fasta)
ORGANIZEBDB=$(mkdir "$1"BLASTdb | mv "$1"bdb.* $1BLASTdb)
echo "$2 have $(grep -c '^>' "$2"_contigs.fasta) hits in $1"
done

#step 3: BLAST each retrieved contig against GenBank and gather the first 10 hits
#retrieves the corresponding accession numbers and the contig Id
for i in "$2"_contigs.fasta; do
    BLASTNCBI=$(blastn -db nr -query $i -outfmt "10" -max_target_seqs 10 -perc_identity 98 -max_hsps 1 -remote > BLAST.csv)
    ACCESSIONS=$(awk "\""^\""*\""*\""\"*print $2\"" BLAST.csv > accessionumbers.txt)
    done
echo "Accession numbers retrieved... looking for taxonomy"

#step 4: use the accession numbers to retrieve the taxonomy of the organism (according to GenBank).
#produces a .csv file containing the contig ID - 10 accessions and their taxonomy.
#removes useless files
echo "BLASTing $(grep -c '^>' "$2"_BLASTres.fasta) contigs"
for i in $(cat accessionumbers.txt); do
    TAXLINAG=$(esearch -db nucleotide -query $i -sort id | efetch -db nucleotide -format fasta | awk -F '>' '{print $2}' | sed -n 1p | sed 's/,//g' >> Geneld.txt)
    CONCATENATE=$(paste BLAST.csv Geneld.txt > BLASTtax.csv | sed 's/\t/\t/g' BLASTtax.csv > BLAST_"$2"_contigs.csv)
    done
CLEANING2=$(rm accessionumbers.txt | rm Geneld.txt | rm BLASTtax.csv | rm BLAST.csv)
echo "DONE! Output contains: retrieved contig sequences in "$2"_contigs.fasta + BLAST results for each contig in BLAST_"$2"_coontigs.csv"
Parameters used for the MOULINETTE script.

QUERY= SkChyt5_v4_rDNA.fasta

LIST= OSDlist.txt

PERCID= 97

MAXTARGET= 100

READCOVER= 0.8

EVALUE= 1e-130