THESE DE DOCTORAT DE L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité : Biologie marine – Ecole Doctorale Diversité du Vivant

Présentée par

M. Johan Decelle

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

Ecologie et évolution de la photosymbiose chez les acanthaires (radiolaires)

Soutenue le :

Devant le jury composé de :

Pr. David A. Caron, University of Southern California, USA	Rapporteur
Dr. Fabrice Vavre, Université Claude Bernard Lyon 1	Rapporteur
Dr. John R. Dolan, UPMC/CNRS Observatoire Océanologique de Villefranche-sur-Mer	Examinateur
Pr. Philippe Vandenkoornhuyse, Université de Rennes I	Examinateur
Pr. Gaby Gorsky, UPMC/CNRS Observatoire Océanologique de Villefranche-sur-Mer	Examinateur
Dr. Fabrice Not, UPMC/CNRS Station Biologique de Roscoff	Directeur de thèse
Dr. Colomban de Vargas, UPMC/CNRS Station Biologique de Roscoff	co-directeur de thèse

Abstract

Symbiotic relationships are widespread in nature and are fundamental for ecosystem functioning and the evolution of biodiversity. In marine environments, mutualistic symbiosis with microalgae (photosymbiosis) has been extensively studied in coral reef ecosystems but remains poorly understood in the open ocean despite the key role of oceanic plankton in global ecology and biogeochemistry. The Acantharia (Radiolaria), abundant and ubiquitous single-celled eukaryotes characterized by a mineral skeleton in strontium sulfate, are key players in the pelagic environment. Although their photosymbiotic relationship was discovered in the 19th century, it has barely been studied since. Acantharia are a promising model to explore the evolutionary and ecological significance of photosymbiosis in the planktonic realm. The main objective of my thesis was to unveil the identity and diversity of the symbiotic microalgae, and to investigate the influence of the symbiosis on the evolution, the ecology and the life cycle of the Acantharia at different scales, from cells to ecosystems.

In order to explore the evolutionary context of the symbiosis, phylogenetic analyses were performed with 78 and 107 partial sequences of the 18S rDNA and 28S rDNA ribosomal genes, respectively. Each of these sequences was obtained from morphologically-identified single cells of Acantharia. This comprehensive molecular phylogeny, composed of ten clades, showed a global discrepancy with the 100-years-old traditional taxonomy. The most recent and genetically diverse clades (clades E and F) are represented by the symbiotic Acantharia, which have a complex skeleton compared to non-symbiotic taxa. This suggests that symbiosis has been a key innovation for the increase of phenotypic complexity. Phylogenetic analyses were also conducted on the endosymbiotic microalgae in different host species and geographic regions. We found that all symbiotic Acantharia of clades E and F consistently host cells of the haptophyte genus *Phaeocystis*, a key phytoplankton taxon in marine ecosystems, from which the symbiotic mode was unknown. Molecular clock and cophylogenetic analyses showed that this partnership originated in the highly oligotrophic Mesozoic ocean (ca. 175 Mya), and that biogeography, rather than host taxonomy, is the main determinant of the association. In one early diverging acantharian species, another association with microalgal symbionts from different lineages (haptophytes and dinoflagellates) within a single host cell was characterized. Like *Phaeocystis*, these microalgae are abundant and widely distributed in free-living, occasionally forming extensive blooms. This contrasts with the widely held view in terrestrial and marine symbioses characterized to date, whereby the symbiont is typically elusive outside its host. This original mode of interaction would be a strategy to ensure repeated capture of an appropriate partner in the vast and microbia-diluted open ocean.

Finally, the worldwide distribution and relative abundance of acantharian taxa was explored by a metabarcoding approach. The morpho-genetic database was used to link the environmental V9 sequences of the 18S rDNA gene to a specific morphotype. Symbiotic Acantharia were predominantly present in the photic zone at all latitudes, hence highlighting the ecological success and flexibility of this association in the world's oceans. Contrarily, the non-symbiotic taxa mainly occurred in the meso-and bathypelagic waters. We provided evidences that these acantharian accomplish a phase of their life cycle in deep waters through formation of fast-sinking biomineralized cyst and contribute to the carbon and strontium biogeochemical cycle in the ocean. This work raises hypotheses about morphological evolution of Acantharia, and the functioning and evolution of pelagic photosymbiosis, highlighting the evolutionary forces inherent to the vast and dynamic nature of the pelagic environment.

Résumé

Les relations symbiotiques sont répandues dans la nature et sont fondamentales pour le fonctionnement des écosystèmes et l'évolution de la biodiversité. Dans le milieu marin, la symbiose mutualiste avec des microalgues (photosymbiose) a été très étudiée dans les écosystèmes récifaux, mais reste incomprise dans l'océan ouvert malgré le rôle majeur du plancton dans l'écologie et la biogéochimie globale. Les acanthaires (radiolaires), unicellulaires eucaryotes caractérisés par un squelette minéral en sulfate de strontium, sont des acteurs clés dans le milieu pélagique. Bien que découverte au 19ème siècle, leur relation photosymbiotique a été très peu étudiée depuis. Les acanthaires sont un modèle prometteur pour explorer l'importance évolutive et écologique de la photosymbiose dans le domaine planctonique. Le principal objectif de ma thèse était de dévoiler l'identité et la diversité des microalgues symbiotiques, et d'étudier l'influence de la symbiose sur l'évolution, l'écologie et le cycle de vie des acanthaires à différentes échelles, de la cellule aux écosystèmes.

Pour mettre en évidence le contexte évolutif de la symbiose, des analyses phylogénétiques ont été réalisées avec 78 et 107 séquences partielles de l'ADNr 18S et l'ADNr 28S, respectivement. Chacune de ces séquences a été obtenue à partir de cellule unique d'acanthaire. Cette phylogénie, composée de dix clades, a révélé des conflits avec la taxonomie traditionnelle qui date de 100 ans. Les clades les plus récents et les plus divers génétiquement (clades E et F) sont représentés par les acanthaires symbiotiques, qui ont un squelette complexe par rapport aux non-symbiotiques. Cela suggère que la symbiose a été une source d'innovation pour la complexité phénotypique de ces microorganismes. Les analyses phylogénétiques ont également été menées sur les microalgues symbiotiques à partir de différentes espèces hôtes et de régions géographiques distinctes. Nous avons constaté que tous les acanthaires des clades E et F vivent constamment avec la microalgue *Phaeocystis* (haptophyte), une espèce phytoplanctonique clé dans les écosystèmes marins, dont le mode symbiotique était inconnu. Une horloge moléculaire et des analyses cophylogénétiques ont montré que l'origine de cette symbiose date du Mésozoïque (~175 Mya) dans des conditions sévères d'oligotrophie, et que la biogéographie, plutôt que la taxonomie de l'hôte, est le principal déterminant de l'association. Chez une espèce d'acanthaire d'un clade basal, une autre association a été caractérisée avec des symbiotes de différentes lignées de microalgues (haptophytes et dinoflagellés) au sein d'une seule cellule hôte. Comme *Phaeocystis*, ces microalgues sont abondantes et ubiquitaires en phase libre, formant parfois de grandes efflorescences. Cela contraste avec les symbioses terrestres et marines, où le symbiote est généralement rare en dehors de son hôte. Ce mode original d'interaction serait une stratégie pour assurer la capture récurrente d'un partenaire approprié dans l'océan ouvert, un monde vaste et moins dense en organismes.

Enfin, la distribution géographique et l'abondance relative des groupes d'acanthaires ont été explorées par une approche de *metabarcoding*. La base de données morpho-génétique a été utilisée pour lier les séquences environnementales V9 de l'ADNr 18S à un morphotype. Les acanthaires symbiotiques étaient prédominants dans la zone photique sous toutes les latitudes, soulignant le succès écologique et la flexibilité de cette association. Au contraire, les groupes non-symbiotiques se situaient principalement dans les eaux méso- et bathypélagiques. Nous avons montré que ces acanthaires ont une partie de leur cycle de vie dans ces eaux profondes *via* la formation de kystes très denses, contribuant au cycle biogéochimique du carbone et du strontium. Ce travail de thèse soulève des hypothèses sur l'évolution morphologique des acanthaires, ainsi que sur le fonctionnement et l'évolution de la photosymbiose planctonique, mettant en évidence les forces évolutives intrinsèques de l'immense milieu pélagique.

Table des matières

Préambule

Le milieu océanique est le plus vaste mais probablement le moins compris des écosystèmes de notre biosphère. Une myriade de microorganismes d'une diversité et d'une abondance vertigineuse y vit et interagit depuis des millions d'années. Au 19ème siècle, le rôle écologique de ces microorganismes était encore énigmatique, et certains les considéraient comme de la « *poussière philosophique* » (J. Müller). Puis, un certain Victor Hensen, inventeur du mot plancton, reconnait que le petit peut nourrir le plus grand dans l'océan, et que cet invisible doit être perçu désormais comme le « *métabolisme de la mer* ». Aujourd'hui, nul doute que les microorganismes, bien que représentant une infime biomasse, sont des acteurs clefs dans les réseaux trophiques des océans et les cycles biogéochimiques globaux.

Le plancton marin est constitué d'organismes aussi divers que les virus, bactéries, archées, protistes (eucaryotes unicellulaires) et métazoaires qui, en concentrations parfois très élevées, sont en continuelle interaction. Darwin reconnut à l'époque que les interactions écologiques entre les espèces, qu'il nomma « *the entangled bank* », étaient parmi les processus les plus importants qui façonnent l'évolution et la diversification des espèces. L'évolution de la biodiversité est fondamentalement une histoire d'évolution des interactions entre espèces. Aucun organisme n'est isolé dans l'environnement, chaque individu a ses compétiteurs, parasites, prédateurs, mais peut également dépendre d'interactions mutualistes avec une ou plusieurs espèces. Les conditions abiotiques et les interactions antagonistes ont longtemps été mises en avant pour décrire la structure et l'organisation des communautés. Pourtant, des écosystèmes entiers reposent sur des relations mutualistes, la symbiose marine dans le milieu récifal et la symbiose terrestre chez les plantes et lichens en sont des parfaites illustrations.

Les interactions mutualistes émergent généralement dans un environnement où les capacités métaboliques des organismes sont contraintes. Le milieu océanique, pauvre en nutriments, présente des conditions favorables à l'établissement de ce type d'associations à bénéfice mutuel. Quels sont ces microorganismes en symbiose et comment interagissent-ils? Quel est l'impact de la relation symbiotique sur leurs traits de vie et sur l'écosystème? Les études sur les interactions entre les microorganismes planctoniques sont essentielles pour comprendre leur évolution et leur écologie mais aussi le fonctionnement des écosystèmes pélagiques. Ce travail de thèse se propose d'étudier la photosymbiose planctonique, une interaction symbiotique entre protistes communément observée mais méconnue à la surface des océans.

Introduction générale

1. Les unicellulaires marins au cœur de la découverte de la symbiose

Les premières observations de la symbiose dans le monde vivant n'ont été faites que très tardivement, à la fin du 19ème siècle, et ont bouleversé la pensée scientifique mais aussi philosophique et sociétale (Perru, 2006). Le vivant était considéré comme individualiste ou comme unité isolée évoluant parmi d'autres dans un environnement. A cette époque, les seules associations connues et décrites entre des organismes différents sont de nature parasitique, et la mentalité pasteurienne considère tous les microbes comme des parasites synonymes d'effets délétères dans un corps hôte. Les premiers individus découverts en symbiose feront alors systématiquement l'objet d'analogies avec les parasites.

Au milieu du 19ème siècle, des granulations chlorophylliennes vertes sont observées chez les animaux et protozoaires, comme l'hydre, les plathelminthes et les ciliés. Les propriétés biochimiques de cette chlorophylle « animale » sont étonnement très proches de celles des végétaux, et mènent certains scientifiques à penser que les cellules animales sont aussi capables de faire la photosynthèse. En 1851, le jeune anglais Thomas Henry Huxley, embarqué sur une expédition scientifique dans le Pacifique, observe des petites cellules jaunes dans les colonies de radiolaires (grands protistes hétérotrophes). Il n'a aucune idée quant à l'origine de ces cellules et la bizarrerie de ces organismes le découragent à poursuivre ses investigations. Ces cellules jaunes sont retrouvées également chez d'autres radiolaires par Johannes Muller quelques années plus tard. A cette époque, où ni le mot ni le phénomène symbiose sont décrits, ces cellules jaunes ne peuvent être que des parasites ou des cellules faisant partie de l'animal. Il faut attendre les travaux du russe Cienkowski en 1871 sur les radiolaires, qui démontrent que ces cellules ne sont pas des parties de l'animal mais plutôt des petites algues autonomes pouvant vivre aussi bien à l'intérieur qu'à l'extérieur (Cienkowski, 1871).

Au même moment, l'union de deux organismes est aussi observée dans le monde végétal sur terre. Albert Bernhard Frank découvre la relation symbiotique entre les plantes et des champignons, appelés mycorhizes (Frank, 1877). La double nature des lichens est mise en évidence par Simon Schwendener, qui observe à la fois des fils de champignons et la multiplication de petites algues lors de la croissance du lichen (Schwendener, 1868). Ces découvertes bouleversent le monde scientifique et lancent un débat en Europe. Certains scientifiques alertent la communauté sur la nécessité de promulguer une définition biologique pour qualifier ces nouveaux modes de vie.

Le mot symbiose apparait pour la première fois dans un contexte politique au IIe siècle avant J.C par l'historien Plèbe, puis resurgit au 17ème siècle pour définir l'organisation de la vie

politique où des communautés de citoyens (symbiotes) s'associent. En 1877, Frank qualifie une association entre individus de «*Zusammenleben*» (vivre ensemble en allemand) et inventa le mot «*symbiotismus*». Selon lui, ce n'est pas du simple parasitisme car les deux partenaires ont des rôles nutritifs complémentaires. Travaillant sur le lichen, Anton de Bary reprend la définition de Frank et utilise plutôt le terme «*symbiosis*» (de Bary, 1879). Il définit **la symbiose comme une association physique entre des espèces différentes**, et inclue ainsi le parasitisme. Le mot symbiose a aujourd'hui une forte connotation mutualiste et sa définition dans le dictionnaire décrit en effet une relation à bénéfice mutuel. Nous utiliserons plutôt la définition de de Bary au cours de cette thèse, car cela évite de statuer définitivement la nature des associations observées. En effet, il est très difficile de qualifier avec certitude une relation de mutualiste à partir d'une simple observation. Les relations symbiotiques sont très versatiles dans le temps et dans l'espace, passant parfois d'un mutualisme à un parasitisme (Sachs et Wilcox, 2006). De plus, le mutualisme implique que les deux partenaires ont une valeur sélective (ou fitness) plus importante lorsqu'ils entrent en interaction (maximisation du nombre de descendants). Mesurer la fitness demande un dispositif expérimental avec des organismes en cultures, ce qui en pratique est irréalisable pour de nombreuses symbioses.

De Bary décrit la symbiose dans le règne végétal mais sa définition prend très vite du sens dans le monde animal. A la fin du 19ème siècle, sous l'impulsion des scientifiques français, russes et allemands, l'étude de la symbiose animale prend son essor. Cependant, l'éternel débat face au phénomène persiste : quelle est la vraie nature de cette relation? Cienkowski pense que les petites algues sont des parasites des animaux, tandis que Patrick Geddes et Karl Brandt penchent plus vers l'hypothèse du bénéfice mutuel. Ces derniers réalisent des expériences physiologiques sur les organismes symbiotiques afin de dévoiler la réelle fonction des deux partenaires. Patrick Geddes présente ses travaux en 1878 sur «*La fonction de la chlorophylle avec les planaires vertes*» (Geddes, 1878), entrepris à la Station Biologique de Roscoff avec le soutien d'Henri de Lacaze-Duthiers. Il y démontre le rôle nutritif des microalgues dans les petits vers, qui accumulent de l'amidon grâce à l'activité photosynthétique de leur partenaire. Geddes se rend ensuite à la Stazione Zoologica de Naples pour effectuer des travaux similaires sur les Actinies (anémones). Ses expériences confirment que les microalgues ne sont pas des parasites, mais qu'il existe bien une complémentarité physiologique entre le végétal et l'animal. En 1881, Karl Brandt fait aussi en parallèle des expériences physiologiques avec des radiolaires. Il veut s'assurer par lui-même que ces corps chlorophylliens appelés «corps verts» sont autonomes et ne sont pas des algues parasites. Il place des colonies de radiolaires (collodaires) avec leurs symbiotes à l'obscurité et à la lumière et constate que sans lumière les organismes meurent très rapidement. Les symbiotes ont donc un rôle bénéfique *via* leur photosynthèse et auraient une fonction physiologique analogue à celle des corps chlorophylliens des végétaux. Brandt nomme les cellules jaunes des radiolaires coloniaux et des actinies, les zooxanthelles (*Zooxanthella nutricula*), et les cellules vertes observés chez d'autres animaux comme les ciliés, les zoochlorelles (Brandt, 1881). En 1887, Ernst Haeckel confirme le rôle nutritif des algues en observant des grains d'amidon dans les cellules jaunes de radiolaires collectés par le HMS

Challenger (Haeckel, 1887), mais malgré toutes ces observations et expériences, certains maintiennent que cette symbiose est plus proche du parasitisme que du mutualisme.

Définie en 1875, la définition du « mutualisme » existait déjà à l'époque mais restait ambigu. Van Beneden, son auteur, a des difficultés à trouver les limites entre les trois types de symbiose: le parasitisme, le commensalisme et le mutualisme (Van Beneden, 1875). La notion de mutualisme ne s'impose pas facilement dans les mentalités, tout particulièrement chez les évolutionnistes fervents défenseurs de la récente théorie darwinienne (Saap, 1994). Dans son « *Origine des espèces* », Darwin rejette toute entraide entre organismes. La compétition, vision malthusienne, est le mode d'interaction qui prévaut dans le monde vivant. Au 20ème siècle, les études sur la symbiose animale se sont faites alors très discrètes et ce sont plutôt les relations antagonistes, prédation et parasitisme, qui ont été privilégiées pour comprendre le fonctionnement des écosystèmes.

2. Rôle de la symbiose dans le règne vivant

« *We are symbionts on a symbiotic planet, and if we care to, we can find symbiosis everywhere.* » (Lynn Margulis)

Squelette d'une feuille d'hortensia
(http://en.wikipedia.org/wiki/Image:Vein_sceleton_hydrangea_ies.jpg)
Selon Lynn Margulis, l'évolution des organismes se fait par anastomose, rencontres entre plusieurs branches du vivant.

2.1 Rôle dans l'évolution

Le passage de la vie unicellulaire simple, sans compartiment, à l'émergence de corps plus grands, plus complexes morphologiquement s'est fait grâce à l'établissement d'interactions symbiotiques (Maynard-Smith, 1989; Margulis et Fester, 1991). Apparues relativement tôt au cours de l'histoire de la terre, la diversification de ces nouvelles unités chimériques et leur dispersion dans de multiples écosystèmes ont transformé la biosphère. Sous l'impulsion de Konstantin Mereschkowsky (1905), puis de Lynn Margulis (1970), il est aujourd'hui clairement reconnu que la symbiose, union d'organismes, de génomes et de métabolismes phylogénétiquement éloignés, est une, sinon la source d'innovation majeure dans l'évolution des

organismes (théorie de la symbiogénèse). L'exemple le plus remarquable est l'évolution des eucaryotes dont la condition est fondamentalement symbiotique. Grâce à leur cytosquelette et leur capacité d'endocytose, les eucaryotes ont évolué par endosymbioses successives avec des bactéries puis avec d'autres eucaryotes pour acquérir leurs différentes machineries cellulaires.

Les eucaryotes ont d'abord acquis leur mitochondrie en entrant en symbiose avec des protéobactéries il y aurait environ 1,5-2 milliards d'années (Embley et Martin, 2006). Au cours de l'évolution, les bactéries symbiotiques sont devenues de simples organites fondamentales pour l'énergie de la cellule. La symbiose a également permis l'acquisition puis la propagation de la photosynthèse dans les différentes lignées eucaryotes (Fig. 1). Entre 1.6 et 1.2 milliards d'années, une cellule eucaryote dotée d'une mitochondrie et d'un flagelle a phagocyté et a retenu en symbiose une cyanobactérie photosynthétique libre (Embley et Martin 2006; Reyes-Prieto et al., 2007). Cette union eucaryote hétérotrophe-procaryote phototrophe, appelée endosymbiose primaire, serait un événement unique au cours de l'évolution, qui a fait émerger trois lignées eucaryotes, regroupées sous le nom de Plantae, ou Archaeplastida : les glaucophytes, les rhodophytes (algues rouges) et les chlorophytes (algues vertes). Puis, certaines de ces microalgues ont été à leur tour phagocytées et retenues en symbiose dans un autre eucaryote hétérotrophe. C'est l'endosymbiose secondaire, union eucaryote hétérotrophe-eucaryote phototrophe, qui a eu lieu plusieurs fois et a donné naissance à plusieurs lignées eucaryotes majeures (Fig. 1). Les plastes secondaires sont appelés parfois « Cheval de Troie » car ils sont synonymes d'arrivée massive de gènes nucléaires et plastidiaux du symbiote dans le noyau de l'hôte. Les algues de la lignée rouge contenant un plaste secondaire, comme les haptophytes et les dinoflagellés (Fig. 1), sont actuellement les organismes photosynthétiques les plus divers génétiquement et morphologiquement dans les océans (Falkowski et al., 2004).

L'évolution des eucaryotes s'est faîte ainsi d'imbriquements et de fusions de cellules, de génomes, de métabolismes, tel un « bricolage » de l'évolution selon François Jacob. Les protistes et les métazoaires sont de véritables chimères, une mosaïque de gènes et d'organites complexes venant de plusieurs sources extérieures. Retracer l'évolution de ces organites et plus largement des grandes lignées eucaryotes est une motivation majeure dans la biologie moderne. Le paysage évolutif de l'acquisition de la photosynthèse, dont une grande partie reste énigmatique, se retrouve dans les océans actuels. Une multitude de protistes vivent en symbiose étroite avec des cyanobactéries (endosymbioses primaires) ou des microalgues eucaryotes (endosymbioses secondaires et tertiaires). Ces différents types d'associations, appelées photosymbioses, se situent le long d'un continuum d'intégration, de la symbiose entre cellules à la rétention de plastes (kleptoplastidie). « Capture » et « esclavagisation » sont des termes communément adoptés pour parler de l'acquisition initiale de la photosynthèse dans l'évolution, tandis que les relations contemporaines dans les océans sont plutôt considérées comme symétriques et à bénéfices réciproques. Les microalgues symbiotiques sont-elles choisies « au hasard » ou existe-t-il des critères physiologiques et écologiques ? **Etudier ces symbioses contemporaines entre protistes marins présente ainsi un intérêt tout particulier pour comprendre un peu plus l'évolution et les mécanismes généraux d'acquisition de la photosynthèse.**

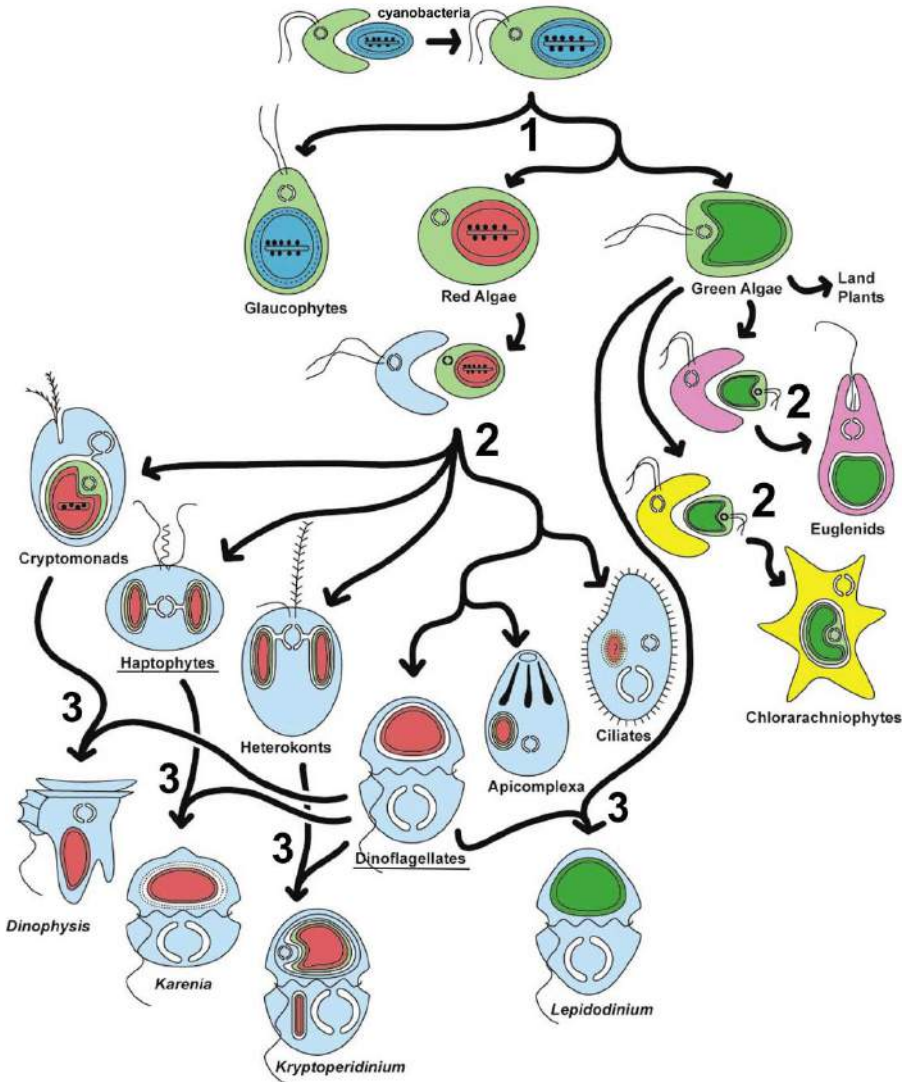


Figure 1. Evolution du plaste chez les eucaryotes par endosymbiose primaire (1), secondaires (2) et tertiaires (3) qui ont donné lieu aux glaucophytes et aux lignées rouges et vertes (indiquées par le code couleur). Les haptophytes et les dinoflagellés (soulignés) seront abordés au cours de cette thèse (figure modifiée d'après Keeling, 2004).

Tous ces événements successifs d'endosymbioses ont été des sources d'innovations phénotypiques, génomiques, mais aussi écologiques. Les océans puis ensuite les continents terrestres ont été peuplés respectivement d'une incroyable diversité de microalgues et de plantes, qui avec leur nouvel organite, le plaste, ont façonné l'atmosphère de la terre. Par exemple, le phytoplancton marin ne représente que 1% de la biomasse des organismes photosynthétiques sur terre mais est responsable de quasiment la moitié de la production primaire (Field et al., 1998). La diversification de ces producteurs primaires a bouleversé les réseaux trophiques des écosystèmes et les cycles biogéochimiques de la terre. L'acquisition de la mitochondrie et de la

photosynthèse ne sont pas les seuls exemples d'unions symbiotiques entraînant des impacts écologiques majeurs.

2.2 Rôle écologique dans la biosphère

La symbiose mutualiste concerne chaque branche du vivant et tient un rôle écologique majeur dans de nombreux écosystèmes. Cette relation entre différents niveaux trophiques permet aux organismes de coloniser et de prospérer dans des environnements auparavant inaccessibles. L'atout écologique majeur est le couplage métabolique très étroit entre les partenaires sans perte par diffusion dans le milieu et sans coût de rapprochement. Le recyclage des déchets métaboliques permet d'augmenter cette efficacité métabolique. Une protection mutuelle chimique et/ou physique peut également s'établir (Selosse, 2000). Le succès écologique des symbioses mutualistes a permis l'apparition et la diversification de nombreux organismes au cours de l'évolution.

De nombreuses associations symbiotiques font intervenir un partenaire photosynthétique. Les mycorhizes sont probablement le système symbiotique le mieux connu occupant une place indiscutable dans différents écosystèmes terrestres. Plus de 90% des espèces végétales reposent sur leur interaction symbiotique avec des champignons pour leur nutrition minérale, plus particulièrement dans des sols pauvres (Bonfante, 2003). Quant aux plantes légumineuses (Fabacées), elles forment des associations avec des bactéries aérobies du sol du genre *Rhizobium,* qui leur confèrent une capacité à fixer l'azote atmosphérique. Les lichens comprennent plusieurs dizaines de milliers d'espèces et couvriraient environ 8% de la surface terrestre du globe (Larson, 1987). Cette association entre un champignon et une microalgue ou une cyanobactérie permet de coloniser des milieux très contraignants marqués par la sécheresse, la salinité ou bien des températures extrêmes (Fig. 2). Dans ces écosystèmes inhospitaliers, les lichens tiennent des rôles importants comme producteur primaire et fixateur d'azote (Potts, 2000).

Figure 2 . De gauche à droite : Photosymbioses terrestres (lichens, J.A. Wearn©) et marines (bénitiers et coraux, Felicia Mc Caullerey©).

Plongé dans une obscurité totale, le fond des océans n'est pas un désert sans vie. Les sources hydrothermales et suintement froids abritent une faune importante, qui repose sur des symbioses mutualistes. Ces interactions entre invertébrés et bactéries endosymbiotiques chimiolithotrophes supportent une biomasse équivalente à celle des invertébrés du littoral (Selosse, 2010). Dans la zone lumineuse des océans, les récifs coralliens, exemples iconiques de la symbiose marine, abritent une incroyable biodiversité et sont de véritables architectes des écosystèmes côtiers tropicaux et subtropicaux (Fig. 2). Alors qu'ils ne représentent que 0.2 % de la surface de la terre (Porter et Tougas, 2000), les écosystèmes récifaux sont d'une importance écologique et économique majeure. Les coraux mais aussi les anémones, les gorgones, les bénitiers, les éponges et des protistes benthiques prospèrent grâce à leur association photosymbiotique avec des microalgues (Stanley, 2006; Trench, 1993).

Alors que la place de la symbiose dans l'évolution des organismes et l'écologie des écosystèmes est indiscutable, paradoxalement ce type d'interaction a été relativement peu étudié dans le milieu pélagique. Les microorganismes planctoniques sont des acteurs clefs dans les réseaux trophiques et les cycles biogéochimiques globaux de la terre. Il apparait donc primordial d'identifier les associations symbiotiques entre ces microorganismes, et de comprendre comment ils interagissent et coévoluent dans l'immensité océanique.

3. La photosymbiose planctonique chez les eucaryotes unicellulaires

"In the sea, there are certain objects concerning which one would be at a loss to determine whether they be animal or vegetable" (Aristote)

3.1 Une adaptation dans les déserts océaniques

Le milieu océanique couvre 70% de la surface du globe et est composé principalement de grandes régions pauvres en nutriments (oligotrophie) car loin de tout apport terrestre. L'activité biologique y est moins importante par rapport aux écosystèmes côtiers, mais l'immensité de ces régions océaniques et les processus microbiens y existants leur donnent une place prépondérante dans la biogéochimie globale et l'équilibre climatique de la planète. Les microorganismes planctoniques présentent différentes stratégies pour faire face à l'oligotrophie. La réduction de leur taille qui permet d'augmenter le rapport surface/volume en est un premier exemple. Les cyanobactéries (*Prochlorococcus* et *Synechococcus*) et les protéobactéries (SAR11) sont les microorganismes les plus abondants à la surface des océans, tandis que les eucaryotes sont représentés principalement par de petits haptophytes et chrysophytes (Morris et al., 2002; Liu et al., 2009; Shi et al., 2011). Dans ces écosystèmes très lumineux, certaines grandes cellules eucaryotes hétérotrophes ont adopté la photosymbiose en s'associant avec un organisme photosynthétique. Ce dernier, généralement le plus petit partenaire, appelé symbiote, est soit une cyanobactérie ou une microalgue eucaryote vivant dans le cytoplasme d'un hôte (le plus grand

partenaire). Dans la majorité des cas, les photosymbioses sont des associations bipartites, impliquant seulement deux types de partenaires (Douglas, 2001).

La photosymbiose est considérée comme mutualiste car les deux partenaires en tirent un bénéfice mutuel, leur procurant une fitness plus importante en symbiose. Les symbiotes photosynthétiques peuvent transférer à l'hôte une grande part du carbone qu'il fixe (Muscatine, 1990; Yellowlees et al., 2008). En retour, l'hôte fournit une protection vis-à-vis des prédateurs, des parasites et virus, et ses déchets métaboliques sont une source nutritionnelle pour le symbiote. L'holobionte (hôte et symbiotes) devient alors un microcosme ou «*cytocosm*» (Taylor, 1982) où la concentration en nutriments est bien plus importante que celle dans le milieu environnant. L'association devient largement autotrophe vis-à-vis du carbone, mais de nombreux hôtes maintiennent cependant leur capacité de prédation et sont qualifiés de mixotrophes (Bé et al., 1981; Caron et al., 1981). Ce couplage trophique étroit entre producteur primaire et consommateur est la principale raison du succès écologique de l'association dans les zones oligotrophiques, plus particulièrement dans les eaux tropicales et subtropicales où l'intensité lumineuse est maximale. Mais la photosynthèse intracellulaire et un fort ensoleillement (source de rayonnements ultraviolets) représentent en contrepartie un stress oxydatif car ils générèrent des radicaux libres d'oxygène ou ERO (espèces réactives de l'oxygène), qui endommagent l'ADN, les protéines, et les membranes (Hörtnalg et Sommaruga, 2007). L'holobionte doit ainsi développer des mécanismes pour prévenir les dommages causés par ces «toxines » cellulaires.

Nos connaissances actuelles sur la photosymbiose marine reposent principalement sur les coraux, les anémones et autre invertébrés récifaux. Les observations faites au 19ème siècle et celles de ces dernières décennies montrent que ce mode de vie est également répandu dans le plancton unicellulaire. **Cependant, la diversité, l'évolution et le fonctionnement de ces photosymbioses planctoniques dans l'un des plus grands écosystèmes de la biosphère restent largement méconnus.**

3.2 La photosymbiose chez les radiolaires et les foraminifères planctoniques

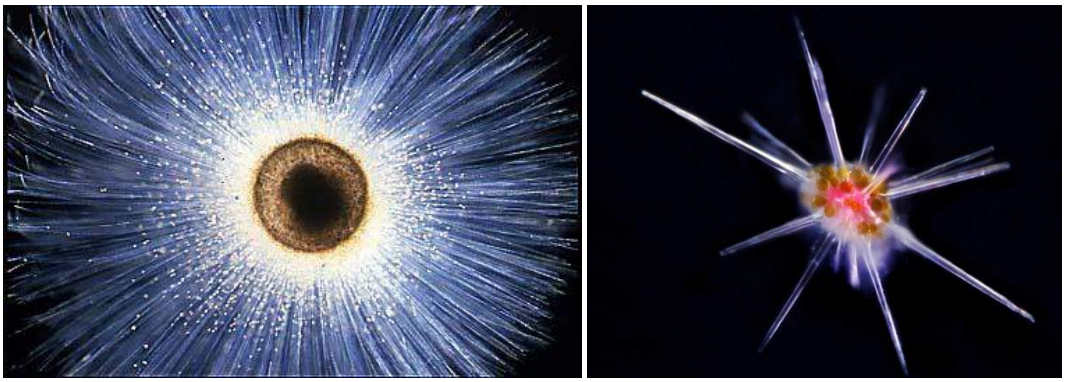


Figure 3. Photos d'un foraminifère (gauche, H. Spero©) et d'un acanthaire (droite, L. Gutiérrez©) vivant en symbiose avec des microalgues (petites cellules vertes-jaunes au centre de la cellule hôte).

Parmi les premières symbioses animales observées, la photosymbiose pélagique est commune chez les radiolaires et les foraminifères (Fig. 3), unicellulaires holoplanctoniques appartenant au grand groupe eucaryote des Rhizaria (Caron, 2000; Moreira et al., 2007, Pawlowski et Burki, 2009). Apparus il y a plus de 500 millions d'années (Pawlowski et al., 2003; Braun et al., 2007), les radiolaires et les foraminifères évoluent aujourd'hui dans tous les océans mais leur relation symbiotique se concentre plus particulièrement à la surface des eaux oligotrophiques aux basses latitudes. Dans le Pacifique équatorial, plus de 90% des foraminifères, et 50% des radiolaires ont été observés avec des microalgues endosymbiotiques (Stoecker et al., 1996). Tel un microcosme dans l'immensité océanique, leur grande cellule amiboïde de 20 µm à 1 mm, dotée d'un cytosquelette et d'un système de phagocytose très performants, peut abriter de nombreuses algues unicellulaires symbiotiques, jusqu'à plusieurs millions chez les colonies de radiolaires (Spero et Parker, 1985; Taylor, 1982). Cette association est obligatoire pour les radiolaires et les foraminifères symbiotiques, car la prévalence de microalgues intracellulaires est maximale et l'inhibition de la photosynthèse entraine un ralentissement de leur croissance voire leur mort (Bé et al., 1982; Anderson, 1983). Il est important de souligner également que cette interaction est transitoire dans la mesure où, à chaque génération, l'hôte doit acquérir *de novo* ses algues symbiotiques depuis l'environnement (transmission horizontale).

Les radiolaires et les foraminifères jouent des rôles clés dans les réseaux trophiques des océans oligotrophiques. *Via* la photosymbiose, ils participent indirectement à la production primaire, parfois jusqu'à 20% (Michaels, 1988). Ils sont également des prédateurs importants de proies extrêmement diverses: des bactéries, microalgues, ciliés jusqu'aux métazoaires. (Anderson, 1983; Swanberg et Caron, 1991). Certaines études montrent que la digestion de leurs propres microalgues endosymbiotiques peut être une source nutritionnelle supplémentaire, et serait un processus permettant de réguler la population des symbiotes à l'intérieur de la cellule hôte (Bé et al., 1983; Anderson 1983).

En plus de leur rôle trophique, les radiolaires et les foraminifères se positionnent comme des acteurs importants dans les flux biogéochimiques de matière vers l'intérieur des océans, en

dehors de la zone photique. En particulier, leur squelette minéral en carbonate de calcium pour les foraminifères et en silice ou en sulfate de strontium pour les radiolaires influence de manière significative les cycles globaux de ces éléments minéraux (Schiebel, 2002; Lampitt et al., 2009). Ce squelette, tel un ballaste, augmente la vitesse de sédimentation de ces microorganismes, qui peuvent se retrouver nombreux dans les trappes à sédiments (Gowing et Coale, 1989; Michaels, 1991), ce qui favorise la séquestration à long terme du carbone et des éléments minéraux du squelette dans les profondeurs.

3.3 Identité des microalgues symbiotiques

Le nom initial donné par Karl Brandt aux microalgues symbiotiques (*Zooxanthellae nutricola*) est aujourd'hui désuet car il n'a aucune valeur taxonomique et ces microalgues s'avèrent être en fait d'une grande diversité. Dans le milieu marin, plus de 70% de ces microalgues ont un plaste rouge, issus d'endosymbioses secondaires, incluant les dinoflagellés, les diatomées et les haptophytes, tandis qu'en eau douce, il s'agit majoritairement d'algues vertes du genre *Chlorella* (Stoecker et al., 2009). Les microalgues symbiotes des invertébrés récifaux (cnidaires, mollusques, gorgones etc.) sont pour la grande majorité des dinoflagellés de l'ordre des Suessiales appartenant au genre *Symbiodinium* (Fig. 4A et 4B), qui comportent de nombreux clades moléculaires (LaJeunesse, 2001; Pochon et al., 2006). Notons que tous ces organismes hôtes sont pour la plupart benthiques et vivent dans les zones côtières peu profondes. Dans le milieu planctonique, les microalgues décrites en symbiose avec les foraminifères appartiennent à deux grandes lignées: les haptophytes (le genre *Chrysochromulina*) et les dinoflagellés (le genre *Pelagodinium*, groupe frère des *Symbiodinium* dans l'ordre des Suessiales, Fig. 4C) (Spero, 1987; Faber et al., 1988; Shaked et de Vargas, 2006). Les radiolaires, comprenant les acanthaires, les spumellaires, les nassellaires et les collodaires (Fig. 5), sont tous capables de former des symbioses avec une grande diversité de microalgues, majoritairement des dinoflagellés (Gast et Caron, 2001), mais aussi des haptophytes (Anderson, 1983), des prasinophytes (Anderson, 1976) et des cyanobactéries (Anderson et Matsuoja, 1992; Yuasa et al., 2012) (Fig. 5).

Grâce à l'utilisation de plusieurs marqueurs phylogénétiques hautement résolutifs, l'identité des symbiotes des invertébrés récifaux est aujourd'hui claire et sans ambigüité, allant jusqu'à l'espèce voire au génotype (LaJeunesse, 2001). Leur diversité géographique est également assez bien connue car la plupart des grands récifs coralliens des océans ont été étudiés. A contrario, nos connaissances sur l'identité et la diversité des microalgues symbiotiques dans le milieu océanique restent rudimentaires. Leur identification s'est largement basée sur l'ultrastructure *in hospite* observée au microscope électronique, mais s'est avérée peu précise car la morphologie de la microalgue en état symbiotique change considérablement. Les critères diagnostiques, comme les flagelles, le squelette et certains organites, sont grandement altérés et peuvent même disparaitre. La mise en culture de l'algue symbiotique après microdissection de l'hôte est une alternative pour l'identification taxonomique car les caractères morphologiques réapparaissent généralement en phase libre. Cependant, comme c'est le cas pour la grande

majorité des microorganismes marins, les chances d'obtenir une culture sont généralement très minces. L'identité des symbiotes impliqués dans les photosymbioses pélagiques reste ainsi à des niveaux taxonomiques faibles, ce qui prévient toutes hypothèses sur leur évolution, leur écologie, et le fonctionnement de la symbiose en général.

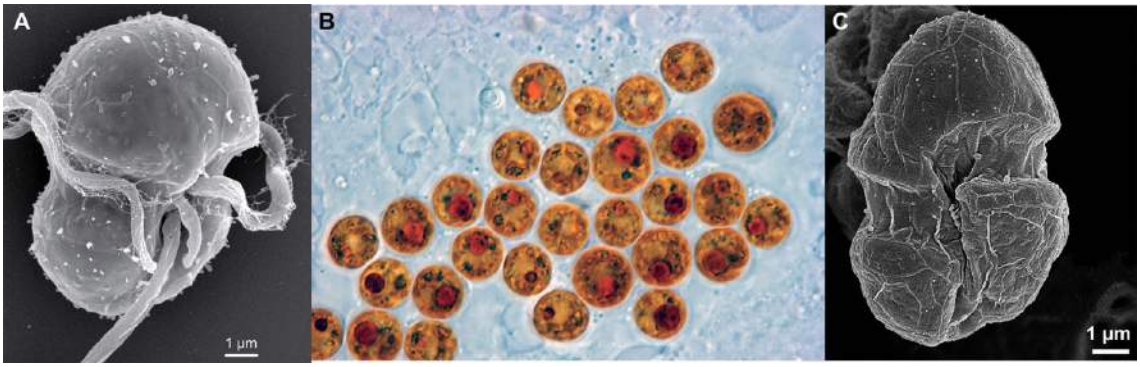


Figure 4. Microalgues dinoflagellés symbiotiques : *Symbiodinium* en culture (A) et dans un hôte (B, chez la méduse *Cassiopea xamachana*), et *Pelagodinium béii* en culture (Siano et al., 2010). **Photos de Gert Hansen (A), de T.C. LaJeunesse (B) et de R. Siano (C).**

3.4 Spécificité de la relation entre partenaires photosymbiotiques

Contrairement aux ectosymbioses, qui sont le plus souvent des relations opportunistes très transitoires, les endosymbioses avec des microalgues montrent un niveau de régulation plus complexe et une certaine spécificité. L'hôte établit généralement une relation spécifique avec un genre de microalgue (ex: *Symbiodinium*, *Pelagodinium*, Fig. 5). Mais la relation peut présenter une certaine flexibilité en dessous de ce niveau taxonomique où, selon les conditions environnementales et la zone géographique, les espèces et/ou génotypes des microalgues peuvent changer à l'intérieur de l'hôte (Shaked et de Vargas, 2006; Little et al., 2004).

En ce qui concerne la microalgue, une espèce ou un génotype peut se retrouver en symbiose avec une multitude d'hôtes phylogénétiquement très éloignés. Par exemple, les dinoflagellés *Scrippsiella nutricula* peuvent être symbiotes de radiolaires mais aussi de méduses (Caron, 2000), et le dinoflagellé *Pelagodinium béii* est en symbiose avec au moins quatre espèces de foraminifères pélagiques (Shaked et de Vargas, 2006). Les clades de *Symbiodinium* sont symbiotiques chez de nombreux invertébrés et protistes récifaux (foraminifères, acœles, méduses, cnidaires, éponges etc.; Pochon et al., 2006). Le «partage» de certaines microalgues entre les hôtes semble donc être un mode répandu. Selon certains auteurs, ces microalgues généralistes auraient une compétence symbiotique qui les prédispose à vivre en symbiose dans un hôte (Gast et Caron, 2001), mais l'état actuel de nos connaissances ne permet pas de définir les raisons biologiques sous-jacentes. Enfin, la présence de ces symbiotes généralistes et la flexibilité de la relation expliquerait qu'aucune coévolution stricte, où les spéciations des partenaires symbiotiques sont simultanées et parallèles, n'ait été mise en évidence à ce jour dans les photosymbioses (Shaked et de Vargas, 2006).

3.5 Ecologie des microalgues en dehors de leur hôte

L'écologie des microalgues symbiotiques en phase libre est très peu connue mais pourtant primordiale pour comprendre l'évolution et les mécanismes de la relation (Douglas, 1998; LaJeunesse, 2001; Knowlton et Rohwer, 2003). Dans l'écosystème récifal, les microalgues sont en très grandes densités dans les tissues des invertébrés benthiques, atteignant plusieurs millions de cellules par cm^2 (Stimson et al., 2001), mais sont relativement rares dans l'environnement. La détection de la forme libre, qui passe par un effort important d'échantillonnage et des outils de biologie moléculaire spécifiques, a montré que les *Symbiodinium* sont plus nombreux dans le sédiment (1000–2000 cells/mL) que dans la colonne d'eau (< 20 cells/mL) (Littman et al., 2008; Coffroth et al., 2006). Une population symbiotique restreinte en dehors de l'hôte est également observée dans les symbioses terrestres et semble donc être une configuration commune (Ahmadjian, 1988). Cela s'expliquerait par le développement d'une dépendance vis-à-vis de l'hôte, critique pour la survie du symbiote (ex: dépendance métabolique). Selon le type de photosymbiose, cette dépendance varierait selon un continuum, de facultative à obligatoire. On ignore si ce continuum reflète l'histoire macroévolutive de l'association ou bien les conditions environnementales de l'habitat.

Dans le milieu océanique, le peu d'information disponible à ce jour sur la taxonomie et la diversité des microalgues symbiotiques n'a pas permis de mettre en évidence leur abondance et leur distribution en phase libre. Par exemple, la population libre de *Pelagodinium*, symbiotes de certaines espèces de foraminifères pélagiques, est totalement inconnue. Leur degré de dépendance vis-à-vis de la symbiose reste par conséquent énigmatique. Est-ce que la symbiose est obligatoire pour les microalgues dans les régions océaniques? Est-ce que la microalgue augmente sa fitness en entrant en symbiose avec un hôte ?

Etant donné le caractère transitoire, mais obligatoire, de la relation pour les hôtes comme les radiolaires et les foraminifères, les microalgues devraient être suffisamment abondantes dans l'environnement pour permettre une acquisition rapide et opportune au cours du cycle de vie de l'hôte. De plus, ces hôtes unicellulaires ont des temps de génération extrêmement rapides par rapport aux invertébrés récifaux, ce qui impose un «*reset*» complet et fréquent de l'association. La transmission verticale permettrait de faire face à ces conditions non propices à la rencontre et à l'établissement d'interactions étroites (Douglas, 1998), mais ce n'est pas le mode adopté par les radiolaires et les foraminifères pélagiques. Tout ceci mène à penser qu'il y aurait des mécanismes spécifiques pour optimiser la rencontre hôte-symbiotes dans l'immensité du milieu océanique, turbulent et plutôt pauvre en organismes. Pourquoi et comment la photosymbiose horizontale s'est maintenue au cours de l'évolution et n'a pas évolué vers une transmission verticale? Notre compréhension sur l'évolution et le mécanisme de la photosymbiose planctonique passe inévitablement par l'étude de l'écologie des microalgues symbiotiques.

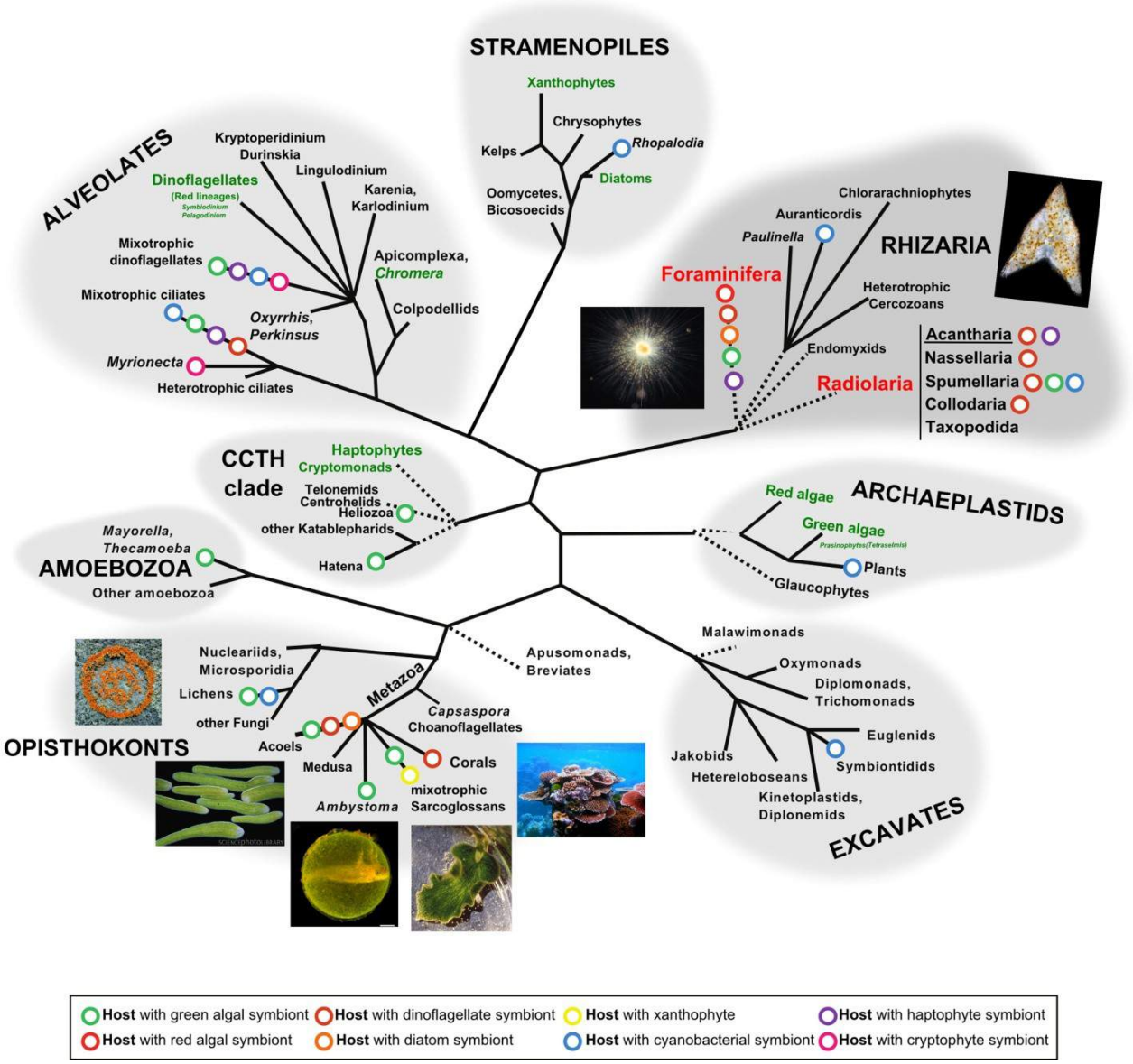


Figure 5. Arbre phylogénétique des eucaryotes indiquant les hôtes en photosymbiose qui abritent différent types de microalgues. Les taxa de microalgues symbiotiques sont indiqués en vert. Ce travail de thèse s'intéressera plus particulièrement aux photosymbioses océaniques rencontrées chez certains protistes comme les foraminifères et les radiolaires appartenant au super-groupe des Rhizaria, qui peuvent vivre avec des dinoflagéllés (Alvéolés) ou des haptophytes (clade CCTH) (modifiée d'après Dorrell et Howe, 2012).

3.6 Choix du modèle d'étude

Parmi les protistes pélagiques photosymbiotiques, les acanthaires (radiolaires) sont ubiquitaires et généralement plus abondants que les autres radiolaires et foraminifères à la surface des grandes régions océaniques. La photosymbiose chez les acanthaires représente clairement un succès écologique et fait de ce taxon un modèle de choix pour étudier les mécanismes de cette relation chez les grandes cellules du plancton. De plus, le fait que certaines espèces d'acanthaires ne vivent pas en photosymbiose permet de comparer l'effet de ce mode de vie sur la morphologie, l'évolution, l'écologie et le cycle de vie de ces organismes, et ce à l'échelle des océans planétaires. Enfin, malgré ses nombreuses particularités étudiées dès le 19ème siècle (Fig. 6), la biologie des acanthaires reste paradoxalement méconnue et son exploration laisse présager des découvertes importantes pour mieux appréhender leur évolution et leur écologie dans les océans. Les acanthaires constituent le modèle photosymbiotique de cette étude.

« *Allons au fond de l'inconnu pour trouver du nouveau* » Charles Baudelaire.

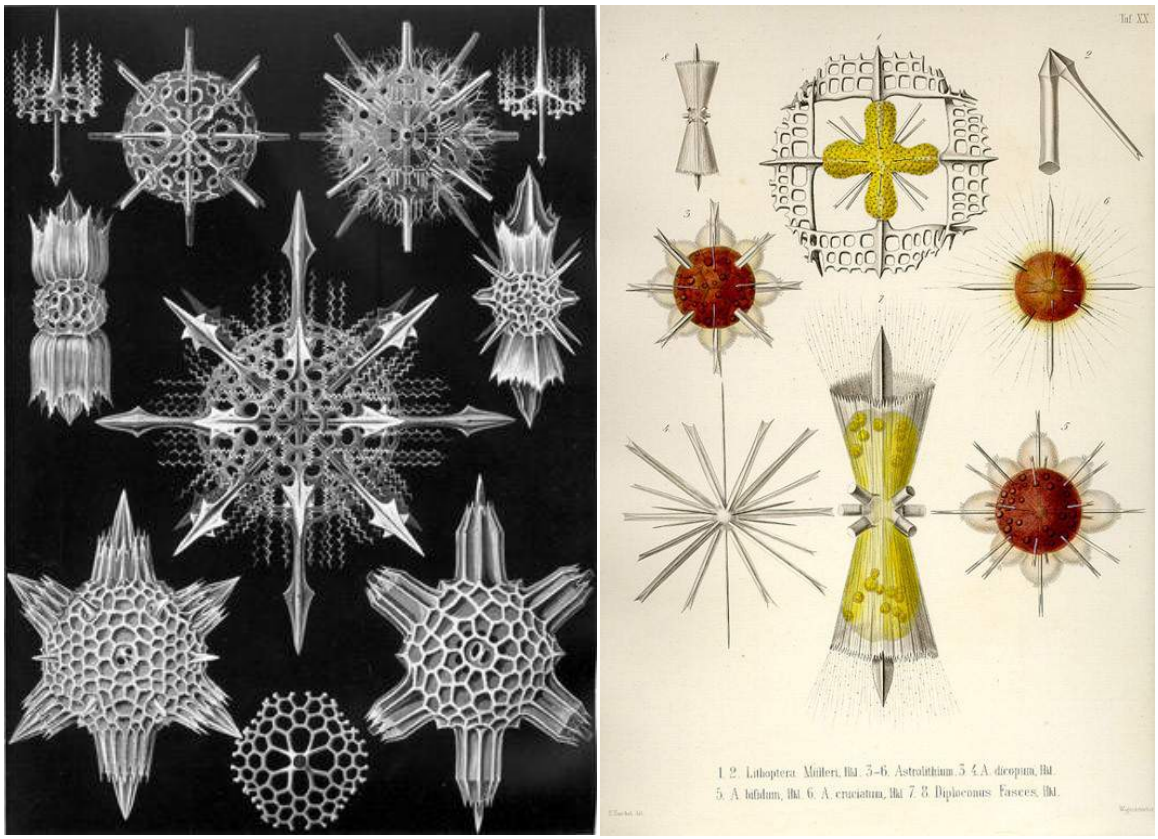


Figure 6. Planches d'acanthaires dessinées par Ernst Haeckel.

4. Les acanthaires

Dans la mythologie grecque, Acanthe (du grec ancien ἄκανθα, « épine ») est une nymphe. Désirée par Apollon (dieu du Soleil), qui tente de l'enlever, elle lui griffe le visage. Pour se venger, le dieu la métamorphose en une plante épineuse qui aime le soleil, et qui porte depuis son nom : l'acanthe.

4.1 Le squelette et la cellule

Le squelette des acanthaires est unique dans le sens où il est interne à la cellule (Fig. 7C), et est constitué de sulfate de strontium (celestite). Ce sont les seuls organismes marins connus à ce jour capables de minéraliser spécifiquement le strontium et d'en faire des structures très complexes. Le rôle du strontium, élément très proche du calcium, est encore méconnu, mais il serait impliqué dans les processus de biominéralisation chez les gastéropodes, coraux mais aussi chez l'homme (Hanlon et al., 1989; Nielsen, 2004). Le squelette des acanthaires est constitué de dix ou vingt spicules, chacun représentant un seul cristal de celestite, qui s'agencent selon une organisation particulière (la loi de Müller; Haeckel, 1862). L'extrémité des spicules émergent à des points cardinaux précis, telle une terre virtuelle : 8 extrémités polaires, 8 extrémités tropicales et 4 extrémités équatoriales (Fig. 7A).

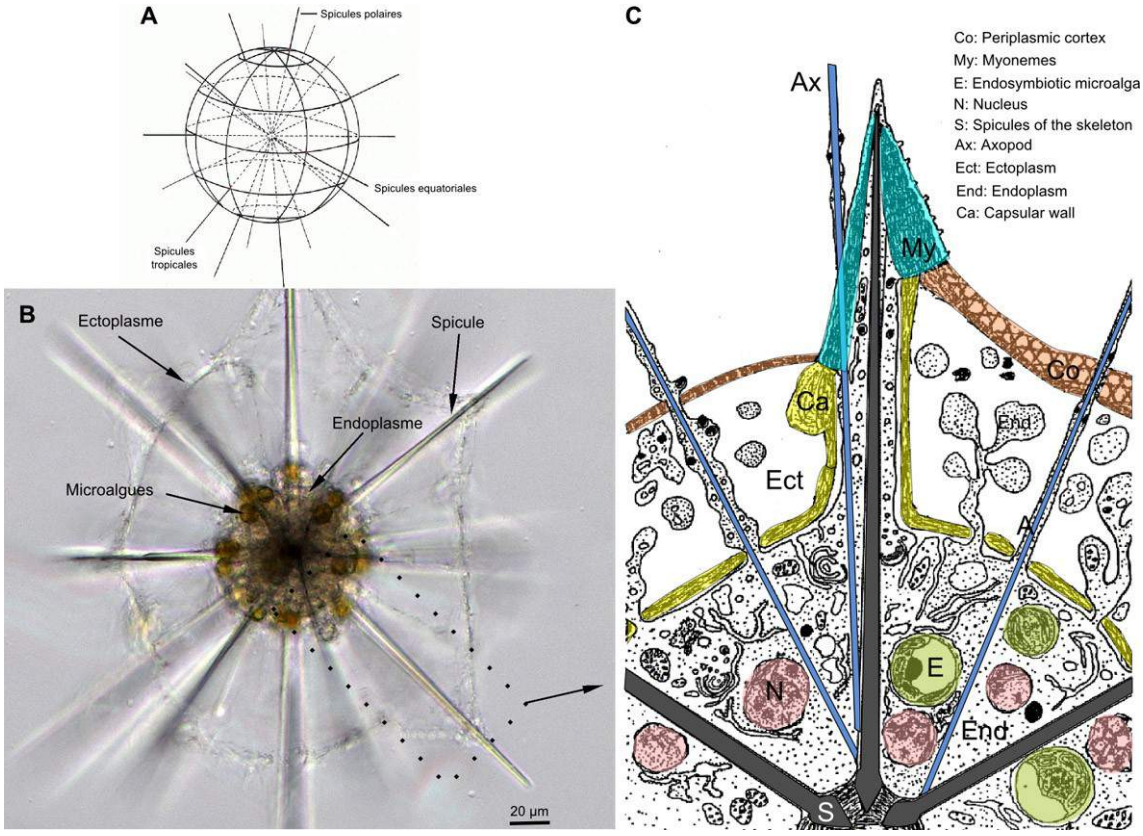


Figure 7. Organisation du squelette d'acanthaire (agencement des spicules selon la loi de Müller, A), et de la cellule contenant des microalgues symbiotiques dans l'endoplasme (B et C). C) Détail de l'ultrastructure cellulaire autour d'un spicule.

L'originalité et la complexité de la cellule des acanthaires attirent les biologistes cellulaires durant la seconde moitié du 20ème siècle. La cellule amiboïde est divisée en deux compartiments principaux: l'ectoplasme et l'endoplasme, délimités respectivement par le cortex periplasmique et la paroi capsulaire (Fig. 7B, C). L'ectoplasme est une extension cytoplasmique d'échange avec l'extérieur, principalement impliqué pour la capture et la digestion des proies. L'endoplasme est le cytoplasme vrai où se trouvent la plupart des organites cellulaires comme les noyaux, les mitochondries et l'appareil de golgi (Febvre et al., 2000). Contrairement aux autres radiolaires et foraminifères, c'est dans l'endoplasme que se trouvent les microalgues symbiotiques des acanthaires**.** Le cortex periplasmique est attaché aux extrémités de chacun des spicules du squelette par des filaments contractiles de 8 à 60 µm de long, appelés myonèmes (Fig. 7C). Selon l'espèce, il y a entre 2 et 60 myonèmes autour de chaque spicule (Febvre, 1981; Febvre et al., 2000). Tels des élastiques qui tendent une toile de tente, les myonèmes se contractent et se relâchent rapidement, variant ainsi la surface cellulaire. Cette activité est contrôlée en partie par le calcium et permettrait à l'acanthaire de moduler sa portance, et donc sa flottabilité dans la colonne d'eau (Febvre, 1981; Febvre et Febvre-Chevalier, 1989). Des axopodes, assemblages de microtubules communs chez les radiolaires et les héliozoaires, sont principalement responsables de la capture des proies, mais également du transport de vésicules et divers organites entre différents compartiment cellulaires (Febvre-Chevalier et Febvre, 1986).

4.2 Taxonomie et évolution

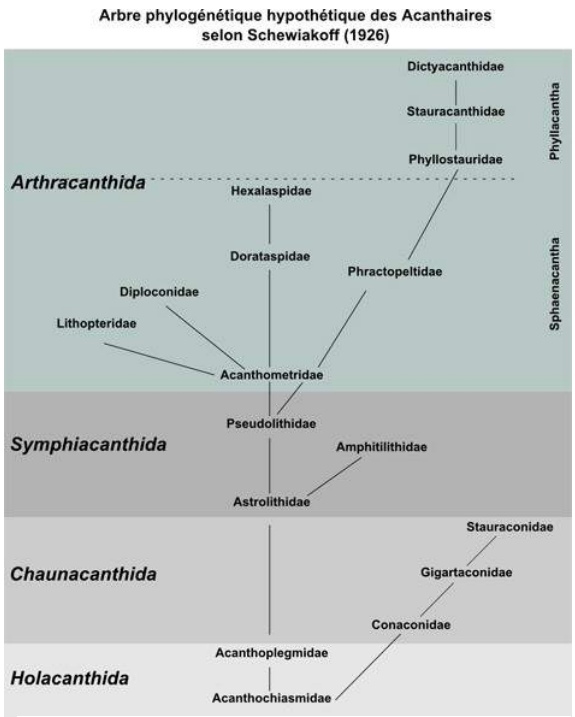


Figure 8. Evolution et schéma taxonomique des acanthaires selon Schewiakoff (1926).

Ernst Haeckel dresse le premier système taxonomique des acanthaires en 1887 dans son ouvrage « *Die Radiolarien* », où environ 372 espèces y sont décrites (Fig. 6). D'abord appelé Acanthometren par son professeur Johannes Müller en 1858, Haeckel crée le nom Acantharia en 1881. Cette première classification verra quelques modifications par la suite par Mielck (1907) et Popofsky (1904a, b, 1906). Les caractères diagnostiques utilisés furent principalement des éléments du squelette comme la longueur, la forme et la manière dont se joignent les spicules au milieu de la cellule (jonction centrale). Wladimir Schewiakoff de l'Université de St Petersburg consacre les trois dernières années de sa carrière à l'étude des acanthaires à la Station Zoologique de Naples. Contrairement à ces prédécesseurs, Schewiakoff travaille principalement sur des cellules vivantes,

ce qui va lui permettre de définir de nouveaux caractères taxonomiques (myonèmes, capsule centrale etc.), mais aussi de connaitre un peu plus l'écologie et la physiologie de ces organismes. Présenté dans une monographie parue en 1926, son nouveau système taxonomique comprend 130 espèces. Cette classification exclusivement basée sur la morphologie des acanthaires fait encore autorité aujourd'hui et très peu de modifications ont été apportées (Trégouboff, 1953; Reshetnyak, 1981; Bernstein et al., 1999; Febvre et al., 2000). Actuellement, la classe des acanthaires comprend environ 50 genres et 150 espèces, distribués dans 18 familles et 4 ordres: Holacanthida, Chaunacanthida, Symphiacanthida et Arthracanthida (Fig. 8). Schewiakoff considérait les deux premiers ordres comme des formes « primitives » par rapport aux deux derniers. Contrairement aux autres radiolaires et aux foraminifères qui présentent un bilan fossile complet, utilisé comme archive stratigraphique, les acanthaires ne fossilisent pas car leur squelette de celestite se dissout immédiatement dès la mort de l'organisme (Beers et Stewart, 1970). L'évolution des acanthaires reste par conséquent énigmatique.

L'identification morphologique des protistes non cultivables comme les acanthaires reste une tâche très ardue. Les caractères morphologiques sont peu nombreux et peuvent demander des préparations longues et délicates sous le microscope pour les mettre en évidence. De plus, certains de ces caractères peuvent se détruire durant la collecte et/ou changer drastiquement au cours de l'ontogénèse de l'organisme. Ces dernières années, la phylogénie moléculaire appliquée à certains groupes de protistes a mise en évidence de nombreux conflits avec la classification traditionnelle, révélant notamment la nature polyphylétique des différents niveaux taxonomiques (Gontcharov et Melkonian, 2000; Cavalier-Smith et von der Heyden, 2007; Bachy et al., 2012). Une phylogénie moléculaire des acanthaires s'impose comme une première étape nécessaire non seulement pour revisiter la taxonomie de ces organismes, mais aussi pour permettre d'élucider l'histoire évolutive de la morphologie et des modes de vie comme la symbiose.

4.3 La photosymbiose

A la surface des régions océaniques, les acanthaires possédants des microalgues symbiotiques représentent 80% de la biomasse des acanthaires (Stoecker et al., 1996). Ces acanthaires peuvent intervervenir dans la production primaire océanique à hauteur de 20% dans les cas les plus extrêmes (Michaels, 1988). Taylor (1982) observait dans la mer des Sargasses que la plupart de la chlorophylle au dessus de la couche de chlorophylle maximum (DCM) pourrait être représentée par les acanthaires, après les microalgues pico-planctoniques et les efflorescences de cyanobactéries filamenteuses. Selon l'espèce, les acanthaires à l'état adulte possèdent de 10 à 100 cellules symbiotiques dans leur cellule. Le nombre de microalgues augmente au cours de l'ontogénèse de l'acanthaire (Michaels, 1991), probablement avec la taille du cytoplasme. Il n'a pas été clairement montré si la microalgue continue de se diviser dans l'hôte, et/ou s'il y a une acquisition récurrente au cours du cycle de vie de l'acanthaire.

Il y a plus de 30 ans, basée sur des critères diagnostiques comme l'organisation du plaste, l'observation en microscopie électronique *in hospite* a indiqué que les symbiotes des acanthaires appartiennent à deux grandes lignées de microalgues: les haptophytes et les dinoflagellés

(Hollande et Carré, 1974; Febvre et Febvre-Chevalier, 1979). En tout et pour tout, les haptophytes ont été observés dans trois espèces d'acanthaires (*Lithoptera muelleri*, *Acanthometra pellucida* et *Amphilonche elongata*) appartenant aux ordres des Arthracanthida et Symphiacanthida qui en contiennent plus de 100 environs. Quant aux dinoflagellés, leur hôte n'a pas été déterminé dans l'étude concernée. Basée uniquement sur ces observations d'ultrastructure, l'identité des microalgues symbiotiques des acanthaires reste imprécise. A ce jour, aucune séquence génétique n'est venue apporter une identification plus précise au niveau du genre ou de l'espèce. De plus, avec seulement trois espèces d'acanthaires venant d'une même localité (Mer Méditerranée, Villefranche-sur-Mer), la spécificité et la biogéographie de la relation ne peuvent être appréciées correctement. D'un point de vue macroévolutif, l'absence de bilan fossile et de phylogénies moléculaires n'a pas permis de mettre en évidence si la symbiose est apparue une ou plusieurs fois dans l'histoire des acanthaires.

4.4 Ecologie et biogéographie

L'étude de l'écologie des acanthaires a été entravée par des biais techniques lors de la collecte et la préservation des cellules. Ce sont des organismes fragiles qui résistent plus ou moins bien aux traits de filets de plancton, et surtout peuvent «disparaître» dans les échantillons car le squelette se dissout très rapidement dans les produits de fixation et de préservation classiquement utilisés en biologie marine (Bernstein et Stewart, 1970). C'est une des raisons pour laquelle les acanthaires, malgré leur abondance dans les communautés planctoniques, ont été mésestimés. Reposant sur relativement peu d'études, notre connaissance sur la distribution biogéographique des acanthaires est parcellaire (Fig. 9), et ne prend pas en considération les différents niveaux taxonomiques. Ces protistes ubiquitaires sont particulièrement abondants dans les eaux tropicales et subtropicales. Typiquement plus nombreux que les autres radiolaires et les foraminifères planctoniques, les abondances rapportées à ce jour varient de 1 à 40 cellules par litre selon la région océanique (Beers et al., 1975; Bottazzi, 1978; Michaels, 1988; Michaels et al., 1995). De véritables efflorescences ont été décrites dans la Mer Méditerranée et la Mer des Sargasses où la densité peut atteindre 500 000 individus par m^2 (Massera-Bottazzi et Andreoli, 1978, 1981; Febvre, 1990). Même si la concentration d'acanthaires n'est en aucun cas comparable aux efflorescences de microalgues, ce phénomène est considéré dans la littérature comme un «*bloom*», où les acanthaires dominent largement le microzooplancton. Quel est la nature des liens entre la relation photosymbiotique et l'écologie des acanthaires? Si les acanthaires symbiotiques vivent avec le même type de microalgue, leur distribution géographique devrait être bornée à celle de leur partenaire photosynthétique. L'étude de l'écologie des acanthaires symbiotiques doit donc reposer également sur l'écologie de sa microalgue en phase libre.

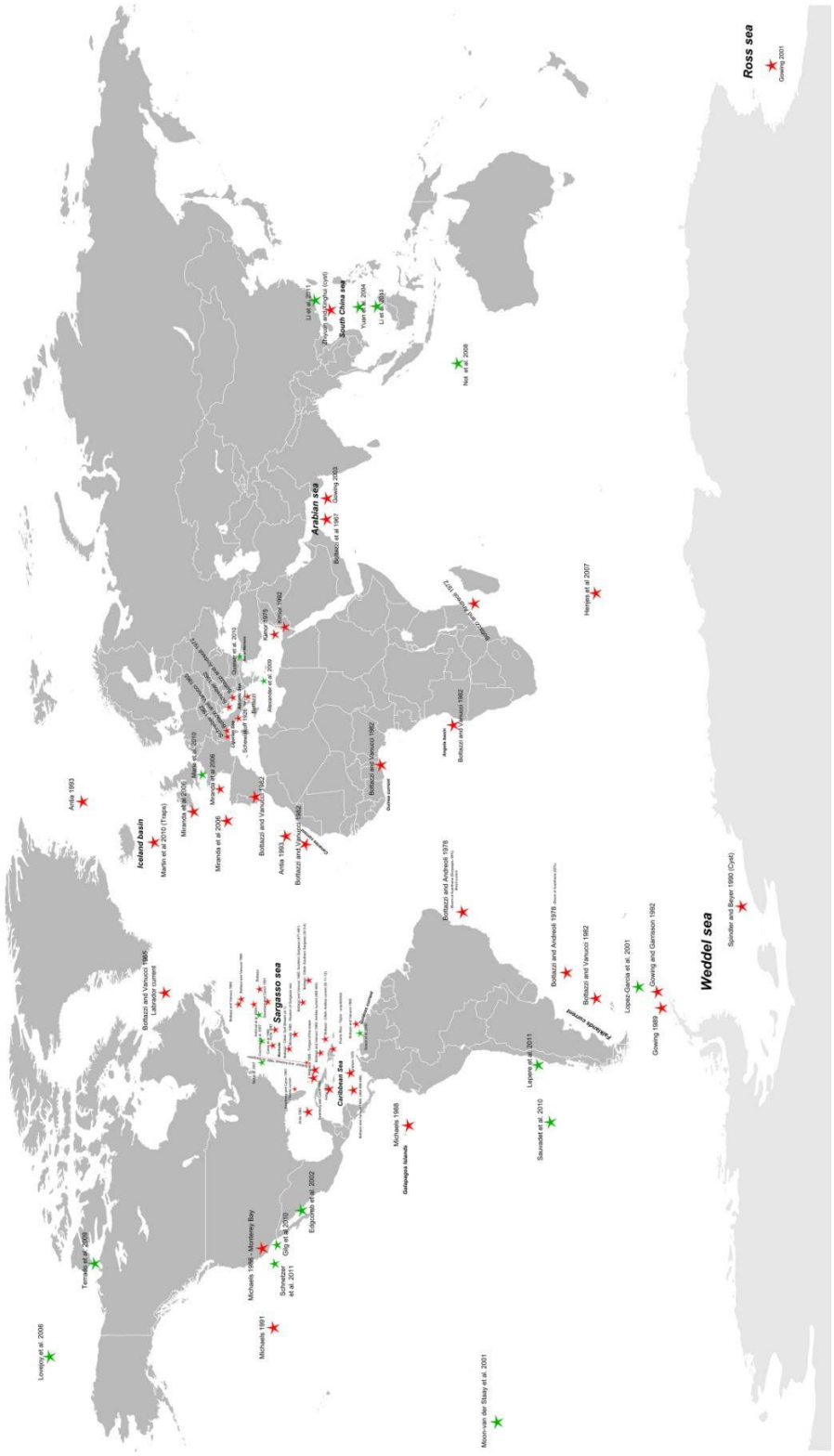


Figure 9. Biogéographie des acanthaires connue à ce jour. Les étoiles rouges représentent des observations sous le microscope et les étoiles vertes indiquent la présence de séquences génétiques environnementales affiliées aux acanthaires obtenues avant ce travail de thèse.

La photosymbiose semble influencer la distribution verticale des acanthaires dans la colonne d'eau. Les acanthaires symbiotiques, photosynthèse oblige, se situent plutôt dans la partie supérieure de la colonne d'eau (zone photique), tandis que les formes non symbiotiques peuvent se trouver plus en profondeur jusqu'à 1000-2000 mètres (Schewiakoff, 1926; Taylor 1982). Ces observations, principalement faites par Schewiakoff en Mer Méditerranée, doivent être confirmées et comparées dans plusieurs zones

Ces dernières années, l'arrivée de la biologie moléculaire a révolutionné la protistologie marine. De nombreuses séquences affiliées aux acanthaires (gène codant pour la petite sous-unité du ribosome, ADNr 18S) ont été obtenues dans une multitude d'études de diversité environnementale par librairies de clones, à des latitudes et profondeurs différentes : en milieu côtier (Marie et al., 2010), en profondeur (Not et al., 2007; Quaiser et al., 2010), en Antarctique et Arctique (López-García et al., 2001; Lovejoy et al., 2006), dans les eaux anoxiques profondes (Alexander et al., 2009; Stoeck et al., 2003), jusqu'aux sédiments des sources hydrothermales (López-García et al., 2003; Edgcomb et al., 2002). Une grande partie de ces séquences a été retrouvée dans des fractions de taille pico-planctoniques (0.8 à 5 µm), une gamme de taille qui ne correspond pas aux grandes cellules d'acanthaires à l'état adulte (20 à 500 µm). Ces résultats soulèvent beaucoup de questions sur l'écologie et le cycle de vie des acanthaires. L'avènement du séquençage à haut débit constitue une avancée significative pour les biologistes qui cherchent à explorer de manière rapide et exhaustive la communauté de protistes, pour la majorité incultivables. Mais à l'instar des clones environnementaux, l'interprétation pertinente de ces données massives de séquences d'ADN marqueurs de diversité (*metabarcoding*) passe nécessairement par l'obtention de séquences références, qui est l'unique moyen de faire le lien entre la séquence et le phénotype. Le faible nombre de séquences références disponibles pour les acanthaires (environ 10 dans GenBank avant cette thèse acquises depuis Amaral-Zettler et al., 1997) empêche toutes interprétations sur l'écologie et la distribution géographique des acanthaires. Les séquences environnementales assignées aux acanthaires forment une phylogénie comprenant des clades moléculaires dont la morphologie et le mode de vie sont inconnus (Not et al., 2007).

4.5 Cycle de vie et rôle biogéochimique

Nos connaissances sur la reproduction de ces organismes non cultivables restent extrêmement floues. La reproduction asexuée a été observée très rarement, et selon Schewiakoff (1926) elle serait spécifique à une seule famille (Acanthochiasmidiae) de l'ordre des Holacanthida. En revanche, la reproduction sexuée avec gamétogénèse a été observée plus fréquemment et se déroule selon deux processus distincts : soit directement à partir de la forme adulte (végétative, Fig. 10), ou soit *via* un enkystement préalable de la cellule (Fig. 11; Hollande et Enjumet, 1957; Hollande et al., 1965; Massera Bottazzi, 1966, 1973). Lorsque la gamétogénèse ou sporogénèse se déclenche, le nombre de noyaux dans la cellule augmente rapidement par une série de divisions, puis chaque noyau s'individualise pour former une spore ou «*swarmer*» (Fig. 10 et 11), cellule biflagellée de moins de 5 µm (Febvre, 1977). Des dizaines

de milliers de ces spores sont libérées dans le milieu après déchirures des enveloppes cytoplasmiques, soit en profondeur pour les acanthaires qui s'enkystent, soit plus en surface pour les acanthaires à gamétogénèse direct. Ces spores correspondraient probablement aux gamètes sexuels, mais leur ploïdie n'a jamais été étudiée.

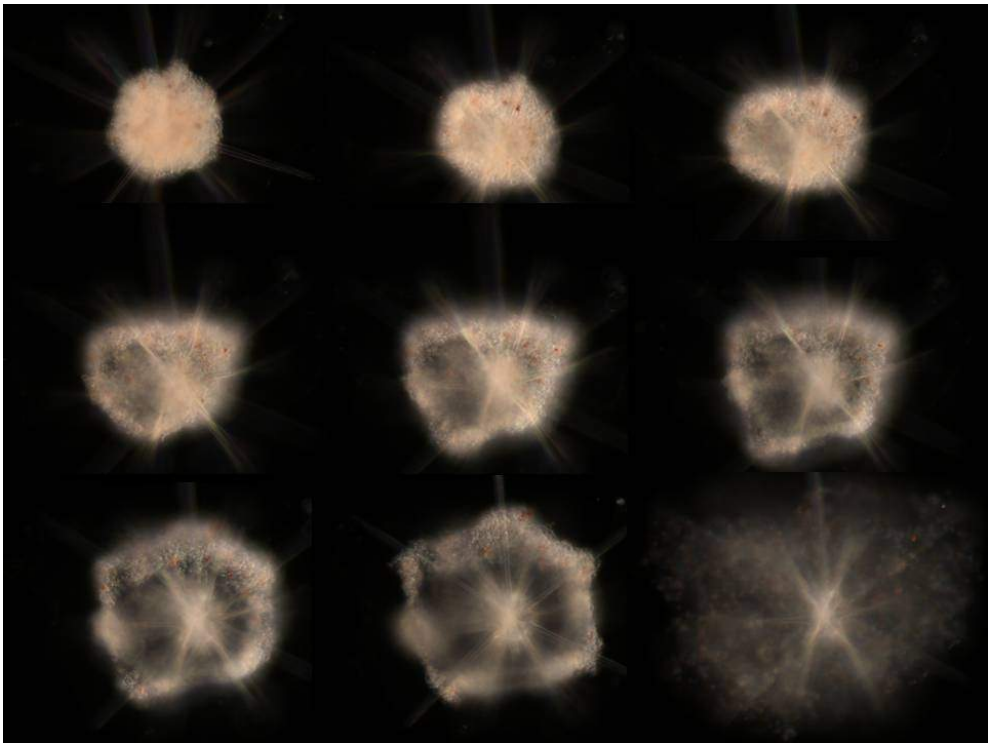


Figure 10. Processus de la gamétogénèse « directe » de la forme végétative, où un acanthaire adulte en fin de cycle de vie relargue des dizaines de milliers de spores dans l'environement (photos J. Decelle).

Certains acanthaires sont capable de changer drastiquement leur morphologie au cours du cycle de vie, pour former un kyste. Ce kyste ressemble à une graine dense mesurant de 200 à 1000 µm et pouvant prendre des formes très différentes (Fig. 11). Le kyste d'acanthaire ne serait pas une forme de dormance pour résister aux conditions environnementales non favorables comme chez beaucoup de protistes (Smetacek, 1985; Kremp et al., 2009), mais constitue une stratégie active de reproduction (Schewiakoff, 1926; Hollande et al., 1965; Masserra Bottazzi 1973). Due à sa grande densité, le kyste plonge rapidement en profondeur avec une vitesse de sédimentation pouvant atteindre 200 mètres par jour et peut parvenir dans le domaine mesopélagique (200-1000 m) et bathypélagique (>1000 m) (Antia et al., 1993). L'enkystement des acanthaires représente donc un flux vertical non négligeable de carbone et de strontium, qui participe à la séquestration à long terme de ces éléments dans les couches profondes (Bernstein et al., 1987; Antia et al., 1993; Michaels et al., 1995). Il a été démontré que la concentration du strontium le long d'un gradient verticale est influencée par la présence des acanthaires (De Deckker, 2004; Bernstein et al., 1987): elle est moins importante en surface où évoluent

principalement les acanthaires qui pompent l'élément, et augmente en profondeur *via* la dissolution des squelettes.

L'enkystement des acanthaires, qui à notre connaissance est un mode de reproduciton unique chez les protistes, soulève de nombreuses questions quant à son rôle et à son influence sur la biogéochimie du carbone et du strontium. La formation de kystes serait propre aux ordres des Holacanthida et Chauncanthida tandis que la reproduction directe de la forme végétative se retrouverait chez les Arthracanthida et les Symphiacanthida, dont certaines espèces sont photosymbiotiques (Febvre et al., 2000). Cette dichotomie du mode de reproduction reste à confirmer car, les critères disparaissant totalement lors du passage entre la forme adulte et le kyste, il n'y a aucun moyen d'identifier taxonomiquement un kyste. De plus, le processus d'enkystement est extrêmement rare à observé sous le microscope après une collecte. Est-ce que les acanthaires photosymbiotiques ont maintenu une partie de leur cycle de vie dans les eaux profondes ?

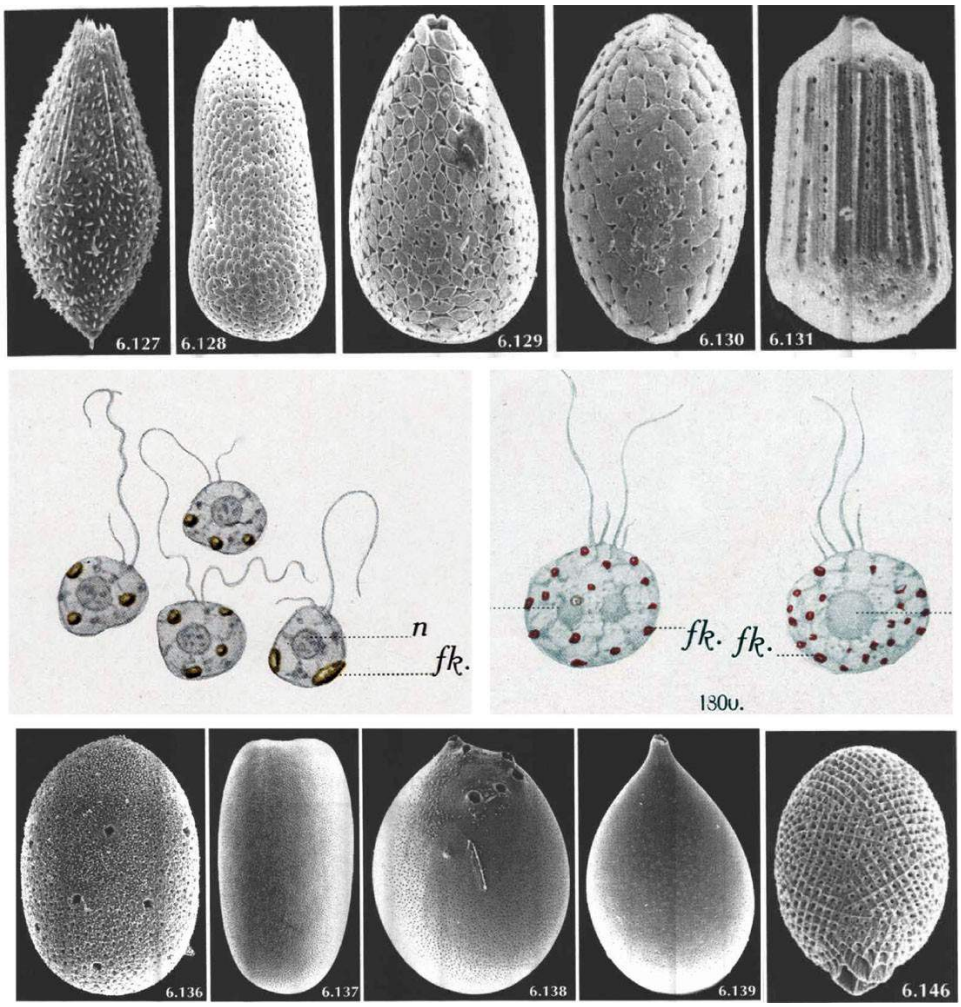


Figure 11. Kystes d'acanthaires observés au microscope électronique à balayage (en haut et en bas ; Bernstein et al., 2004) et dessins de Schewiakoff (1926) des «*swarmers*» d'acanthaires (au milieu).

.

Objectifs de la thèse

- Afin de mettre en évidence l'écologie et l'évolution de la photosymbiose chez les acanthaires, le premier objectif de ce travail sera de définir un cadre morpho-génétique de réference. Une phylogénie moléculaire à partir de cellules uniques, photographiées et identifiées sur la base de leurs caractères morphologiques, sera établie pour faire le lien entre la position phylogénétique, la morphologie et le mode de vie symbiotique (**chapitre 1)**.

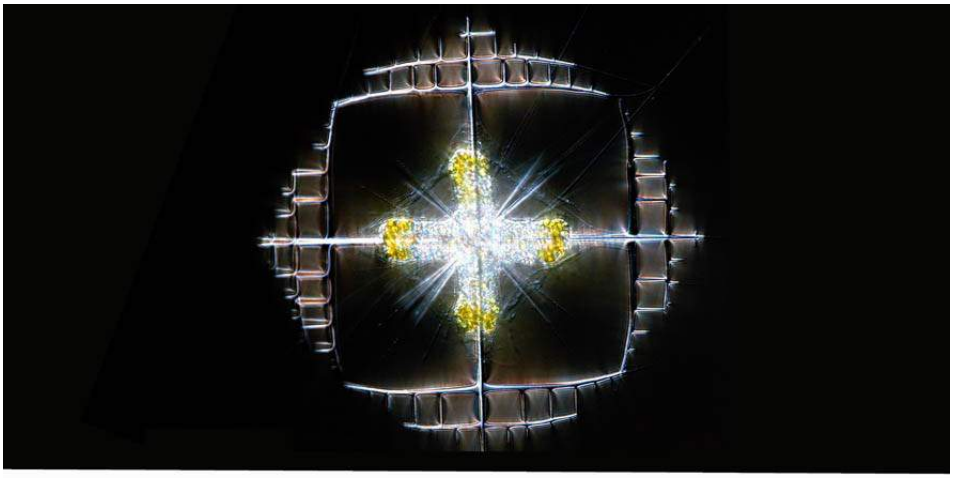
- Dans un deuxième temps, cette thèse se penchera sur l'identité des microalgues symbiotiques des photosymbioses observées, ainsi que sur la spécificité de la relation selon la zone géographique et la taxonomie de l'hôte. Cette étape permettra de voir si les microalgues des acanthaires sont déjà connues pour être en photosymbiose avec d'autres hôtes ou représentent plutôt un nouveau groupe d'algues symbiotiques. Pour mettre la symbiose dans un cadre macroévolutif, une datation moléculaire tentera de définir quand et dans quel contexte environnemental la relation symbiotique chez les acanthaires s'est établie (**chapitre 2**).

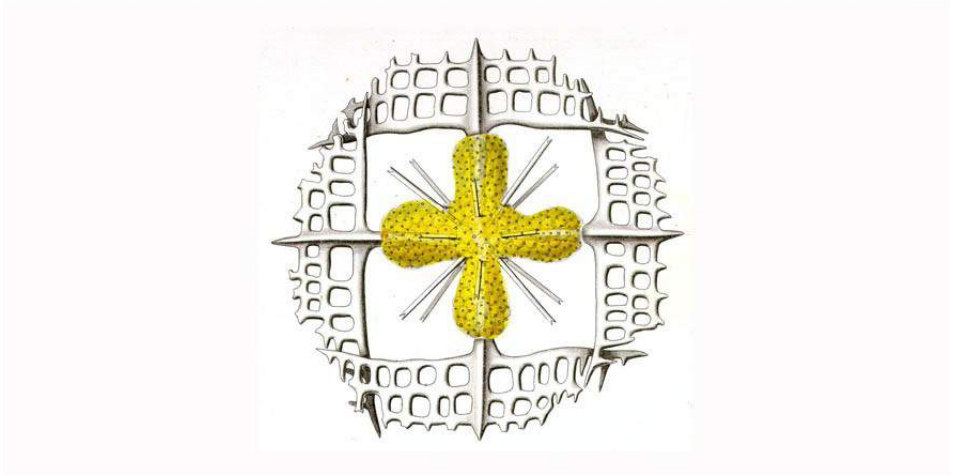
- Afin d'examiner l'influence de la photosymbiose sur l'écologie des acanthaires, une approche moléculaire de type « *metabarcoding* » sera adoptée à l'échelle globale des océans. Cette étape s'appuiera sur les travaux précédents, notamment la base morpho-génétique de référence de l'hôte et l'identité génétique des microalgues symbiotiques. L'abondance relative et la distribution géographique des acanthaires symbiotiques et non symbiotiques pourront être comparées, ainsi que celles des microalgues. Cette partie tentera d'élucider les mécanismes d'acquisition et de transmission des symbiotes à l'échelle de l'écosystème dans le milieu pélagique (**chapitre 3**).

- Enfin, l'impact de la symbiose sur le cycle de vie des acanthaires sera exploré et permettra de d'identifier plus précisément le rôle biogéochimique respectif des acanthaires symbiotiques et non symbiotiques. La distribution verticale des différents groupes d'acanthaires dans la colonne d'eau sera également examinée entre la zone photique et la zone mesopélagique (**chapitre 4**).

Chapitre I

Molecular phylogeny and evolutionary history of the skeleton of the Acantharia (Radiolaria)





Molecular Phylogeny and Morphological Evolution of the Acantharia (Radiolaria)

Johan Decelle[a], Noritoshi Suzuki[b], Fredéric Mahé[a], Colomban de Vargas[a], and Fabrice Not[a]

[a]CNRS, UMR 7144 & Université Pierre et Marie Curie Paris 6, Station Biologique de Roscoff, Equipe Evolution du Plancton et PaléoOcéans, Place Georges Teissier, 29682 Roscoff, France

[b] Institute of Geology and Paleontology, Graduate School of Science, Tohoku University, 6–3Aoba, Aramaki, Aoba-ku, Sendai, 980–8578 Japan

Published in ***Protist,*** **2012 vol.163: 435-450**

Abstract

Acantharia are ubiquitous and abundant rhizarian protists in the world ocean. The skeleton made of strontium sulphate and the fact that certain harbour microalgal endosymbionts make them key planktonic players for the ecology of marine ecosystems. Based on morphological criteria, the current taxonomy of Acantharia was established by W.T. Schewiakoff in 1926, since when no major revision has been undertaken. Here, we established the first comprehensive molecular phylogeny from single morphologically-identified acantharian cells, isolated from various oceans. Our phylogenetic analyses based on 78 18S rDNA and 107 partial 28S rDNA revealed the existence of 6 main clades, sub-divided into 13 sub-clades. The polyphyletic nature of acantharian families and genera demonstrates the need for revision of the current taxonomy. This molecular phylogeny, which highlights the taxonomic relevance of specific morphological criteria, such as the presence of a shell and the organisation of the central junction, provides a robust phylogenetic framework for future taxonomic emendation. Finally, mapping all the existing environmental sequences available to date from different marine ecosystems onto our reference phylogeny unveiled another 3 clades and improved the understanding of the biogeography and ecology of Acantharia.

Introduction

Acantharia are marine protists taxonomically affiliated to the super-group Rhizaria, the phylum Retaria and to the first rank taxon Radiolaria (Adl et al. 2005; Moreira et al. 2007). Their characteristic star-shaped morphology consists of a skeleton of 10 or 20 spicules made of celestite (strontium sulphate; Odum 1951; Suzuki and Aita, 2011), arranged according to the geometric law of Müller (1859). The skeleton supports a typical rhizarian amoeboid cell whose shape and motility are controlled by-characteristic axopods and myonemes (Febvre et al. 1981). The Acantharia are the only known organisms able to biomineralize strontium sulfate as the principal component of the skeleton.

In surface waters of marine ecosystems, Acantharia consistently outnumber their rhizarian counterparts, such as Foraminifera and Polycystinea (Caron and Swanberg 1990; Michaels et al. 1995; Stoecker et al. 2009). Their abundance tends to increase in oligotrophic waters from equatorial to subtropical latitudes (Massera Bottazzi and Andreoli 1982), and they can form blooms at certain periods of the year, reaching densities up to 500 000 individuals.m^{-2} (Massera Bottazzi and Andreoli 1978, 1981). Acantharian cells have also been found hundreds and even thousands of meters deep in the water column (Antia et al. 1993; Bernstein et al. 1987; Martin et al. 2010). Within the marine food web, the Acantharia are active predators (Swanberg and Caron 1991), and they significantly contribute to carbon flux and biogeochemical cycles of strontium and barium in the oceans (Bernstein 1987; Martin et al. 2010). Furthermore, they indirectly contribute to primary production through endosymbiotic relationships with tens to hundreds of microalgae per cell (Michaels 1988, 1991). Despite holding a key position in marine ecosystems, the Acantharia have been largely overlooked in ecological studies, essentially due to dissolution of the skeleton in classical fixatives (Beers and Stewart 1970) and the lack of success in culturing them. Our knowledge of their biology and diversity is therefore still in its infancy.

The first classification of Acantharia was initiated by Johannes Müller (1856, 1859), and completed by his student Ernst Haeckel (1887, 1888). This classification initially comprised 372 species, and further studies added around 80 more (Mielck 1907; Popofsky 1904a, b, 1906). All of the diagnostic characters used by these authors were exclusively based on the morphology of the skeleton, such as the length, form and central junction of the spicules. Working on living specimens, W.T. Schewiakoff emended this classification in 1926 by taking into account various features of the cell body (e.g. structure and colour of the cytoplasm, absence or presence of the central capsule, myonemes). In his remarkable monograph based on accurate observations of 500 living cells, he confirmed a total of 130 species, and erected the main taxonomic framework for the Acantharia (Schewiakoff 1926). Minor modifications have since been made to this classification (Bernstein et al. 1999; Febvre et al. 2000; Reshetnjak 1981; Tan 1998; Trégouboff 1953). The class Acantharia currently comprises around 50 genera and 150 species, which are grouped into 18 families distributed in 4 orders: Holacanthida, Chaunacanthida, Symphiacanthida and Arthracanthida (Bernstein et al. 1999). The distinction between the four orders is mainly based on the way the spicules cross the cytoplasm. In the Holacanthida, which was considered by Schewiakoff to be the most basal order, 10 diametral spicules loosely cross

the centre of the cell, where they tangle to form a central body (Acanthocollidae) or do not join at all (Acantochiasmidae and Acanthoplegmidae). The Chaunacanthida are characterized by having 20 spicules that are more or less joined at the cell centre and that can be easily dissociated. The Symphiacanthida and the Arthracanthida have 20 tightly joined spicules. The spicules in the Symphiacanthida are attached to each other at the centre by their basal parts, forming a uniform central body. The Arthracanthida, which are characterized by the presence of a thick central capsule, were suggested to represent the most derived forms of Acantharia (Schewiakoff 1926), and are divided into two sub-orders, the Sphaenacantha and the Phyllacantha.

Hitherto, there has been little effort to validate this morphology-based acantharian taxonomy using molecular phylogenetics. Because of the elusive nature of Acantharia and the difficulty to perform accurate morphological identification on living specimens, very few cells have been isolated, morphologically identified and sequenced (about 20 18S rDNA sequences are publicly available to date). For instance, only 4 of the 30 genera of Arthracanthida are represented in GenBank from isolated specimens (Gilg et al. 2009; Oka et al. 2005; Zettler et al. 1997). Molecular phylogenies including these sequences have nevertheless highlighted inconsistencies within the existing morphological classification. The orders Symphiacanthida and Arthracanthida are mixed, the Chaunacanthida includes specimens identified as Symphiacanthida, and the Holacanthida is simply missing in these analyses (Gilg et al. 2009; Oka et al. 2005). Yet, together with more recent investigations (Krabberød et al. 2011), these studies demonstrated the monophyly of the Acantharia among the Radiolaria.

In addition to sequences from isolates, numerous sequences assigned to Acantharia have been retrieved from environmental surveys of genetic diversity in various environments, including coastal (Marie et al. 2010), deep (Not et al. 2007; Quaiser et al. 2010), Antarctic and Arctic (Lopez-Garcia et al. 2001 and Lovejoy et al. 2006 respectively), and anoxic (Alexander et al. 2009; Stoeck et al. 2003) waters, as well as, sediments from the mid-Atlantic ridge and the southern Guaymas vent field (Lopez-Garcia et al. 2003 and Edgcomb et al. 2002 respectively). This considerable diversity of environmental 18S rDNA sequences from Acantharia has no associated morphological information. This phenomenon will undoubtedly be further amplified with the advent of environmental surveys using high-throughput DNA sequencing technologies. As for many protist groups, reference sequences (from morphologically identified organisms) are fundamental anchors to taxonomically characterize the coming profusion of environmental data.

The current vision of the systematics of Acantharia is very obscure. Molecular tools can be helpful to examine the relationships among Acantharia, for which morphology-based classification is unstable (Febvre et al. 1989). The present study aims at producing a comprehensive molecular phylogeny of Acantharia, and shedding light on the morphological evolution of this ecologically key group of marine protists. To do so, we isolated, morphologically identified and sequenced ribosomal DNA markers (18S rDNA and partial 28S rDNA) for more than 100 acantharian specimens collected worldwide. This morpho-molecular

approach on single-cells allowed to assess the validity of the current morphological taxonomy, and to explore the evolution of the group.

Results

Molecular Phylogeny of the Acantharia

The entire 18S rDNA and the D1 and D2 regions of the 28S rDNA were sequenced from acantharian cells isolated in the Mediterranean sea, the Red sea, the English Channel and the West Pacific ocean. In total, 107 partial 28S rDNA and 78 18S rDNA sequences were obtained (Supplementary Table S1). In the 28S rDNA gene, 40.6 % of the 778 characters sequenced are parsimoniously informative compared to 20% of the 1743 characters in the 18S rDNA gene. Overall, the 28S rDNA gene evolves twice as fast as the 18S rDNA, with high regularity (r^2=0.99) across all of the acantharian taxa sampled here (Supplementary Fig. S1).

From the 28S rDNA phylogeny (Fig. 1), 6 clades (A-B-C-D-E-F) sub-divided into 13 sub-clades (C1, C2, D1, etc.) were defined. The genetic divergence of this gene within clades (based on the uncorrected p-distance) is higher in clades C (6.5%) and D (5.7%) compared to clades E (3.9%) and F (3.5%), the latter containing the highest number of individuals (n = 54). The phylogenetic relationships between the 6 clades are generally well resolved, except for the position of clades B and C with bootstrap values (BV) lower than 65%. With weak support (BV 49 %), clade B presents a sister-group relationship with clades D, E and F in the 28S rDNA tree (Fig. 1), but clusters with clade C in the 28S + 18S concatenated phylogeny with higher support (BV 67%; Fig. 2). The position of Ei 68 in clade B is confirmed by the two trees, 28S rDNA and concatenated, and receives stronger support in the latter (49 % compared to 63 %, respectively). Clade A, which contains a single taxon (Ei 59), holds the most basal position and clade C comprises 4 sub-clades (C1 to C4) that are well defined (Figs. 1 and 2). The branching pattern of clades D, E and F are robustly resolved in both phylogenies, especially for 28S rDNA (BV > 94%), and highly supported sub-clades were defined within these clades (D1-D2, E1 to E4, and F1 to F3, respectively). Further sub-groups appear in the largest sub-clade F3, but none of these were unambiguously supported by our overall analysis.

Comparison between Molecular Phylogeny and Morphological Taxonomy

Our sampling included representatives of the 4 orders of Acantharia, including the first sequence data for Holacanthida. Based on morphological criteria, we identified 24 (of the 49 described) genera and 14 (of 18) families. Clades A, B, C and D encompass the two acantharian orders Holacanthida and Chaunacanthida, except for *Phyllostaurus echinoides* (Ei 68) and *Acanthostaurus purpurascens* (Vil 45) in clade B (Figs. 1 and 2), both originally described as belonging to the Arthracanthida. The Holacanthida (genera *Acanthocyrtha*, *Acanthocolla*, *Staurolithium* and *Acanthoplegma*) is spread over the four clades (A, B, C, and D). The genus *Acanthoplegma* (family Acanthoplegmidae), represented by a unique individual (Ei 59) forming clade A (Figs. 1 and 2), is the earliest diverging taxon. Clade B is composed of 7 sequences, of which 5 were from *Acanthocyrtha haeckeli* (family Acanthocollidae). Two individuals of

apparently the same species (Ei 47 and Ei 48) are also found in clade C (sub-clade C4). Members of the family Acanthocollidae, *Staurolithium* sp. and *Acanthocolla cruciata*, are also present in sub-clade D2. The order Chaunacanthida has representatives only in clade C, but the two morphologically-defined families are not clearly distinguished. The family Gigartaconidae is for instance scattered over various sub-clades: *Gigartacon fragilis* (Oki 33 and Oki 91) in sub-clade C2, *Gigartacon muelleri* in C1 (Vil 110) and C4 (Vil 105, Oki 30 and Oki 23), and *Litholophus* sp. (Ei 71, Oki 79) in C4. Similarly, the family Stauraconidae is represented in sub-clades C3 and C4 with respectively *Heteracon biformis* (Vil 126, Vil 47, Vil 65), and *Amphiacon denticulatus* (Vil 131). Acantharia from clade D1 (Ros 6, Ros 7 and Ros 8; Fig. 3D) were identified as *Trizona brandti*, which belongs to the suborder Actinastra (Popofsky 1906, p.56) or to the division Astrolophi (Campbell 1954, D32).

Clades E and F mainly contain representatives from the Arthracanthida, and a few from the Symphiacanthida. The specimens forming clade E are easier to identify compared to other Acantharia, due to the presence of a shell (Fig. 3G-L). In the traditional taxonomy, they belong to the families Dorataspidae, Phractopeltidae, Diploconidae, and Hexalaspidae, in the sub-order Sphaenacantha. Sub-clade E1 includes species from the Phractopeltidae with *Phractopelta sarmentosa* (Oki 45 and 72) and *Phractopelta dorataspis* (Oki 42), and species from the Dorataspidae with *Lychnaspis giltschii* (Oki 46, Oki 55 and Oki 36). Sub-clade E1 contains as well the individual Oki 73, which was originally described by Haeckel (1887, p. 612, pl. 50, fig. 8) as a Spumellaria (*Larcidium dodecanthum*). Sub-clades E2 and E3 consist of three *L. giltschii* (Oki 11, Oki 49 and Oki 54) and three *Diploconus fasces* (Diploconidae) sequences, respectively. Sub-clade E4 contains four Acantharia belonging to the Dorataspidae: *Coleaspis vaginata* (Vil 86 and Vil 25) and *Dorataspis loricata* (Oki 4 and Oki 74). Like all of the taxa from clade E, the families Acanthometridae and Lithopteridae belong to the sub-order Sphaenacantha in the traditional classification. However, our phylogenetic analyses show that Acanthometridae (*Acanthometra fusca*, *Acanthometra pellucida*, *Amphilonche elongata*) and Lithopteridae (*Lithoptera fenestrata*) fall in sub-clades F3 and F1, respectively. In addition to the Acanthometridae, sub-clade F3 contains representatives of two other families: the Phyllostauridae (*Acanthostaurus* sp., *Phyllostaurus* sp. and *Amphistaurus* sp.) and the Stauracanthidae (*Xiphacantha* sp. and *Stauracantha* sp.). These two families are part of the sub-order Phyllacantha in the Arthracanthida. Finally, sub-clade F2 contains three individuals of one genus *Amphibelone*, which belongs to the order Symphiacanthida. This is the only isolated representative from this order recovered in this study.

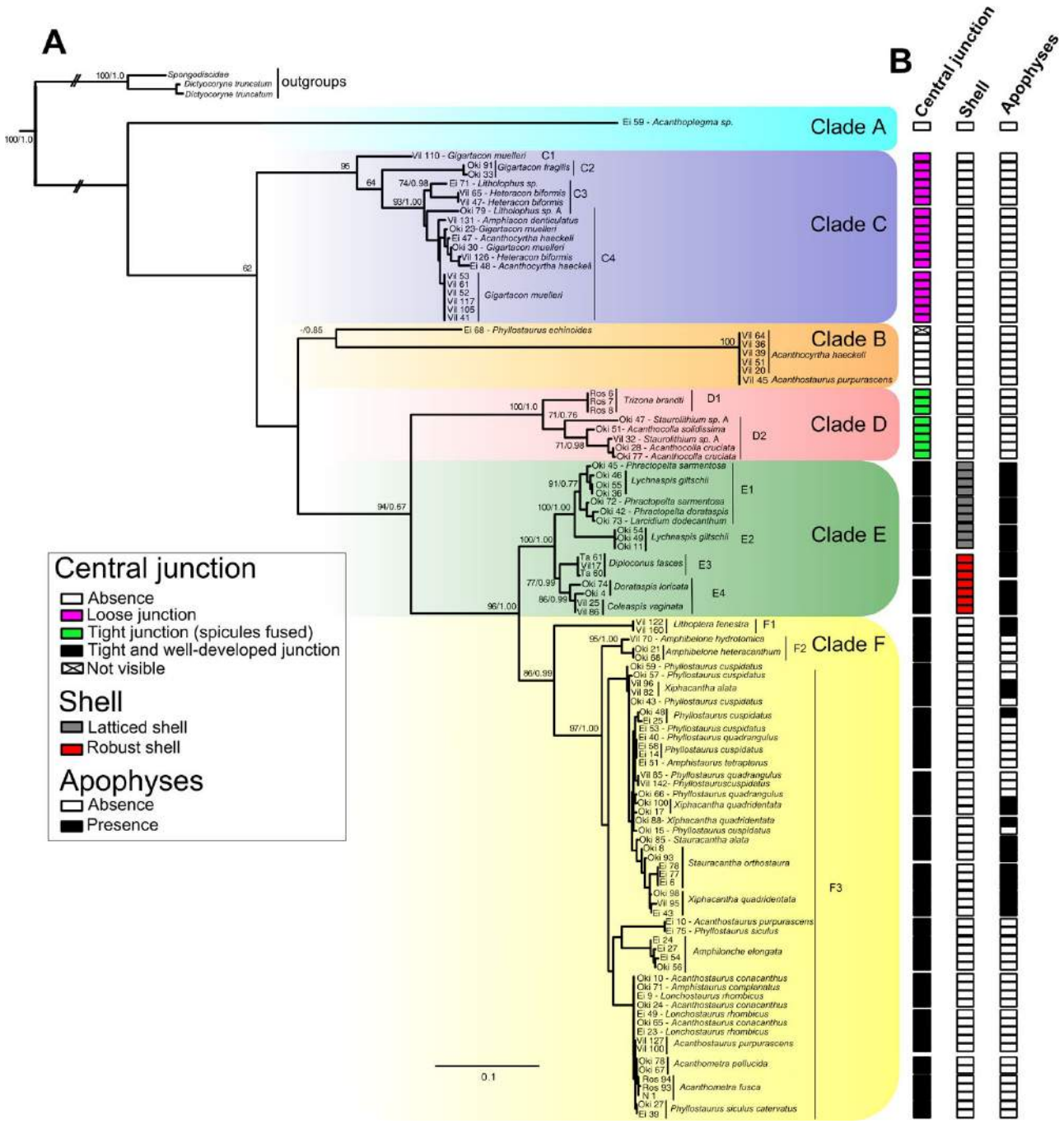


Figure 1. 28Sr DNA molecular phylogeny and mapping of morphological characters of Acantharia sampled in this study. A. Maximum likelihood tree inferred from 28S (D1 + D2) rDNA sequences (110 taxa with 702 aligned positions) and a GTR+I+G model of nucleotide substitutions. PhyML bootstrap percentages based on 1000 pseudo-replicates and bayesian posterior probabilities (PP) are indicated at each node when support values are higher than 60% and 0.8, respectively. Six major clades (A-F) and thirteen sub-groups have been defined based on the support values and morphological criteria. The tree is rooted with sequences of three other radiolarian (Spumellaria) species. Branches with a double barred symbol are reduced to half length for clarity, and the scale bar corresponds to 0.1 substitutions per site. B. Skeleton traits identified for each acantharian cell mapped onto the molecular phylogeny: 1) junction of spicules at the cell centre, 2) type of the shell and 3) presence/absence of apophyses.

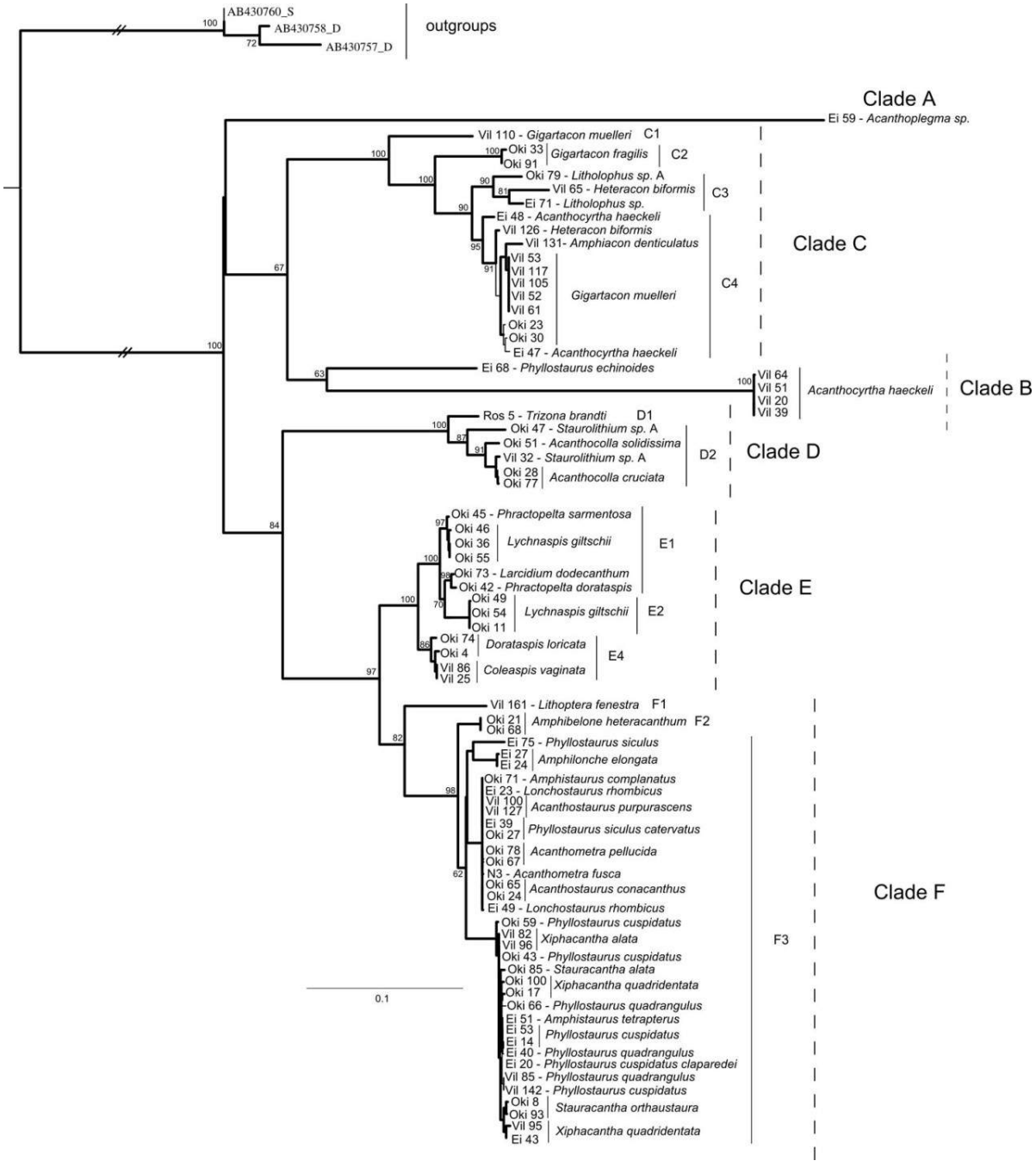


Figure 2. Concatenation (entire 18S + partial 28S rDNA) topology obtained from Maximum-Likelihood (ML) analysis. The tree was constructed with 1000 bootstraps, using an alignment of 82 taxa with 2172 positions. The tree is rooted with sequences of three spumellarian species. Only ML bootstraps values higher than 60% are shown. Branches with a double barred symbol are reduced to half length for clarity. The reference clades defined in this study are shown (A-B-C-D-E-F).

Evolutionary History of the Acantharian Skeleton

Mapping morphological characters onto the molecular phylogeny highlights dominant trends in acantharian evolution. We focused on morphological characters of the skeleton that we considered to be the most informative: central junction, shell, and apophyses (Fig. 1). Regarding the central junction, *Acanthoplegma* sp. defining clade A has 10 diametral spicules crossing the cytoplasm, a unique feature among all of the taxa sampled in this study. Specimens from clades B and C exhibit 20 radial spicules embedded in contractile matter without fused central junction. We observed that the spicules of these two clades can be folded like an umbrella (Fig. 3B). The organisation of the spicules in the taxon Ei 68 (clade B) was unfortunately not visible inside the cell due to a particularly dense cytoplasm (Fig. 3C). The simple spicule arrangement without junction of clades A, B and C could be considered as an ancestral character due to their early divergent positions, when compared to clades D, E and F (Figs. 1 and 2). In clade D, the 20 spicules are fused in the cell centre, forming a ball-like junction (Fig. 3D-F), and in clades E and F, they are tightly joined and fused, creating a well developed central body (Fig. 3M-O).

Apophyses, which are small lateral celestite extensions on each spicule, characterize all the taxa from clades E (Fig. 3J-L), and some lineages in clade F. In this last clade, the family Lithopteridae, with its unique wing-like fenestrated apophyses (Fig. 3N) forms a separate lineage in the molecular phylogeny (sub-clade F1; Figs. 1 and 2).

Finally, the shell appears exclusively in clade E. This character does not appear to be homoplasic and could constitute a key diagnostic character for this clade. Taxa from sub-clade E1 can even build a double concentric shell (*Tessaropelma* sp.; Fig. 3K). The molecular phylogeny distinguishes two kinds of shells: latticed shells (Fig. 3J-L) formed by the junction of apophyses in sub-clades E1 and E2 (*Lychnaspis* sp., *Phractopelta* sp. and *Larcidium* sp.), and armoured shells (Fig. 3G-I) consisting of robust plates in sub-clades E3 and E4 (*Coleaspis* sp., *Dorataspis* sp. and *Diploconus* sp).

Environmental Genetic Diversity of the Acantharia

We used the pplacer software (Matsen et al. 2010) to map all the 18S rDNA sequences affiliated to Acantharia (142 environmental and 24 reference sequences from GenBank), whatever the length and the region, to the full-length 18S rDNA reference phylogeny produced in this study. For some Acantharia isolated in this study, we could not obtain the entire 18S rDNA gene but a sequence representing the first (750 pb) or the second part (1000 pb) of this gene (e.g. the second part for Ei 59, the first part for Vil 119). These partial 18S rDNA sequences (18 in total) were also mapped in this analysis as query sequences. Numerous environmental sequences are placed within the reference clades defined above (Fig. 4): 2 in clade B, 50 in clade C, 2 in clade D, 5 in clade E and 18 in clade F. However, 65 environmental sequences could not be associated with these clades (unplaced sequences) and form 5 distinct branches, labelled 37, 49, 148, 149 and 150. Some partial sequences from isolated acantharian cells are placed in some of these branches. The taxa Vil 119 (*Litholophus* sp.) is found in branch 149 and Ei 59 (*Acanthoplegma* sp.) in branch 150 (Fig. 4 and Supplementary Table S2). To better understand the phylogenetic

relationships of the unplaced environmental sequences contained in the 5 branches, further analyses were carried out on the first 662 nucleotides (as most environmental sequences correspond to this region) of 179 reference and environmental 18S rDNA sequences (Fig. 5). The unplaced environmental sequences are forming 4 new clades: Acanth I, II, III and IV. The individual Vil 119 is clearly associated with the clade Acanth IV. The 3 clades Acanth I, II and III hold a basal position in the phylogeny, and none contains reference sequences. In order to see whether Ei 59 (clade A) belongs to one of these three clades, additional phylogenetic analyses were performed with the second part of the 18S rDNA with about 780 bp (data not shown). The individual Ei 59 representing the Clade A groups with the clade Acanth II.

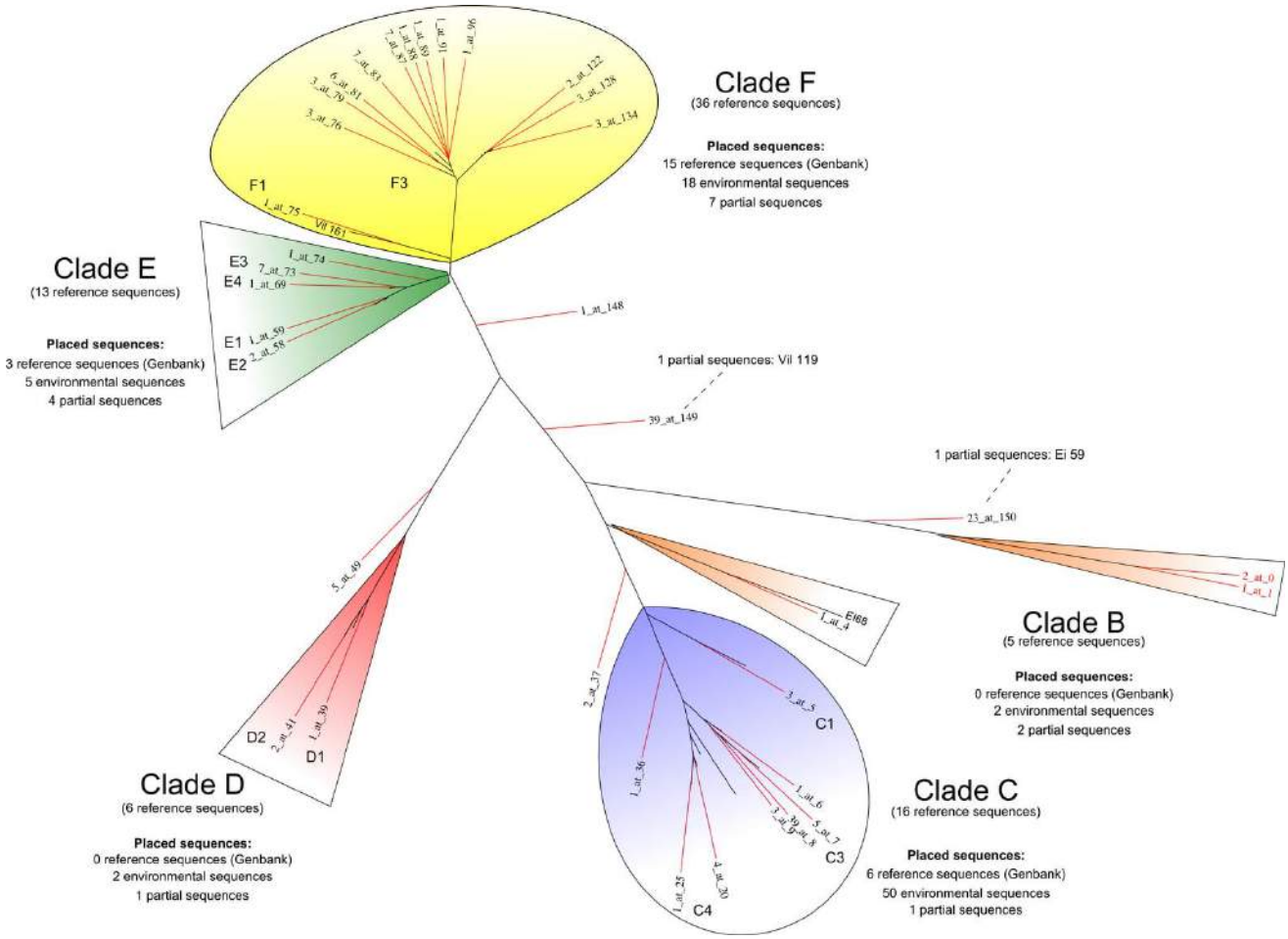


Figure 4. pplacer tree where 3 kinds of sequences were phylogenetically placed onto the 18S rDNA reference phylogeny produced in this study: 24 reference and 142 environmental sequences from GenBank, and 18 partial sequences from isolates produced here. pplacer determined the most probable location for each sequence and represented them as additional branches in the tree. The first number and the second number on each branch correspond respectively to the name of the branch and the amount of sequences forming the branch. 77 environmental sequences are associated with the 5 main clades of Acantharia, B-C-D-E-F, whereas 65 do not have affinity with any of these reference clades and form the branches 37, 49, 148, 149 and 150 (see Supplementary Table S2 for detailed information about each sequence).

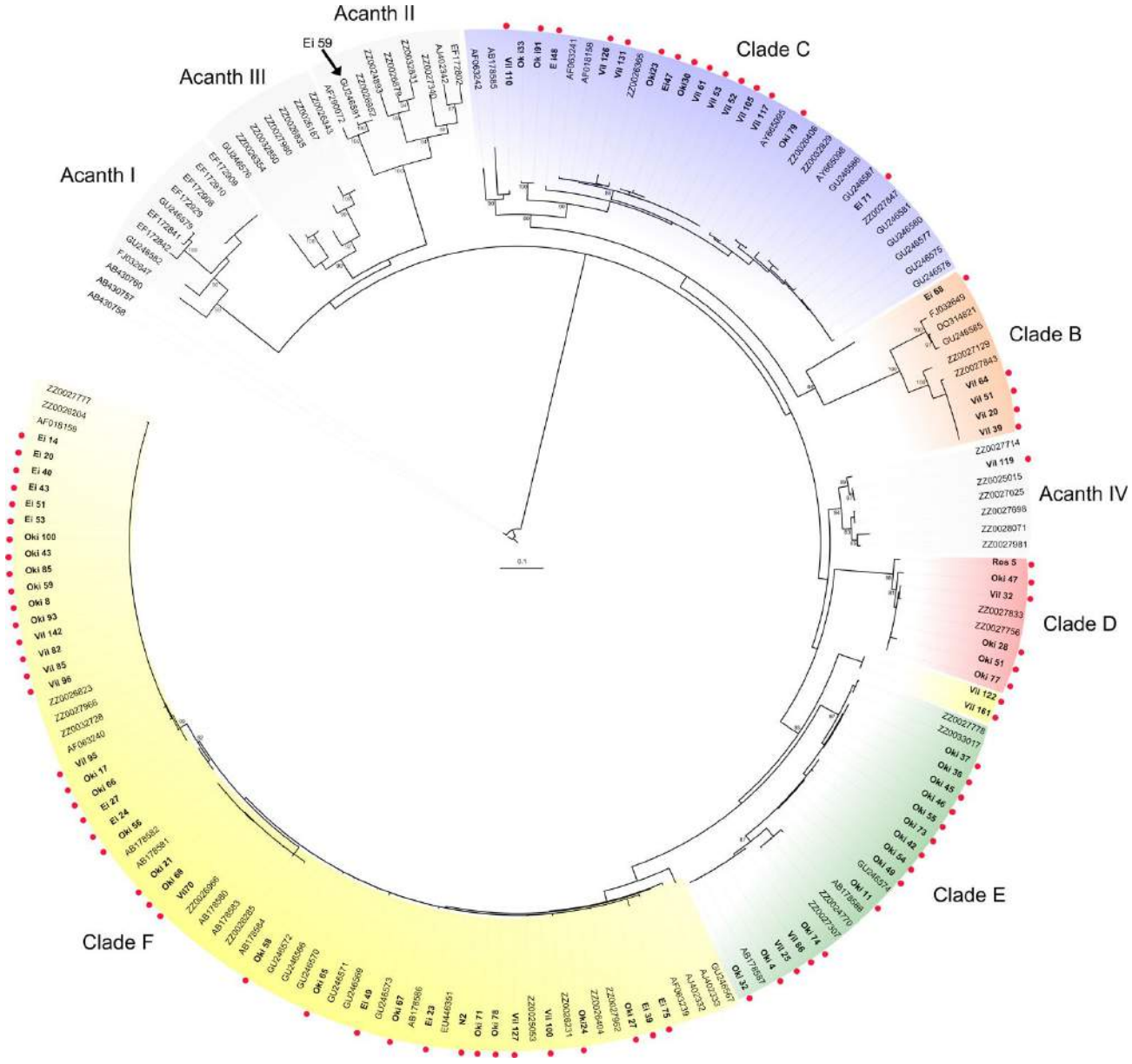


Figure 5. Phylogeny obtained from Maximum-Likelihood (ML) analysis with the first part (662 positions) of the 18S rDNA gene with 179 taxa (model TPM2uf+I+G). Reference sequences (symbolized by small dots) and environmental sequences recovered from other studies (see Supplementary Table S2) have been analysed. The tree is rooted with sequences of three spumellarian species like the other phylogenies in this study. ML bootstraps > 80% are shown (1000 pseudo-replicates). The environmental clades Acanth I and Acanth III do not contain any reference sequences, whereas clade Acanth IV contains Vil 119. Ei 59 from clade A belongs to clade Acanth II, based on a phylogenetic analysis with the other part of the 18S rDNA gene (data not shown).

Discussion

Global Genetic Diversity of the Acantharia

We present a comprehensive molecular phylogeny of the Acantharia, obtained by carrying out broad taxonomic sampling on single cells. 107 partial 28S rDNA and 78 full-length 18S rDNA sequences allowed defining 6 main clades (A to F), and at least 13 sub-clades. Many environmental sequences related to Acantharia, from various worldwide oceans, are associated with these different reference entities, providing insights into the biogeography and the phenotype associated to all of these sequences. Our phylogenetic analyses on the first part of the 18S rDNA gene (Fig. 5) recovered all of the new clades highlighted in the pplacer analysis, indicating that this region gives a good picture of extant acantharian genetic diversity. Two environmental clades (Acanth I and Acanth III) have no association to any reference sequences, and Acanth II and Acanth IV contain only one reference sequence each (Ei 59 and Vil 119, respectively). Acanth II would therefore represent the family Acanthoplegmidae. As the morphological identification of Vil 119 is uncertain (*Litholophus* sp. from Chaunacanthida), we are not able to associate clade Acanth IV with a family or order from the traditional taxonomy. Clades Acanth I and Acanth II correspond to clades Rad I and Rad II in Not et al. 2007, respectively. The basal clades Acanth I-II-III-mainly include environmental sequences sampled at depth (e.g. 1000, 2000 and 3000 meters, Countway et al. 2007; Gilg et al. 2009; Not et al. 2007; Quaiser et al. 2010) and in relatively extreme environments (hydrothermal vents, Arctic and Antarctic waters, Edgcomb et al. 2002; Lopez-Garcia et al. 2001, 2003; Terrado et al, 2009). Since we mostly sampled surface waters, this could explain the reason why we did not cover all of the existing genetic diversity. Acantharia are known to thrive in the upper water column (Massera Bottazzi and Andreoli 1982), but it is not known whether some species live deeper. Schewiakoff (1926) observed that Acanthoplegmidae can be found very deep, down to 1000 m, supporting the idea that this family could be represented by the clade Acanth II. Some of the deep acantharian sequences could also be linked to the complex and poorly understood life cycle of Acantharia. Indeed, sexual reproduction may occur deep in the water column where thousands of gametes are shed from sinking cysts or vegetative adults (Antia et al. 1993; Bernstein et al. 1987; Martin et al. 2010; Schewiakoff, 1926). In addition to the acantharian diversity structure according to depth, there may be differences between waters of distinct nutrient status. We might have missed species which preferentially thrive in nutrient-rich waters. Finally, our sampling did not encompass polar regions, and some acantharian species could be endemic to these areas, as it was observed in planktonic foraminifera (Darling et al. 2007). Sampling of mesopelagic and bathypelagic waters from different oceans, with the same single-cell approach adopted here, could be considered in future studies to unveil more genetic diversity in Acantharia. The reference framework presented here, which bridges molecular and phenotypic data, will contribute to a better understanding of the distribution, and more broadly the ecology of Acantharia in marine ecosystems. This is of paramount importance considering the ongoing next-generation sequencing era that is revolutionising environmental surveys of protist diversity.

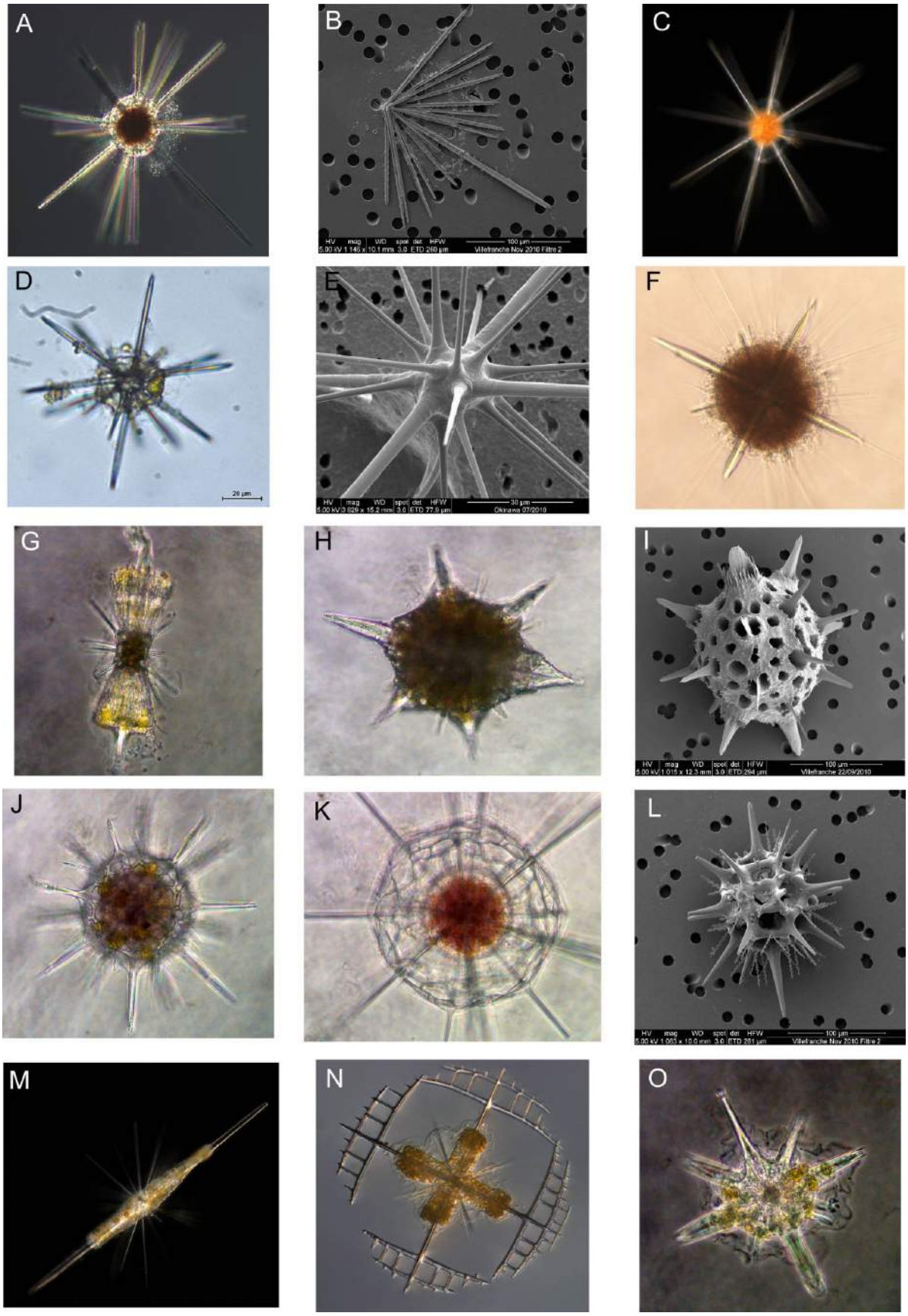

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O

Figure 3. A panel of selected acantharian species photographed by light and scanning electronic microscopy and representing clade C (pictures A and B), clade B (C), clade D (D-F), clade E (G – L) and clade F (M-O). Some morphological characters are visible: central junction can be loose (picture A and B), tight with fused spicules (D-F), or tight and well-developed (G-O); presence of a robust shell (G-I) or a latticed shell (J –L); presence of apophyses (N).
A: *Gigartacon muelleri* (vil 110) – Chaunacanthida. B: *Heteracon biformis* (Vil 65 and Vil 47) – Chaunacanthida. C: *Phyllostaurus echinoides* (Ei68) – Arthracanthida. D: *Trizona brandti* (Ros 6) – Actinastra or Astrolophi. E and F: *Staurolithium cruciatum* (Oki 47) - Holacanthida. G: *Diploconus fasces* (Vil 17) – Arthracanthida. H and I: *Hexaconus* sp. – Arthracanthida. J and L: *Lychnaspis giltschii* (Oki 46 and Oki 49) – Arthracanthida. K: *Lychnaspis* sp. – Arthracanthida. M: Amphilonche elongate (Ei 54) – Arthracanthida. N: *Lithoptera fenestra* (Vil 160) – Arthracanthida. O: *Acanthostaurus conacanthus* (Oki 24) – Arthracanthida.

Probing the Congruence between DNA-based Phylogeny and Traditional Taxonomy

The comparison between our molecular phylogeny and the current morphological taxonomy of J. Müller, E. Haeckel, W. Mielck, A. Popofsky and W.T. Schewiakoff reveals an overall discrepancy, as it is often the case when ancient classifications and molecular work are confronted (Gontcharov and Melkonian 2011). The 14 families and 24 genera that we sampled do not appear as independent lineages, and some species can be found across one or even two molecular clades. However, a certain consistency with the traditional taxonomy can be observed by considering general patterns from the molecular phylogeny. The two orders Chaunacanthida and Holacanthida are represented by one large group, which encompasses the most basal clades A, B and C, and clade D. The Arthracanthida and Symphiacanthida correspond to the most recently diverged clades E and F (Fig. 1 and 2). The Chaunacanthida appears to be monophyletic as all representatives fall in clade C. This order was also reported to be a distinct lineage in a recent study (Gilg et al. 2009). The isolated specimen "Symphiacanthid 211" in clade C (Fig. 5) is probably a misidentification as suggested by Gilg et al (2009). Although the family Conaconidae (that contains one species, *Conacon foliaceus*) is missing in our dataset, we can note that no monophyletic patterns emerged at the family level within the Chaunacanthida. The order Holacanthida is polyphyletic, being distributed across clades A, B and D. The morpho-molecular phylogeny distinguishes different kinds of central junction within the Holacanthida, which is not the case in traditional taxonomy. This suggests that central junction is a reliable diagnostic character that should be reconsidered for future revisions of the taxonomy.

Clades E and F, representing the orders Arthracanthida and Symphiacanthida, are the most recently diverged clades (Fig. 1 and 2). All of the shell-bearing species of clade E (Fig. 3G-L) belong to the sub-order Sphaenacantha (Arthracanthida), but there is no clear pattern at finer taxonomic levels among the sub-groups E1 to E4. Nevertheless, the molecular phylogeny reflects the way the shell is built, which was not considered as an important diagnostic criterion in the traditional taxonomy. Members of sub-clades E1 and E2 have a shell exhibiting a latticed architecture (Fig. 3J-L), whereas the shell is more thick and robust in sub-clades E3 and E4 (Fig. 3G-I). In the current taxonomy, for instance, the family Dorataspidae includes the two kinds of shell. The non shell-bearing families Acanthometridae and Lithopteridae also belong to the Sphaenacantha in Schewiakoff's taxonomy, but our molecular phylogeny rejects this position.

Indeed, several genera of these families, such as *Amphilonche*, *Acanthometra* and *Lithoptera* are not found in clade E, but rather in clade F. We can therefore argue that the presence of a shell is a specific character of clade E, the Sphaenacantha, and characteristics of the shell have to be re-examined to improve the classification within this sub-order.
Several subgroups appear within clade F (Figs. 1, 2 and 5), but only the Lithopteridae (F1) and *Amphibelone* sp. (F2) are statistically supported. Although we sampled only one symphiacanthid genus (*Amphibelone*) out of 9 morphologically described (Bernstein et al. 1999; Schewiakoff 1926), it is very unlikely that the Symphiacanthida has a monophyletic nature in molecular phylogeny. Other species (*Amphibelone* sp. and *Haliommatidum* sp.) from GenBank sequences do not group together within clade F (Figs. 4 and 5), being mixed within the Arthracanthida. We note that the existence of the Symphiacanthida as a distinct order has already been questioned because of the morphological similarity with the Arthracanthida (Hollande et al. 1965). The large sub-group F3 encompasses members of the families Phyllostauridae, Stauracanthidae and Acanthometridae, and would therefore represent the second sub-order of the Arthracanthida, the Phyllacantha. Although these families can be distinguished by their spicules (e.g. length and characteristics of the apophyses), this is not reflected in the molecular phylogeny as sub-group F3 lacks resolution. The use of more rapidly evolving genes could allow fine-tuning of morpho-genetic comparisons within the Acantharia, especially in clade F.

Hypothesis on the Evolutionary Pathway of Acantharian Skeleton

The skeleton of Acantharia is unique among protists, raising many questions about its origin and evolutionary history. Due to a rapid dissolution in seawater after death, there is no fossil record of Acantharia, and the evolution of the morphology is therefore enigmatic. The acantharian spicules are single crystals of strontium sulphate, radiating from the cell centre (Schreiber et al. 1962; Wilcock et al. 1988). In Collodaria (Radiolaria), swarmer cells (first reproductive stages) have been observed to contain several strontium sulphate crystals in separated vesicles (Anderson et al. 1990; Hollande and Martoja 1974; Hughes et al. 1989; Perry and Hughes 1990). These crystals are present for a brief period and disappear in later development stages. Their precise function is unclear, but it has been suggested that they may be involved in buoyancy control (Anderson et al. 1990). Although these crystals have not been observed in Acantharia so far, it can be hypothesized that these crystals are a plesiomorphic state within Radiolaria and developed in the acantharian ancestor, leading to the formation of spicules. From single crystals, a wide variety of symmetrical skeletons would have emerged throughout the evolution of the Acantharia. The mapping onto the phylogeny of skeletal features, such as the central junction, shell and apophyses, shed light on the evolutionary pathway of the acantharian skeleton. The early divergent taxon (clade A) has 10 long and similar needle-like spicules, which simply cross the cytoplasm without any junction. Then, spicules increase in number to 20 in the other acantharian clades, and can be of different length and shapes. These spicules are either embedded in an organic matrix or loosely joined at the cell centre. In the most recently diverged clades (E and F), the central junction can become very tight and develop into a wide and robust plate-like

structure. Considering this evolutionary pattern, one may argue that the skeleton of the environmental clade Acanth I, which diverged before all the other clades (Fig. 5), may be different than the skeletons observed in this study, and could be elementary or even absent. In parallel to the evolved central junction, some Acantharia display complex mineral structures, such as apophyses, latticed or armoured shells. These milestones in the evolutionary history of the acantharian skeleton are a good example of how single-celled organisms can transform mineral crystals through time to develop increasingly complex skeletons.

Methods

Single-cell acantharian collection

Acantharia were collected in 2010 at 5 different locations: in the Mediterranean Sea (Naples 40°50'24"N 14°15'58" E, and Villefranche sur Mer 43°42′18″N 7°18′45″E), in the Red Sea (Gulf of Eilat 29°33'27"N 34°57'69"E), in the Pacific (Akajima islands, Okinawa 26°11'56"N 127°16'5"E) and in the English Channel (off Roscoff 48°43'36" N 3°59'10" E) (Supplementary Table S1). Acantharian cells were harvested at the subsurface by slowly towing nets (64 and 200 µm mesh size). Micropipette isolation of cells using an inverted microscope was conducted as soon as possible after sampling. Each sorted cell was maintained for several hours in 0.2 µm filtered seawater to allow self-cleaning (debris and particles are removed and prey digested). Following incubation, each acantharian cell was further cleaned and rinsed several times in 0.2 µm filtered seawater, and finally individually isolated into a guanidine-containing extraction buffer (GITC, de Vargas et al. 2002). Before transferring the cell into GITC, photographs were taken under an inverted microscope. Eppendorf tubes containing the samples were stored at -20 °C until processing.

Morphological identification

For this study, all of the published work regarding acantharian taxonomy, including emendations, was examined (see in references list) and a taxonomic synonym database was built using PaleoTax (http://www.paleotax.de/index.htm). The morphological criteria used in previous classifications were evaluated with a database of about 6000 acantharian individuals collected in many areas, including those from the present study. In addition, the original microscope slides and working notebooks of Schewiakoff (1926) were thoroughly studied. Each acantharian specimen sampled in this study was identified based on morphological criteria observed in vivo through the microscope and photographs. A matrice of about 25 morphological characters for all the taxa we sampled was built. Three characters (central junction, presence of a shell and apophyses) turned to be more important than the others as their different states seemed to carry a phylogenetic signal. The isolates were coded according to the sampling site (e.g. Oki and Ei for Okinawa and Eilat, respectively).

DNA extraction, amplification, sequencing and alignment: For single-cell DNA extraction, tubes were thawed at 70 °C for 20 minutes. The nucleic acids were precipitated at -20 °C

overnight by adding 100µl of isopropanol. Tubes were then centrifuged for 45 minutes at maximum speed and supernatant was carefully removed. Ethanol was added as a rinsing step and tubes were again centrifuged for 30 minutes. The supernatant was discarded and tubes were dried for 3 hours at room temperature. DNA was finally eluted in 20 µl TE buffer. Single-cell Polymerase Chain Reaction (PCR) was conducted to amplify two ribosomal DNA genes: 18S and 28S (region D1 and D2) with primers described in Table 1. The entire 18S rDNA gene was obtained in two steps with 2 pairs of specific primers amplifying the first and the second part of the gene. The 18S and 28S genes were amplified with GoTaq (Promega, Lyon, France) in a 25-µl reaction volume using the following PCR parameters: 5 min at 95 °C, followed by 35 cycles of 30 s denaturation at 95 °C, 30 s annealing at 50 °C for the 28S, 55 °C for the 18S, and 2 min extension at 72 °C, with a final elongation step of 10 minutes at 72 °C. Following amplification, the PCR products were purified by EXOSAP-IT (GE Healthcare Bio-Sciences Corp.), and bidirectionnaly sequenced using the ABI-PRISM Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems). Similar morphospecies have been sequenced at least twice in order to check for putative genetic polymorphisms and confirm their phylogenetic placement.

Table 1. Sequences of eukaryote (SA and SB) and radiolaria-specific primers (S879, S32J, 28S Rad2 and ITSa3) used in this study to amplify the entire 18S rDNA and the partial 28S rDNA genes. SA/S879 and S32J/SB amplify the first 792-838 nucleotides and the last 1090-1109 nucleotides of the 18S rDNA gene, respectively.

Name	Gene	Sequence (5'-3')	Direction	Reference
SA	18S (1st part)	AAC CTG GTT GAT CCT GCC AGT	F	Medlin et al. 1988
S879	18S (1st part)	CCA ACT GTC CCT ATC AAT CAT	R	This study
S32 J	18S (2nd part)	CCA GCT CCA ATA GCG TAT AC	F	This study
SB	18S (2nd part)	TGA TCC TTC TGC AGG TTC ACC TAC	R	Medlin et al. 1988
28S Rad2	28S	TAA GCG GAG GAA AAG AAA	F	Ando et al. 2009
ITS-a3	28S	TCA CCA TCT TTC GGG TCC CAA CA	R	Ando et al. 2009

Phylogenetic analysis

The different matrices of 18S rDNA and partial 28S rDNA sequences were aligned using the default parameters of T-COFFEE 8.98 "meta mode", which constructs a consensus alignment based on the results of 12 multiple sequence alignment (MSA) methods (Moretti et al. 2007; Wallace et al. 2006). From these alignments, various data were derived. Firstly, 18S rDNA and 28S rDNA pairwise-distances were calculated with T-COFFEE (uncorrected-p distance) and compared in order to detect evolution rate differences between the two targeted genes. Secondly, the software PAUP* 4.0b10 (Swofford 2001) was used to estimate the phylogenetic signal by assessing the percentage of parsimoniously informative characters present in 18S rDNA and 28S rDNA matrices. Thirdly, the general time-reversible (GTR) model of nucleotide substitution was selected for 18S rDNA (81 taxa; 1404 positions), 28S rDNA (110 taxa; 702 positions) and 18S

+28S rDNA (81 taxa; 2115 positions) data sets according to jModeltest v0.1.1 (Posada 2008) and the corrected Akaike information criterion (Posada 2009; Posada and Buckley 2004). Three sequences from spumellarian species (Radiolaria) were used as outgroups for the analysis of each dataset: 2 sequences of *Dictyocoryne truncatum* and one of Spongodiscidae sp. The phylogenetic inference by Maximum Likelihood (ML) was performed with PhyML v3.0 (Guindon and Gascuel 2003) using invariant sites and 8-class gamma distributions. Robustness of inferred topologies was assessed by 1,000 non-parametric bootstrap re-sampling. The Bayesian inference was conducted on the 28S rDNA dataset using Beast v.1.5.4 and companion softwares (Drummond and Rambaut 2007) under the GTR model taking into account 4-class gamma and invariant sites. Three Markov Chain Monte Carlo (MCMC) chains were run for 30 million generations, sampling every 1000 generations. The first 3000 trees were discarded as burn-in as indicated by the program Tracer v.1.4.1, and the remaining trees were used by TreeAnnotator 1.5.4 to build the consensus tree. The final tree was visualized with FigTree v1.3.1.

Phylogenetic analysis of the acantharian environmental sequences

All of the existing environmental 18S rDNA sequences related to Acantharia available in GenBank (in January 2011) were downloaded (see Supplementary Table S2). Most of these sequences are partial, covering different part of the 18S rDNA gene. In order to be exhaustive and take into account all of these partial sequences, we used the program pplacer 1.0 (Matsen et al. 2010) and the full-length 18S rDNA data set built in this study as a reference. The 142 environmental and 24 single-cell sequences from other studies (query sequences, Supplementary Table S2) were aligned with a HMM built from the reference alignment using tools from the HMMER v3.0 suite (http://hmmer.org/), with default parameters. Partial sequences that we obtained in this study from isolated Acantharia (Supplementary Tables S1 and S2) were not included in the reference alignment but also analysed as query sequences (e.g. Ei 59 from clade A and Vil 119). Using this alignment and the inferred reference tree, pplacer determined the most probable location for each query sequence, whatever the fragment length or position, and represented them as additional branches in the tree. In order to confirm the pplacer results and to further investigate the evolutionary relationships between environmental and reference sequences, we carried out a complementary phylogenetic analysis. To do so, we focused on the first part of the 18S rDNA (662 positions, 179 taxa) as most environmental sequences correspond to this region. The phylogenetic reconstruction was made as explained before with the TPM2uf+I+G model. TPM2uf is a TPM2 model with unequal bases frequencies and substitution rates as such: AC=AT; CG=GT; AG=CT. TPM stands for "three-parameter model" (Kimura 1981).

Acknowledgements

This work was supported by the PhD fellowship DIPHOPE and SYMFORAD project from the Region Bretagne. We thank the marine stations which hosted us for field sampling: LOV from Villefranche sur Mer (John Dolan); the Interuniversity Institute for Marine Sciences in Eilat (EU FP7 ASSEMBLE program funded this sampling mission); Stazione Zoologica Anton Dohrn in Naples; Akajima Marine science Laboratory in the Okinawa islands (CNRS-JST PhytoMetagene project.from Daniel Vaulot); Station Biologique de Roscoff and the marine service crew. We are grateful to Kihara Akira (Hosei University, Japan), Flegra Bentivegna and Christianen Groeben (Statione Zoologica Anton Dohrn, Napoli) for providing access to Schewiakoff's reference material. We thank Ian Probert for valuable comments on previous drafts, and Colette Febvre, who kindly shared her knowledge about Acantharia and encouraged this work.

References

Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol **52**: 399–451

Alexander E, Stock A, Breiner HW, Behnke A, Bunge J, Yakimov MM, Stoeck T (2009) Microbial eukaryotes in the hypersaline anoxic L'Atalante deep-sea basin. Environ Micro- biol **11**: 360-381

Anderson OR, Perry CC, Hughes NP (1990) Transmission and scanning electron microscopic evidence for cytoplasmic deposition of strontium sulphate crystals in colonial radiolaria. Philos Trans R Soc Lond B **329**: 81-86

Ando H, Kunitomo Y, Sarashina I, Iijima M, Endo K, Sashida K (2009) Intraspecific variations in the ITS region of recent radiolarians. Earth Evol Sci **3**: 37-44

Antia AN, Bauerfeind E, Von Bodungen B, Zeller U (1993) Abundance, encystment and sedimentation of Acantharia during autumn 1990 in the East Greenland Sea. J Plankton Res **15**: 99–114

Beers JR, Stewart GL (1970) The preservation of acantharians in fixed plankton samples. Limnol Oceanogr **15**: 825-827

Bernstein R, Kling SA, Boltovskoy D (1999) Acantharia. In Boltovoskoy D (ed) South Atlantic Zooplankton, vol.1. Backhuys Publishers, Leiden, The Netherlands, pp 75-147

Bernstein RE, Betzer PR, Feely RA, Byrne RH, Lamb MF, Michaels AF (1987) Acantharian fluxes and strontium to chlorinity ratios in the North Pacific Ocean. Science **237**: 1490-1494

Campbell AS (1954) Radiolaria. In: Moore RC, (eds.) Treatise on Invertebrate Paleontology, Part D Protista 3. Protozoa (Chiefly Radiolaria and Tintinnina), Geological Society of America and University of Kansas Press, Kansas, pp. D1–D163

Caron D, Swanberg N (1990) The ecology of planktonic sarcodines. Rev Aquat Sci **3**: 147-180

Countway PD, Caron DA, Gast RJ, Savai P (2007) Comparison of protistan diversity in deep (2500m) vs euphotic zone assemblages in the Sargasso Sea and Gulf Stream (N.Atlantic). Environ Microbiol: **9**: 1219-1232

Darling KF, Kucera M, Wade CM (2007) Global molecular phylogeography reveals persistent Arctic circumpolar isolation in a marine planktonic protist. Proc Natl Acad Sci USA **104**: 5002-5007

de Vargas C, Bonzon M, Rees NW, Pawlowski J, Zaninetti L (2002) A molecular approach to biodiversity and biogeography in the planktonic foraminifer *Globigerinella siphonifera* (d'Orbigny). Mar Micropaleontol **45**: 101-116

Edgcomb VP, Kysela DT, Teske A, Gomez AD, Sogin ML (2002) Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. Proc Natl Acad Sci USA **99**: 7658-7662

Febvre J (1981) The myoneme of the Acantharia (Protozoa): a new model of cellular motility. BioSystems **14**: 327-336

Febvre J (1989) Phylum Actinopoda, Class Acantharia. In Marulis L, Corliss JO, Melkonian M, Chapman DJ (Eds) Handbook of Protoctista, Jones and Bartlett Publishers, Boston, pp 363-379

Febvre J, Febvre C, Michaels A (2000) Acantharia Haeckel, 1881. In Lee JJ, Leedale GF, Bradbury P (Eds) An Illustrated Guide to the Protozoa (2nd edition). Organisms Traditionally Referred to As Protozoa, or Newly Discovered Groups, Society of Protozoologists, Kansas, 783-803

Gilg IC, Amaral-Zettler LA, Countway PD, Moorthi S, Schnetzer A, Caron D (2009) Phylogenetic affiliations of mesopelagic acantharia and acantharian-like environmental 18S rRNA genes off the Southern California coast. Protist **161**: 197-211

Gontcharov AA, Melkonian M (2011) A study of conflict between molecular phylogeny and taxonomy in the Desmidiaceae (Streptophyta, Viridiplantae): analyses of 191 *rbc*L sequences. Protist **162**: 253-267

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52**: 694-704

Haeckel E (1887) Report on Radiolaria Collected by H.M.S. Challenger during the Years 1873-1876. In Murray CWT aJ (ed), The Voyage of the H.M.S.Challenger, vol. 18. Her Majesty's Stationary Office, London, pp 1803

Haeckel E (1888) Die Radiolarien (Rhizopoda Radiolaria). Eine Monographie, Dritter Teil, Die Acantharien oder Actinopyleen-Radiolarien. Verlag von Georg Reimer, Berlin, pp 33

Haug GH, Tiedemann R (1998) Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. Nature **393**: 673-676

Hollande A, Martoja R (1974) Identification du cristalloide des isopores de Radiolaires à un crystal de celestite (SrSO4). Détermination de la constitution du cristalloide par voie cytochimique et à l'aide de la microsonde électronique et du microanalyseur à émission ionique secondaire. Protistologica **10**: 603-609

Hollande A, Cachon J, Cachon-Enjumet M (1965) Les modalités de l'enkystement présporogénétique chez les acanthaires. Protistologica **1**: 91-112

Hughes NP, Perry CC, Anderson OR, Williams RJP (1989) Biological minerals formed from strontium and barium sulphates. III. The morphology and crystallography of strontium sulphate crystals from the colonial radiolarian, *Sphaerozoum punctatum*. Philos Trans R Soc Lond [Biol] **238**: 223-233

Kimura M (1981) Estimation of evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci USA **78**: 454-458

Krabberød AK, Brate J, Dolven JK, Ose RF, Klaveness D, Kristensen T, Bjorklund KR, Schalchian-Tabrizi K (2011) Radiolaria divided into Polycystina and Spasmaria in combined 18S and 28S rDNA phylogeny. PLoS ONE 6: e23526 doi:10.1371/journal.pone.0023526

Liu H, Aris-Brosou S, Probert I, de Vargas C (2010) A time line of the environmental genetics of the haptophytes. Mol Biol Evol **271**: 161-176

López-García P, Philippe H, Gail F, Moreira D (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the mid-Atlantic ridge. Proc Natl Acad Sci USA **100**: 697-702

López-García P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature **409**: 603-607

Lovejoy C, Massana R, Pedrós-Alió C (2006) Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent seas. Appl Environ Microbiol **72**: 3085-3095

Marie D, Shi XL, Rigaut-Jalabert F, Vaulot D (2010) Use of flow cytometric sorting to better assess the diversity of small photosynthetic eukaryotes in the English Channel. FEMS Microbiol Ecol **72**: 165-178

Martin P, Allen JT, Cooper MJ, Johns DG, Lampitt RS, Sanders R, Teagle DAH (2010) Sedimentation of acantharian cysts in the Iceland Basin: Strontium as a ballast for deep ocean

particle flux, and implications for acantharian reproductive strategies. Limnol Oceanogr **55**: 604-614

Massera Bottazzi E and Andreoli MG (1978) "Blooming' di Acantari (Protozoa-Sarcodina) nell'Atlantico meridionale. Ateneo P se Ada Nat **14**: 465-475

Massera Bottazzi E, Andreoli MG (1981) ‘‘Blooming’’of Acantharia (Protozoa-Sarcodina) in the southern Atlantic Ocean. Quad Lab Technol **3**: 637

Massera Bottazzi E, Andreoli MG (1982) Distribution of adult and juvenile Acantharia (protozoa, Sarcodina) in the Atlantic- Ocean. J Plankton Res **4**: 757-777

Matsen FA, Kodner RB, Armbrust EV (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinformatics **11**: 538

Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene **71**: 491-499

Michaels AF (1988) Vertical distribution and abundance of Acantharia and their symbionts. Marine Biology **97**: 559-569

Michaels AF (1991) Acantharian abundance and symbiont productivity at the VERTEX seasonal station. J Plankton Res **13**: 399-418

Michaels AF, Caron DA, Swanberg NR, Howse FA, Michaels CM (1995) Planktonic sarcodines (acantharia, radiolaria, foraminifera) in surface waters near Bermuda - abundance, biomass and vertical flux. J Plankton Res **17**: 131-163

Mielck W (1907) Acanthometren von Neu-Pommern. Wissenschaftliche Meeresuntersuchungen, herausgegeben von der Kommission zur Untersuchung der deutschen Meere in Kiel und der Biologischen Anstalt auf Helgoland. Abteilung Kiel. Neue Folge **10**: 41-105

Moreira D, von der Heyden S, Bass D, López-García P, Chao E, Cavalier-Smith T (2007) Global eukaryote phylogeny: Combined small- and large-subunit ribosomal DNA trees support monophyly of Rhizaria, Retaria and Excavata. Mol Phylogenet Evol **44**: 255-66

Moretti S, Armougom F, Wallace IM, Higgins DG, Jongeneel CV, Notredame C (2007) The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. Nucleic Acids Res **35** (web server issue): W645-W648

Müller J (1856) Über die Thalassicollen, Polycystinen und Acanthometren des Mittelmeeres. Monatsber Königl Preuß Akad Wissensch Berlin **1856**: 474-503

Müller J (1859) Über die Thalassicollen, Polycystinen und Acanthometren des Mittelmeeres. Abhandl Königl Akad Wissensch Berlin **1858**: 1-62

Not F, Gausling R, AzamF, Heidelberg JF, Worden AZ (2007) Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. Environ Microbiol **9**: 1233-1252

Not F, Latasa M, Scharek R, Viprey M, Karleskind P, Balagué V, Ontoria-Oviedo I, Cumino A, Goetze E, Vaulot D, Massana R (2008) Protistan assemblages across the Indian Ocean, with a specific emphasis on the picoeukaryotes. Deep-sea Res I **55**: 1456-1473

Odum HT (1951) Notes on the strontium contain of seawater, celestite Radiolaria and strontianite snail shell. Science **114**: 211-213

Oka A, Endo K, Sashida K (2005) Molecular phylogeny of Acantharea (Actinopodea: Protista) based on small subunit ribosomal RNA gene sequences. Sci Rep Inst Geo sci Univ Tsukuba Sec B **26**: 13-22

Perry CC, Hughes NP (1990) Transmission and scanning electron microscopic evidence for cytoplasmic deposition of strontium sulphate crystals in colonial radiolaria. Philos Trans R Soc Lond [Biol] **329**: 81-86

Popofsky A (1904a) System und Faunistik der Acanthometriden der Planktonexpedition. Inaugural-Dissertation zur Erlangung der Doktorwürde der hohen philosophischen Fakultät der Königlichen Christian-Albrecht-Universtät zu Kiel, Kiel, pp 5-77

Popofsky A (1904b) Die nordischen Acantharien. Teil 1: Acanthometriden. Ergebnisse der Plankton-Expedition der Humboldt-Stiftung **3**(Lf. α): 3-158

Popofsky A (1906) Die Acantharia der Plankton-Expedition Teil II: Acanthophracta. Ergebnisse der Plankton-Expedition der Humboldt-Stiftung **3**(Lf. ß): 3-160

Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst Biol **53**: 793-808

Posada D (2008) jModelTest: Phylogenetic Model Averaging. Mol Biol Evol **25**: 1253-1256

Posada D (2009) Selection of Models of DNA Evolution with jModelTest. In Bioinformatics for DNA sequences analysis Methods in Molecular Biology Humana Press, New York, pp 93-112

Quaiser A, Zivanovic Y, Moreira D, Lopez-Garcia P (2010) Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. ISME **5**: 285-304

Reshetnyak VV (1981) Akantarii (Acantharia, Protozoa) Mirovogo okeana. Fauna SSSR **123**: 1–224. (in Russian)

Schewiakoff WT (1926) The Acantharia. Fauna e Flora del Golfo di Napoli **37**: 1-755

Schreiber B, Bottazzi EM, Fano-Schreiber A, Guerra F, Pelati, L (1962) Ricerche sulle presenza dello Sr nel plancton marino in rapporto alla ecologia degli Acantari. Pubbl. Staz. zool. Napoli **32** (suppl.): 400-426

Sinninghe Damste JS, Muyzer G, Abbas B, Rampen SW, Masse G, Allard WG, Belt ST, Robert JM, Rowland SJ, Moldowan JM, Barbanti SM, Fago FJ, Denisevich P, Dahl J, Trindade LAF, Schouten S (2004) The rise of the rhizosolenid diatoms. Science **304**: 584-587

Stoeck T, Taylor GT, Epstein SS (2003) Novel eukaryotes from the permanently anoxic Cariaco Basin (Caribbean Sea). Appl Environ Microbiol **69**: 5656-5663

Stoecker DK, Johnson MD, de Vargas C, Not F (2009) Acquired phototrophy in aquatic protists. Aquat Microb Ecol **57**: 279-310

Suzuki N and Aita Y (2011) Radiolaria: achievements and unresolved issues: taxonomy and cytology. Plankt & Benth Res **6**: 69-91

Swanberg NR and Caron DA (1991) Patterns of feeding in epipelagic oceanic plankton. J Plankton Res **13**: 287-312

Swofford DL (2001) PAUP* Phylogenetic Analysis Using Parsimony (*and other methods). Biosystems (version 4.0). Sinauer Associates, Sunderland, Mass.

Tan Z (1998) Order Acantharia, Order Spumellaria. In Editional Committee of Fauna Sinica, Academia Sinica (Ed) Fauna Sinica, Phylum Protozoa. Science Publishing, Beijing, pp 315. (in Chinese)

Terrado R, Vincent WF, Lovejoy C (2009) Mesopelagic protists: Diversity and succession in a coastal Arctic ecosystem. Aquat Microb Ecol **56**: 25-40

Tregouboff G (1953) Traité de Zoologie, ed. Grasse, Masson, Paris, pp 271-320

Wallace IM, O'Sullivan O, Desmond GH, Notredame C (2006) M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Res **34**: 1692-1699

Wilcock JR, Perry CC, Wilimas RJP, FRS, Mantoura RFC (1988) Crystallographic and morphological studies of the celestite skeleton of the acantharian species *Phyllostautus siculus*. Proc R Soc Lond B **233**: 393-405

Zettler LA, Sogin ML, Caron DA (1997) Phylogenetic relationships between the Acantharea and the Polycystinea: a molecular perspective on Haeckel's Radiolaria. Proc Natl Acad Sci USA **94**: 11411-11416

Supplementary Information

Table S1. Taxon sampling, GenBank accession numbers and sampling sites for the corresponding 18S and partial 28S rDNA sequences obtained from isolated Acantharia.

Organism name	code name	Sampling site	Accession number	
			18S	28S
Stauracantha orthostaura Haeckel	Ei 6	Gulf of Eilat		JN811227
Lonchostaurus rhombicus Haeckel	Ei 9	Gulf of Eilat		JN811228
Acanthostaurus purpurescens Haeckel	Ei 10	Gulf of Eilat		JN811229
Phyllostaurus cuspidatus Haeckel	Ei 14	Gulf of Eilat	JN811148	JN811230
Phyllostaurus cuspidatus claparedei Haeckel	Ei 20	Gulf of Eilat	JN811149	JN811231
Lonchostaurus rhombicus Haeckel	Ei 23	Gulf of Eilat	JN811150	JN811232
Amphilonche elongata Müller	Ei 24	Gulf of Eilat	JN811151	JN811233
Phyllostaurus cuspidatus Haeckel	Ei 25	Gulf of Eilat		JN811234
Amphilonche elongata Müller	Ei 27	Gulf of Eilat	JN811152	JN811235
Phyllostaurus siculus catervatus Haeckel	Ei 39	Gulf of Eilat	JN811153	JN811236
Phyllostaurus quadrangulus Haeckel	Ei 40	Gulf of Eilat	JN811154	JN811237
Xiphacantha quadridentata Müller	Ei 43	Gulf of Eilat	JN811155	JN811238
Acanthocyrta haeckeli Schewiakoff	Ei 47	Gulf of Eilat	JN811156	JN811239
Acanthocyrta haeckeli Schewiakoff	Ei 48	Gulf of Eilat	JN811157	JN811240
Lonchostaurus rhombicus Haeckel	Ei 49	Gulf of Eilat	JN811158	JN811241
Amphiastrus tetrapterus Haeckel	Ei 51	Gulf of Eilat	JN811159	JN811242
Phyllostaurus cuspidatus Haeckel	Ei 53	Gulf of Eilat	JN811160	JN811243
Amphilonche elongata Müller	Ei 54	Gulf of Eilat		JN811244
Phyllostaurus cuspidatus Haeckel	Ei 58	Gulf of Eilat		JN811245
Acanthoplegma sp. A Haeckel	Ei 59	Gulf of Eilat	JN811161	JN811246
Phyllostaurus cuspidatus Haeckel	Ei 68	Gulf of Eilat	JN811162	JN811247
Litholophus sp. B Haeckel	Ei 71	Gulf of Eilat	JN811163	JN811248
Phyllostaurus siculus Haeckel	Ei 75	Gulf of Eilat	JN811164	JN811249
Stauracantha orthostaura Haeckel	Ei 77	Gulf of Eilat		JN811250
Stauracantha orthostaura Haeckel	Ei 78	Gulf of Eilat		JN811251
Acanthometra fusca Müller	N1	Gulf of Naples		JN811252
Acanthometra fusca Müller	N2	Gulf of Naples	JN811165	JN811253
Dorataspis loricata Haeckel	Oki 4	Akajima island	JN811166	JN811254
Stauracantha orthostaura Haeckel	Oki 8	Akajima island	JN811167	JN811255
Acanthostaurus conacanthus Haeckel	Oki 10	Akajima island		JN811256
Lychnaspis giltschii Haeckel	Oki 11	Akajima island	JN811168	JN811257
Phyllostaurus cuspidatus Haeckel	Oki 15	Akajima island		JN811258
Xiphacantha quadridentata Müller	Oki 17	Akajima island	JN811169	JN811259
Amphibelone heteracanthum Haeckel	Oki 21	Akajima island	JN811170	JN811260
Gigartacon muelleri Haeckel	Oki 23	Akajima island	JN811171	JN811261
Acanthostaurus conacanthus Haeckel	Oki 24	Akajima island	JN811172	JN811262
Phyllostaurus siculus catervatus Haeckel	Oki 27	Akajima island	JN811173	JN811263
Acanthocolla cruciata Haeckel	Oki 28	Akajima island	JN811174	JN811264
Gigartacon muelleri Haeckel	Oki 30	Akajima island	JN811175	JN811265
Gigartacon fragilis Haeckel	Oki 33	Akajima island	JN811176	JN811266
Lychnaspis giltschii Haeckel	Oki 36	Akajima island	JN811177	JN811267
Phractopelta dorataspis Haeckel	Oki 42	Akajima island	JN811178	JN811268
Phyllostaurus cuspidatus Haeckel	Oki 43	Akajima island	JN811179	JN811269
Phractopelta sarmentosa Tchang & Tan	Oki 45	Akajima island	JN811180	JN811270
Lychnaspis giltschii Haeckel	Oki 46	Akajima island	JN811181	JN811271
Staurolithium sp. A Haeckel	Oki 47	Akajima island	JN811182	JN811272
Phyllostaurus cuspidatus Haeckel	Oki 48	Akajima island		JN811273
Lychnaspis giltschii Haeckel	Oki 49	Akajima island	JN811183	JN811274
Acanthocolla solidissima Popofsky	Oki 51	Akajima island	JN811184	JN811275
Lychnaspis giltschii Haeckel	Oki 54	Akajima island	JN811185	JN811276

Lychnaspis giltschii Haeckel	Oki 55	Akajima island	JN811186	JN811277
Amphilonche elongata heteracantha Haeckel	Oki 56	Akajima island		JN811278
Phyllostaurus cuspidatus Haeckel	Oki 57	Akajima island		JN811279
Phyllostaurus cuspidatus Haeckel	Oki 59	Akajima island	JN811187	JN811280
Acanthostaurus conacanthus Haeckel	Oki 65	Akajima island	JN811188	JN811281
Phyllostaurus quadrangulus Haeckel	Oki 66	Akajima island	JN811189	JN811282
Acanthometra pellucida Müller	Oki 67	Akajima island	JN811190	JN811283
Amphibelone heteracanthum Haeckel	Oki 68	Akajima island	JN811191	JN811284
Amphistaurus complanatus Haeckel	Oki 71	Akajima island	JN811192	JN811285
Phractopelta sarmentosa Tchang & Tan	Oki 72	Akajima island		JN811286
Larcidium dodecanthum Haeckel	Oki 73	Akajima island	JN811193	JN811287
Dorataspis loricata Haeckel	Oki 74	Akajima island	JN811194	JN811288
Acanthocolla cruciata Haeckel	Oki 77	Akajima island	JN811195	JN811289
Acanthometra pellucida Müller	Oki 78	Akajima island	JN811196	JN811290
Litholophus sp. A Haeckel	Oki 79	Akajima island	JN811197	JN811291
Stauracantha alata Müller	Oki 85	Akajima island	JN811198	JN811292
Xiphacantha quadridentata Müller	Oki 88	Akajima island		JN811293
Gigartacon fragilis Haeckel	Oki 91	Akajima island	JN811199	JN811294
Stauracantha orthostaura Haeckel	Oki 93	Akajima island	JN811200	JN811295
Xiphacantha quadridentata Müller	Oki 98	Akajima island		JN811296
Xiphacantha quadridentata Müller	Oki 100	Akajima island	JN811201	JN811297
Trizoma brandti Popofsky	Ros 6	English channel	JN811202	JN811298
Trizoma brandti Popofsky	Ros 7	English channel		JN811299
Trizoma brandti Popofsky	Ros 8	English channel		JN811300
Acanthometra fusca Müller	Ros 93	English channel		JN811301
Acanthometra fusca Müller	Ros 94	English channel		JN811302
Diploconus fasces Haeckel	Ta 60	Villefranche sur mer		JN811303
Diploconus fasces Haeckel	Ta 61	Villefranche sur mer		JN811304
Diploconus fasces Haeckel	Vil 17	Villefranche sur mer		JN811305
Unable to identify	Vil 20	Villefranche sur mer	JN811203	JN811306
Coleaspis vaginata Haeckel	Vil 25	Villefranche sur mer	JN811204	JN811307
Staurolithium sp. A Haeckel	Vil 32	Villefranche sur mer	JN811205	JN811308
Acanthocyrta haeckeli Schewiakoff	Vil 36	Villefranche sur mer		JN811309
Acanthocyrta haeckeli Schewiakoff	Vil 39	Villefranche sur mer	JN811206	JN811310
Gigartacon muelleri Haeckel	Vil 41	Villefranche sur mer		JN811311
Acanthostaurus purpurescens Haeckel	Vil 45	Villefranche sur mer		JN811312
Heteracon biformis Popofsky	Vil 47	Villefranche sur mer		JN811313
Acanthocyrta haeckeli Schewiakoff	Vil 51	Villefranche sur mer	JN811207	JN811314
Gigartacon muelleri Haeckel	Vil 52	Villefranche sur mer	JN811208	JN811315
Gigartacon muelleri Haeckel	Vil 53	Villefranche sur mer	JN811209	JN811316
Gigartacon muelleri Haeckel	Vil 61	Villefranche sur mer	JN811210	JN811317
Acanthocyrta haeckeli Schewiakoff	Vil 64	Villefranche sur mer	JN811211	JN811318
Heteracon biformis Popofsky	Vil 65	Villefranche sur mer	JN811212	JN811319
Amphibelone hydrotomica Haeckel	Vil 70	Villefranche sur mer		JN811320
Xiphacantha alata Müller	Vil 82	Villefranche sur mer	JN811213	JN811321
Phyllostaurus quadrangulus Haeckel	Vil 85	Villefranche sur mer	JN811214	JN811322
Coleaspis vaginata Haeckel	Vil 86	Villefranche sur mer	JN811215	JN811323
Xiphacantha quadridentata Müller	Vil 95	Villefranche sur mer	JN811216	JN811324
Xiphacantha alata Müller	Vil 96	Villefranche sur mer	JN811217	JN811325
Acanthostaurus purpurescens Haeckel	Vil 100	Villefranche sur mer	JN811218	JN811326
Gigartacon muelleri Haeckel	Vil 105	Villefranche sur mer	JN811219	JN811327
Gigartacon muelleri Haeckel	Vil 110	Villefranche sur mer	JN811220	JN811328
Gigartacon muelleri Haeckel	Vil 117	Villefranche sur mer	JN811221	JN811329
Litholophus sp. Haeckel (uncertained ident.)	Vil 119	Villefranche sur mer	JN811222	
Lithoptera fenestra Müller	Vil 122	Villefranche sur mer		JN811330
Heteracon biformis Popofsky	Vil 126	Villefranche sur mer	JN811223	JN811331
Acanthostaurus purpurescens Haeckel	Vil 127	Villefranche sur mer	JN811224	JN811332

Table S2: Data set for pplacer analysis including environmental and single-cell sequences from GenBank, and partial sequences from this study. Associated information is indicated in this table, such as the position of the sequences within the pplacer topology (number of the branch and reference clade in which the sequence is placed), the associated clade in the phylogeny Figure 5, their accession number, as well as the depth and location of sampling. Ref Partial = partial sequence of isolated Acantharia from this study; Env = environmental sequences from GenBank; GenBank Ref = sequences of isolates from previous studies.

Branch number	Origin	Accession number	Sequence name	Associated clade in pplacer (Fig. 3)	Depth (m)	Location	Reference	Associated clade in Fig.5
0	Ref partial		Vil 56-*Acanthochiasma fusiforme*	B	surface	Med sea	this study	
0	Ref partial		Vil 31-*Acanthochiasma fusiforme*	B	surface	Med sea	this study	
1	Env	ZZ0027843	C9_65m_137	B	surface	Med sea	Viprey et al *subm.*	Clade B
4	Env	DQ918099	ENVP107.00002	B	2500	Sargasso Sea and Gulf Stream	Countway et al 2007	
5	GenBank ref	AF063242	Symphiacanthid 211	C	surface	North atlantic - Bermuda coast	Zettler et al. 1997	Clade C
5	GenBank ref	AB178585	*Amphiacon denticulatus*	C	surface	North East pacific	Oka et al. 2005	Clade C
5	Env	DQ918042	ENVP10203.00243	C	surface	Sargasso Sea and Gulf Stream	Countway et al 2007	
6	Env	AY665095	SCM27_C24	C	DCM	Sargasso sea	Ambrust et al. Unpubl.	Clade C
7	Env	ZZ0032829	C5_55m_21	C	surface	Med sea	Viprey et al *subm.*	Clade C
7	Env	ZZ0026406	CM_110m_89	C	surface	Med sea	Viprey et al *subm.*	Clade C
7	Env	GU246587	MO010.880.0150	C	880	Pacific San Pedro Channel	Gilg et al 2009	Clade C
7	Env	GQ383195	MO010_150.00004	C	150	Pacific San Pedro Channel	Schnetzer et al 2011	
7	GenBank ref	GU246575	Chaunacanthid sp. 6200	C	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade C
8	Env	GQ383281	MO010_500.00199	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383295	MO010_500.00230	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383300	MO010_500.00248	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383396	MO010_880.00133	C	880	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383014	MO010_150.00111	C	150	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383199	MO010_150.00050	C	150	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383207	MO010_150.00072	C	150	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383226	MO010_500.00116	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383232	MO010_500.00127	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383235	MO010_500.00131	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383266	MO010_500.00172	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383303	MO010_500.00253	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383315	MO010_500.00276	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383321	MO010_500.00300	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383324	MO010_500.00306	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383342	MO010_500.00332	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383368	MO010_500.00049	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383411	MO010_880.00154	C	880	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383484	MO010_880.00305	C	880	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383198	MO010_150.00049	C	150	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383307	MO010_500.00261	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	GQ383379	MO010_880.00103	C	880	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	AY046840	C3_E010	C	2004	Pacific Guaymas vent field	Edgcomb et al 2002	
8	Env	ZZ0027847	C9_65m_141	C	surface	Med sea	Viprey et al *subm.*	Clade C
8	Env	FJ169731	CS050S22	C	161	Arctic - Franklin Bay	Terrado et al.2009	
8	Env	FJ169733	CS050S32	C	161	Arctic - Franklin Bay	Terrado et al.2009	
8	Env	FJ169714	CS060L34	C	219	Arctic - Franklin Bay	Terrado et al.2009	
8	Env	GQ383253	MO010_500.00155	C	500	Pacific San Pedro Channel	Schnetzer et al 2011	
8	Env	FJ169736	CS050L21	C	161	Arctic - Franklin Bay	Terrado et al.2009	
8	Env	FJ775665	CS060S27	C	219	Arctic - Franklin Bay	Terrado et al.2009	
8	Env	FJ169728	CS123L18	C	223	Arctic - Franklin Bay	Terrado et al.2009	
8	Env	FJ169734	CS050L17	C	161	Arctic - Franklin Bay	Terrado et al.2009	
8	Env	FJ169722	CS163S48	C	219	Arctic - Franklin Bay	Terrado et al.2009	
8	Env	GU246578	MO010.500.0040	C	500	Pacific San Pedro Channel	Gilg et al 2009	Clade C
8	Env	GU246581	MO010.500.0116	C	500	Pacific San Pedro Channel	Gilg et al 2009	Clade C
8	Env	GU246577	MO010.500.0038	C	500	Pacific San Pedro Channel	Gilg et al 2009	Clade C
8	Env	GU246580	MO010.500.0049	C	500	Pacific San Pedro Channel	Gilg et al 2009	Clade C
8	Env	AY665098	SCM15_C12	C	DCM	Sargasso sea	Ambrust et al. Unpubl.	Clade C
8	Env	GU246586	MO010.880.0133	C	880	Pacific San Pedro Channel	Gilg et al 2009	Clade C
9	Env	FJ169720	CS163S39	C	219	Arctic - Franklin Bay	Terrado et al.2009	
9	Env	EU562092	IND72.15	C	75	Indian ocean	Not et al 2008	
9	Env	FJ169713	CS060L18	C	219	Arctic - Franklin Bay	Terrado et al.2009	
20	GenBank ref	AF018158	Chaunacanthid_218	C	surface	North atlantic - Bermuda coast	Zettler et al. 1997	Clade C
20	GenBank ref	AF063241	Chaunacanthid 217	C	surface	North atlantic - Bermuda coast	Zettler et al. 1997	Clade C
20	Env	ZZ0026365	CM_110m_77	C	surface	Med sea	Viprey et al *subm.*	Clade C
20	Ref partial		Ta 56	C	surface	Med sea	this study	
25	Env	GQ383194	MO010_150.00039	C	150	Pacific San Pedro Channel	Schnetzer et al 2011	
36	GenBank ref	GU246567	Acantharian sp. 6201	C	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
37	Env	DQ918165	ENVP107.00119	**no clade**	2500	Sargasso Sea and Gulf Stream	Countway et al 2007	
37	Env	ZZ0027340	C1_30m_220	**no clade**	surface	Med sea	Viprey et al *subm.*	Acanth II
39	Ref partial	JN811202	Ros 6 -*Trizona brandti*	D	surface	Roscoff	this study	
41	Env	ZZ0027833	C9_65m_129	D	surface	Med sea	Viprey et al *subm.*	Clade D
41	Env	ZZ0027756	CM_50m_177	D	surface	Med sea	Viprey et al *subm.*	Clade D
49	Env	GQ383465	MO010_880.00263	**no clade**	880	Pacific San Pedro Channel	Schnetzer et al 2011	
49	Env	GQ383447	MO010_880.00228	**no clade**	880	Pacific San Pedro Channel	Schnetzer et al 2011	
49	Env	GQ383478	MO010_880.00290	**no clade**	880	Pacific San Pedro Channel	Schnetzer et al 2011	
49	Env	ZZ0027698	CM_50m_123	**no clade**	surface	Med sea	Viprey et al *subm.*	Acanth IV
49	Env	ZZ0027025	C9_65m_76	**no clade**	surface	Med sea	Viprey et al *subm.*	Acanth IV
58	Env	ZZ0027778	CM_50m_9	E	surface	Med sea	Viprey et al *subm.*	Clade E
58	Env	ZZ0033017	C5_25m_53	E	surface	Med sea	Viprey et al *subm.*	Clade E
59	GenBank ref	GU246574	*Acanthometra* sp. 3 813	E	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade E
69	GenBank ref	AB178587	*Hexaconus serratus*	E	surface	North East pacific	Oka et al. 2005	Clade E
73	Ref partial		Ta61 *Diploconus fasces*	E	surface	Med sea	this study	
73	Ref partial		Ta 53 *Diploconus fasces*	E	surface	Med sea	this study	
73	Ref partial		Vil 19 *Diploconus fasces*	E	surface	Med sea	this study	
73	Ref partial		Vil 17 *Diploconus fasces*	E	surface	Med sea	this study	
73	Env	ZZ0024770	C3_5m_47	E	surface	Med sea	Viprey et al *subm.*	Clade E
73	Env	ZZ0027307	C1_30m_190-2	E	surface	Med sea	Viprey et al *subm.*	Clade E
73	GenBank ref	AB178588	*Hexaconus serratus*	E	surface	North East pacific	Oka et al. 2005	Clade E
74	Env	FJ775649	CS123S41	E	225	Arctic - Franklin Bay	Terrado et al.2009	
75	Ref partial		Vil 122-*Lithoptera fenestra*	F	surface	Med sea	this study	
76	Ref partial		Ei 45- *Amphilonche elongata*	F	surface	Red sea	this study	
76	Ref partial		Ei 54 -*Amphilonche elongata*	F	surface	Red sea	this study	
76	Ref partial		Oki 56 -*Amphilonche elongata*	F	surface	North East pacific	this study	
79	GenBank ref	AB178581	*Amphibelone cultellata*	F	surface	North East pacific	Oka et al. 2005	Clade F
79	GenBank ref	AB178582	*Amphibelone anomala*	F	surface	North East pacific	Oka et al. 2005	Clade F
79	Ref partial		Vil 70 - *Amphibelone* sp	F	surface	Med sea	this study	
81	GenBank ref	AB178584	*Amphibelone anomala*	F	surface	North East pacific	Oka et al. 2005	Clade F
81	Env	ZZ0026285	C3_5m_118	F	surface	Med sea	Viprey et al subm.	Clade F
81	GenBank ref	AB178580	*Amphibelone cultellata*	F	surface	North East pacific	Oka et al. 2005	Clade F
81	GenBank ref	AB178583	*Amphibelone anomala*	F	surface	North East pacific	Oka et al. 2005	Clade F

81	Env	ZZ0026966	C9_5m_62	**F**	surface	Med sea	Viprey et al subm.	
81	Env	FN598261	clone BIO3_F9	**F**	500	South Pacific	Sauvadet et al 2010	
83	GenBank ref	AB178586	Dorataspid	**F**	surface	South Pacific	Oka et al. 2005	Clade F
83	GenBank ref	AF063239	Arthracanthid 206	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
83	Env	AJ402332	OLI011_75m_15-5	**F**	75	Equat pacific	Moon-van der Staay et al 2001	Clade F
83	Env	AJ402333	OLI011_75m_16-5	**F**	75	Equat pacific	Moon-van der Staay et al 2001	Clade F
83	GenBank ref	GU246569	Acantharian sp. 6205	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
83	GenBank ref	GU246570	Acanthometra sp. 1 6202	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
83	GenBank ref	GU246572	Acanthometra sp. 1 7201	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
87	Env	FJ431859	RA071004T.056 18S	**F**	surface	Roscoff	Marie et al 2010	
87	GenBank ref	GU246566	*Dorataspis* sp. 813	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
87	GenBank ref	GU246573	*Acanthometra* sp. 2 7202	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
87	Env	ZZ0025053	CM_50m_23	**F**	surface	Med sea	Viprey et al subm.	Clade F
87	Env	ZZ0026231	C1_80m_139	**F**	surface	Med sea	Viprey et al subm.	Clade F
87	Env	ZZ0026404	CM_110m_73	**F**	surface	Med sea	Viprey et al subm.	Clade F
87	Env	ZZ0027962	CD_15m_4	**F**	surface	Med sea	Viprey et al subm.	Clade F
88	GenBank ref	GU246571	Acanthometra sp. 1 6203	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
89	Ref partial		Ros 93-*Acanthometra fusca*	**F**	surface	off Roscoff	this study	
91	Env	HM474660	T149_W01D.011	**F**	surface	South East Pacific	Lepere et al 2010	
96	Env	FJ431871	RA071004T.071	**F**	surface	off Roscoff	Marie et al 2010	
122	GenBank ref	AF063240	Acanthometra 205	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
122	Env	ZZ0026204	C3_5m_87	**F**	surface	Med sea	Viprey et al subm.	Clade F
128	Env	ZZ0026823	CD_50m_135	**F**	surface	Med sea	Viprey et al subm.	Clade F
128	Env	ZZ0027966	CD_15m_53	**F**	surface	Med sea	Viprey et al subm.	Clade F
128	Env	ZZ0032728	C5_25m_41	**F**	surface	Med sea	Viprey et al subm.	Clade F
134	GenBank ref	AF018159	*Haliommatidium* BBSR-235	**F**	surface	Pacific San Pedro Channel	Gilg et al 2009	Clade F
134	Env	ZZ0027777	CM_50m_99	**F**	surface	Med sea	Viprey et al subm.	Clade F
134	Ref partial		Ei77-*Stauracantha orthostaura*	**F**	surface	Red sea	this study	
148	Env	GQ383317	MO010_500.00280	**F**	500	Pacific San Pedro Channel	Gilg et al 2009	
149	Env	ZZ0027981	CD_15m_67	**no clade**	surface	Med sea	Viprey et al subm.	Acanth IV
149	Env	EU562128	IND72.51	**no clade**	75	Indian ocean	Not et al 2008	
149	Env	ZZ0027714	CM_50m_138	**no clade**	surface	Med sea	Viprey et al subm.	Acanth IV
149	Env	GQ383489	MO010_880.00323	**no clade**	880	Pacific San Pedro Channel	Schnetzer et al 2011	
149	Env	ZZ0028071	CD_50m_65	**no clade**	surface	Med sea	Viprey et al subm.	Acanth IV
149	Env	GU246576	MO010.500.0014	**no clade**	500	Pacific San Pedro Channel	Gilg et al 2009	Acanth III
149	Env	ZZ0025015	C3_5m_35	**no clade**	surface	Med sea	Viprey et al subm.	Acanth IV
149	Env	DQ918566	ENVP223.00150	**no clade**	2500	Pacific San Pedro Channel	Countway et al 2007	
149	Env	FJ032647	Ma121 1A14	**no clade**	550	Marmara sea	Quaiser et al 2010	Acanth I
149	Env	GU246582	MO010.500.0307	**no clade**	500	Pacific San Pedro Channel	Gilg et al 2009	Acanth I
149	Ref partial	JN811222	Vil 119	**no clade**	surface	Med sea	this study	
149	Env	EF172909	SSRPC45	**no clade**	3000	Sargasso Sea	Not et al 2007	Acanth I
149	Env	ZZ0026354	CM_110m_68	**no clade**	surface	Med sea	Viprey et al subm.	Acanth III
149	Env	EF172908	SSRPC15	**no clade**	3000	Sargasso Sea	Not et al 2007	Acanth I
149	Env	EF172910	SSRPC07	**no clade**	500	Sargasso Sea	Not et al 2007	Acanth I
149	Env	EF172929	SSRPC49	**no clade**	3000	Sargasso Sea	Not et al 2007	Acanth I
149	Env	ZZ0032850	C5_55m_23	**no clade**	surface	Med sea	Viprey et al subm.	Acanth III
149	Env	ZZ0026835	CD_50m_144	**no clade**	surface	Med sea	Viprey et al subm.	Acanth III
149	Env	ZZ0026952	C9_5m_69	**no clade**	surface	Med sea	Viprey et al subm.	Acanth II
149	Env	GQ382988	MO010_42.00070	**no clade**	surface	Pacific San Pedro Channel	Schnetzer et al 2011	
149	Env	ZZ0026187	C3_5m_109	**no clade**	surface	Med sea	Viprey et al subm.	Acanth III
149	Env	EF172841	SSRPB70	**no clade**	3000	Sargasso Sea	Not et al 2007	Acanth I
149	Env	GU246579	MO010.500.0043	**no clade**	500	Pacific San Pedro Channel	Gilg et al 2009	Acanth I
149	Env	GU246591	MO010.880.323	**no clade**	880	Pacific San Pedro Channel	Gilg et al 2009	Acanth II
149	Env	EF172842	SSRPB59	**no clade**	500	Sargasso Sea	Not et al 2007	Acanth I
149	Env	ZZ0026343	C3_5m_146	**no clade**	surface	Med sea	Viprey et al subm.	Acanth III
149	Env	EU562151	IND2.9	**no clade**	25	Indian ocean	Not et al 2008	
149	Env	ZZ0024893	C1_80m_101	**no clade**	surface	Med sea	Viprey et al subm.	Acanth II
149	Env	GQ382969	MO010_42.00051	**no clade**	surface	Pacific San Pedro Channel	Schnetzer et al 2011	
149	Env	DQ918138	ENVP107.00070	**no clade**	2500	Sargasso Sea and Gulf Stream	Countway et al 2007	
149	Env	EF172802	SSRPB51	**no clade**	500	Sargasso Sea	Not et al 2007	Acanth II
149	Env	DQ918602	ENVP223.00197	**no clade**	2500	Sargasso Sea and Gulf Stream	Countway et al 2007	
149	Env	ZZ0032831	C5_55m_37	**no clade**	surface	Med sea	Viprey et al subm.	Acanth II
149	Env	AJ402342	OLI011_75m_32-5	**no clade**	75	Equat pacific	Moon-van der Staay et al 2001	Acanth II
149	Env	DQ918656	ENVP223.00269	**no clade**	2500	Sargasso Sea and Gulf Stream	Countway et al 2007	
149	Env	ZZ0026879	CD_50m_116	**no clade**	surface	Med sea	Viprey et al subm.	Acanth II
149	Env	DQ918230	ENVP107.00241	**no clade**	2500	Sargasso Sea and Gulf Stream	Countway et al 2007	
149	Env	ZZ0027980	CD_15m_66	**no clade**	surface	Med sea	Viprey et al subm.	Acanth III
150	Ref partial	JN811161	Ei 59 - *Acanthoplegma* sp	**no clade**	surface	Red sea	this study	
150	Env	AF290072	DH147_EKD17	**no clade**	2000	Antarctic Drake passage	Lopez-Garcia et al 2001	Acanth II
150	Env	HM103402	Ma101_1E_23	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	DQ314820	NW614_35	**no clade**	50	Canada Basin Arctic	Lovejoy et al 2006	
150	Env	HM103391	Ma101_1E_13	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	DQ314823	NW614_52	**no clade**	50	Canada Basin Arctic	Lovejoy et al 2006	
150	Env	DQ314822	NW614_60	**no clade**	50	Canada Basin Arctic	Lovejoy et al 2006	
150	Env	HM103395	Ma101_1E_17	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	AY046858	C3_E029	**no clade**	2004	Pacific Guaymas vent field	Edgcomb et al 2002	
150	Env	HM103418	Ma101_1E_39	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	HM103411	Ma101_1E_31	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	HM103427	Ma101_1E_5	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	FJ032649	Ma121 1A29	**no clade**	500	Marmara sea	Quaiser et al 2010	Clade B
150	Env	HM103419	Ma101_1E_40	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	GU246585	MO010.880.0119	**no clade**	880	Pacific San Pedro Channel	Viprey et al subm.	Clade B
150	Env	HM103406	Ma101_1E_27	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	DQ314821	NW614_49	**no clade**	50	Canada Basin Arctic	Lovejoy et al 2006	Clade B
150	Env	HM103399	Ma101_1E_20	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	HM103409	Ma101_1E_3	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	HM103417	Ma101_1E_38	**no clade**	1000	Marmara sea	Quaiser et al 2010	
150	Env	ZZ0027129	C9_65m_200	**no clade**	surface	Med sea	Viprey et al subm.	Clade B
150	Env	AY046843	C3_E013	**no clade**	2004	Pacific Guaymas vent field	Edgcomb et al 2002	

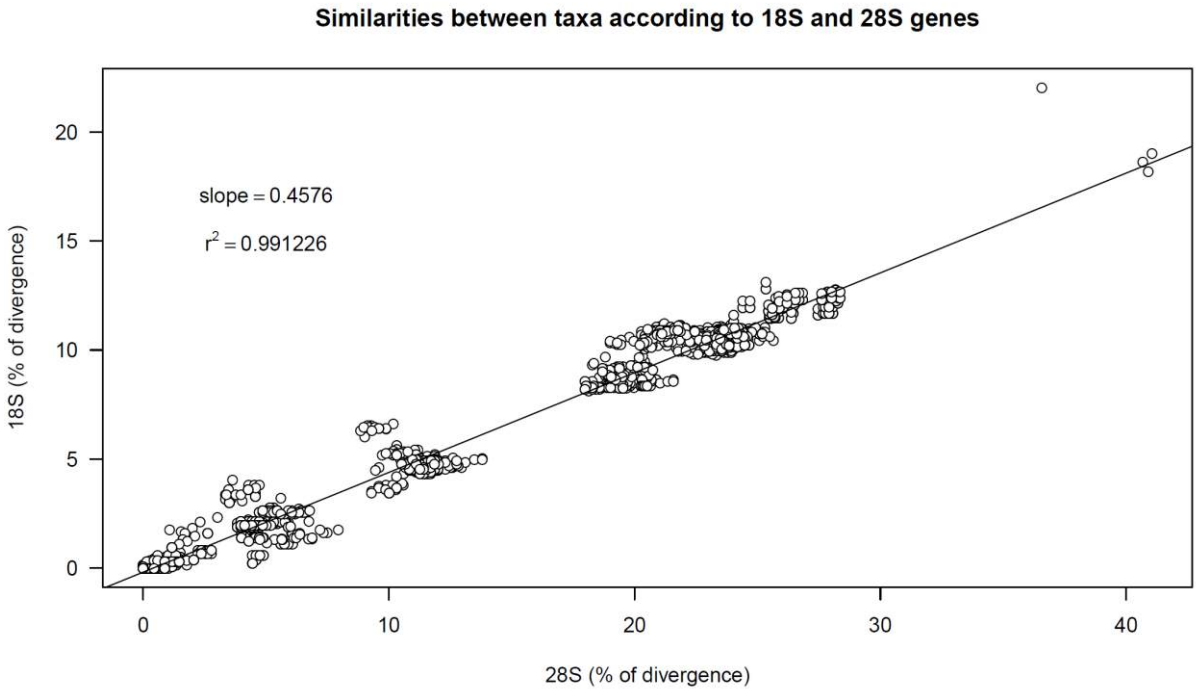


Figure S1. Comparison of evolutionary rate between the 18S and partial 28S rDNA markers.

References in supplementary

Lepere C, Demura M, kawachi M, Romac S, Probert I, Vaulot D (2011) Whole-genome amplification (WGA) of marine photosynthetic eukaryote populations. FEMS Microbial Ecol 1-11

Moon-van der Staay SY, De Wachter R, Vaulot D (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. Nature **409**: 607-610

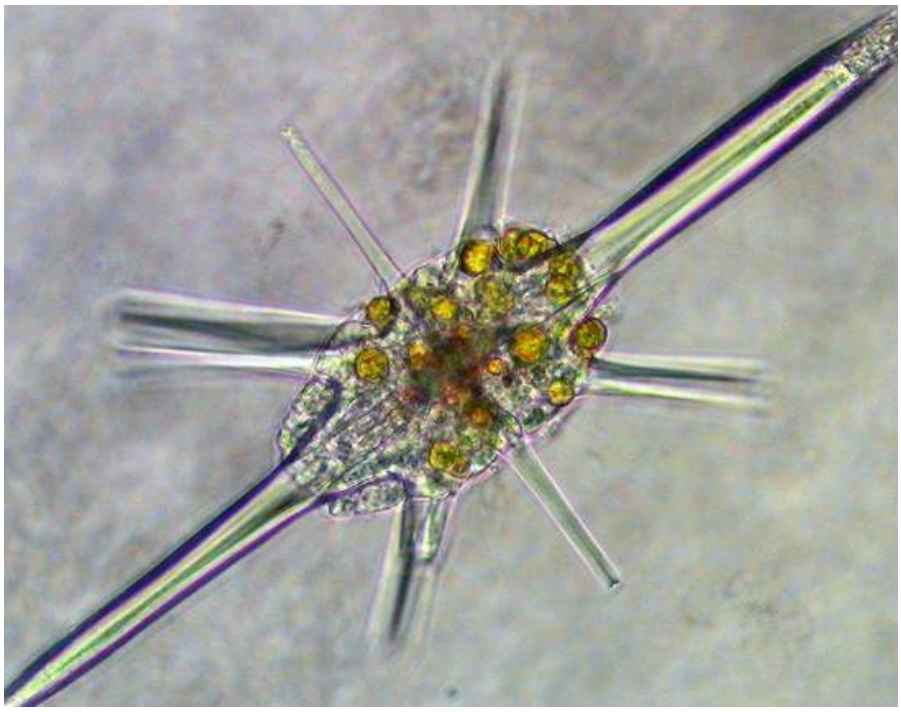
Sauvadet AL, Gobet A, Guillou L (2010) Comparative analysis between protist communities from the deep-sea pelagic ecosystem and specific deep hydrothermal habitats. Environ Microbiol **12**(11):2946-64

Schnetzer A, Moorthi SD, Countway PD, Gast RJ, Gilg IC, Caron D (2011) Depth matters: microbial eukaryote diversity and community structure in the eastern North Pacific revealed through environmental gene libraries. Deep-Sea R I **58**: 16-26

Chapitre II

Photosymbiosis in Acantharia

A) An original mode of symbiosis in the open ocean plankton



An original mode of symbiosis in open ocean plankton

Johan Decelle[1], Ian Probert[2], Lucie Bittner[1], Yves Desdevises[3], Sébastien Colin[1], Colomban de Vargas[1], Martí Galí [4], Rafel Simó[4], Fabrice Not[1]

[1] UPMC Université Paris 6 and CNRS, UMR 7144, EPPO, Station Biologique, Place G. Teissier 29680 Roscoff, France.

[2] UPMC Université Paris 6 and CNRS, FR2424, Station Biologique, Place G. Teissier 29680 Roscoff, France.

[3]UPMC Université Paris 6 and CNRS, UMR 7628, Modèles en Biologie Cellulaire et Evolutive, Laboratoire Arago, Avenue Fontaulé, BP44, 66651 Banyuls-sur-Mer, France.

[4]Institut de Ciències del Mar (CSIC), Passeig Maritim de la Barceloneta 37-49, 08003, Barcelona, Catalonia, Spain.

Published in ***Proceedings of the National Academy of Sciences USA*, 2012 vol.109 no.44 18000-18005**

Abstract

Symbiotic relationships are widespread in nature and are fundamental for ecosystem functioning and the evolution of biodiversity. In marine environments, photosymbiosis with microalgae is best known for sustaining benthic coral reef ecosystems. Despite the importance of oceanic microbiota in global ecology and biogeochemical cycles, symbioses are poorly characterized in open ocean plankton. Here, we describe a widespread symbiotic association between Acantharia, biomineralizing microorganisms that are abundant grazers in plankton communities, and members of the haptophyte genus *Phaeocystis* that are cosmopolitan bloom-forming microalgae. Cophylogenetic analyses demonstrate that symbiont biogeography, rather than host taxonomy, is the main determinant of the association. Molecular dating places the origin of this photosymbiosis in the Jurassic (ca 175 Mya), a period of accentuated marine oligotrophy. Measurements of intracellular dimethylated sulfur indicate that the host likely profits from antioxidant protection provided by the symbionts as an adaptation to life in transparent oligotrophic surface waters. In contrast to terrestrial and marine symbioses characterized to date, the symbiont reported in this association is extremely abundant and ecologically active in its free-living phase. In the vast and barren open ocean, partnership with photosymbionts that have extensive free-living populations is likely an advantageous strategy for hosts that rely on such interactions. Discovery of the Acantharia-*Phaeocystis* association contrasts with the widely held view that symbionts are specialized organisms that are rare and ecologically passive outside the host.

Introduction

Symbiosis, whereby different biological species live together in close and long-term interaction, plays key ecological roles in terrestrial and marine ecosystems and is regarded as a major source of evolutionary innovation (1). Photosymbiosis, whereby microalgae live intracellularly within a heterotrophic host organism, is typically considered mutually beneficial for the two partners: the symbiont provides photosynthetically derived products to the host, which in turn maintains a sheltered and relatively nutrient-rich environment for the symbiont (2, 3). Combining heterotrophy, autotrophy, and often biomineralization, photosymbioses profoundly impact food webs and ecosystem functioning in shallow coastal waters, particularly in oligotrophic settings (4). The most familiar case of photosymbiosis is the association of reef-building corals with 'zooxanthellae' (dinoflagellates of the genus *Symbiodinium*), but other hosts include molluscs, anemones, acoel flatworms, sponges and various protists (4-7). Despite early recognition of the existence of photosymbioses in the plankton realm (8) and more recent emphasis on the fundamental ecological and evolutionary significance of symbiosis, remarkably little attention has been given to the diversity, function and ecological impact of symbioses in open ocean plankton. In most known photosymbioses, the algal symbiont must be acquired from the surrounding environment during each host generation (horizontal transmission), implying the existence of a free-living phase of the symbiont. In shallow waters, free-living populations of known symbionts are typically elusive and often undetectable in the water column, but in the case of *Symbiodinium* are rather concentrated in surface sediments surrounding the coral hosts (9, 10). In open ocean surface waters, where microbiotia is typically very diluted, it is not clear what strategy hosts employ to ensure repeated capture of appropriate symbionts in successive generations.

Large heterotrophic amoeboid protists, such as Radiolaria and Foraminifera from the eukaryote supergroup Rhizaria, are known to host endosymbiotic microalgae in oligotrophic oceanic surface waters (11, 12). Molecular investigations associated with culturing methods have shown that symbionts can be dinoflagellates, prasinophytes or haptophytes (13-15), but in most cases their precise identity, host specificity and biogeography are unknown. The fact that photosymbiosis gave rise to light-harvesting chloroplasts and the spread of photosynthesis across the eukaryotic tree of life (16) means that investigation of modern photosymbioses can help understanding these key evolutionary processes.

The Acantharia are radiolarians characterized by possessing an endoskeleton of barium-enriched strontium sulfate (celestite). They are widely distributed throughout the world's oceans, and are particularly abundant in oligotrophic open ocean waters. Contributing up to 40% of total zooplankton biomass (17), they typically outnumber other biomineralizing protists like Foraminifera and Polycystinea (18). Classically known as grazers in marine ecosystems (19), certain acantharians have long been recognized to also participate, sometimes significantly (20), to primary production by harboring endosymbiotic microalgae. Symbiotic Acantharia are abundant in surface waters, whereas their non-symbiotic relatives typically inhabit deeper, down to the mesopelagic waters (21, 22). Photosymbiosis in Acantharia is associated with significant

morphological innovation and an increased complexity of the celestite skeleton, symbiotic species having thicker spicules, a tighter central association and a more robust shell than non-symbiotic species (23). Acantharian symbionts were taxonomically assigned three decades ago to the division Haptophyta based on an *in hospite* study of their ultrastructure (24), but their exact identity has never been elucidated.

Results and Discussion

Diversity and specificity of the Acantharia-*Phaeocystis* symbiosisWe isolated individual symbiont-bearing acantharian cells from 7 oceanic regions (Fig. 1) and subsequently PCR amplified genetic markers from the host (18S and 28S rDNA) and the symbionts (18S and 28S rDNA, rbcL and psbA). For more than 100 acantharian specimens analyzed, representing 25 distinct morphospecies and the majority of known symbiotic families, we found that symbionts consistently belonged to the well-known haptophyte genus *Phaeocystis*. The microalga *Phaeocystis* is one of the most extensively studied taxa of marine phytoplankton, but has never been reported to occur in symbiosis. Free-living *Phaeocystis* are ubiquitous from poles to tropics and from coastal to open ocean waters (25). They are recognized both as harmful algae and as one of the few keystone phytoplankton genera that shape the structure and functioning of marine ecosystems (26). *Phaeocystis* are not only major contributors to the global carbon budget (27), but they also impact sulfur cycling by producing substantial amounts of dimethylsulfoniopropionate (DMSP) and its volatile catabolite dimethylsulfide (DMS), a climatically active trace gas emitted from the ocean (28). In coastal areas, blooms of *Phaeocystis* are detrimental to the growth and reproduction of shellfish and zooplankton, and strongly impact human activities such as fisheries, aquaculture and tourism (25).

Multi-gene phylogenetic analyses including sequences from both symbiotic and cultured free-living *Phaeocystis* revealed that symbionts belong to different *Phaeocystis* species (Fig. 1). Most *Phaeocystis* species known as free-living forms were found in association with Acantharia, including species such as *P.globosa* and *P. antarctica* that are known to form massive blooms consisting of colonies of cells maintained in a gelatinous matrix. In contrast to the vast majority of terrestrial and marine symbiotic associations described to date, the genetic footprint of *Phaeocystis* symbionts was strictly identical or very similar to species that are known to be very abundant in the free-living state. Results from our multi-gene analysis confirmed previously reported phylogenetic relationships between described *Phaeocystis* species (29), but also revealed additional putative cryptic species. Notably, two *Phaeocystis* clades (named “Phaeo1” and “Phaeo2”) contained only sequences retrieved from symbionts or culture-independent environmental surveys in the < 5µm size fraction (Figs. S2 and S3), demonstrating a broader genetic diversity than previously reported for this extensively studied microalgal genus. Different acantharian species from a single location typically live in symbiosis with the same *Phaeocystis* genotype, while the same host species living in distinct habitats can harbor different *Phaeocystis* genotypes (Fig. 1). For instance, most acantharians from the Antarctic live with the symbiont *P. antarctica*, whereas those from the Mediterranean Sea and Red Sea mostly associate with *P.*

cordata. The strong influence of symbiont biogeography on the association was confirmed by a Mantel test for each genetic marker (Table S1), which revealed a highly significant correlation between the genetic and geographical distances for symbionts (Pearson $R > 0.40$; $P = 0.001$), while no such correlation was found for hosts (Pearson $R < 0.03$; $P > 0.28$).

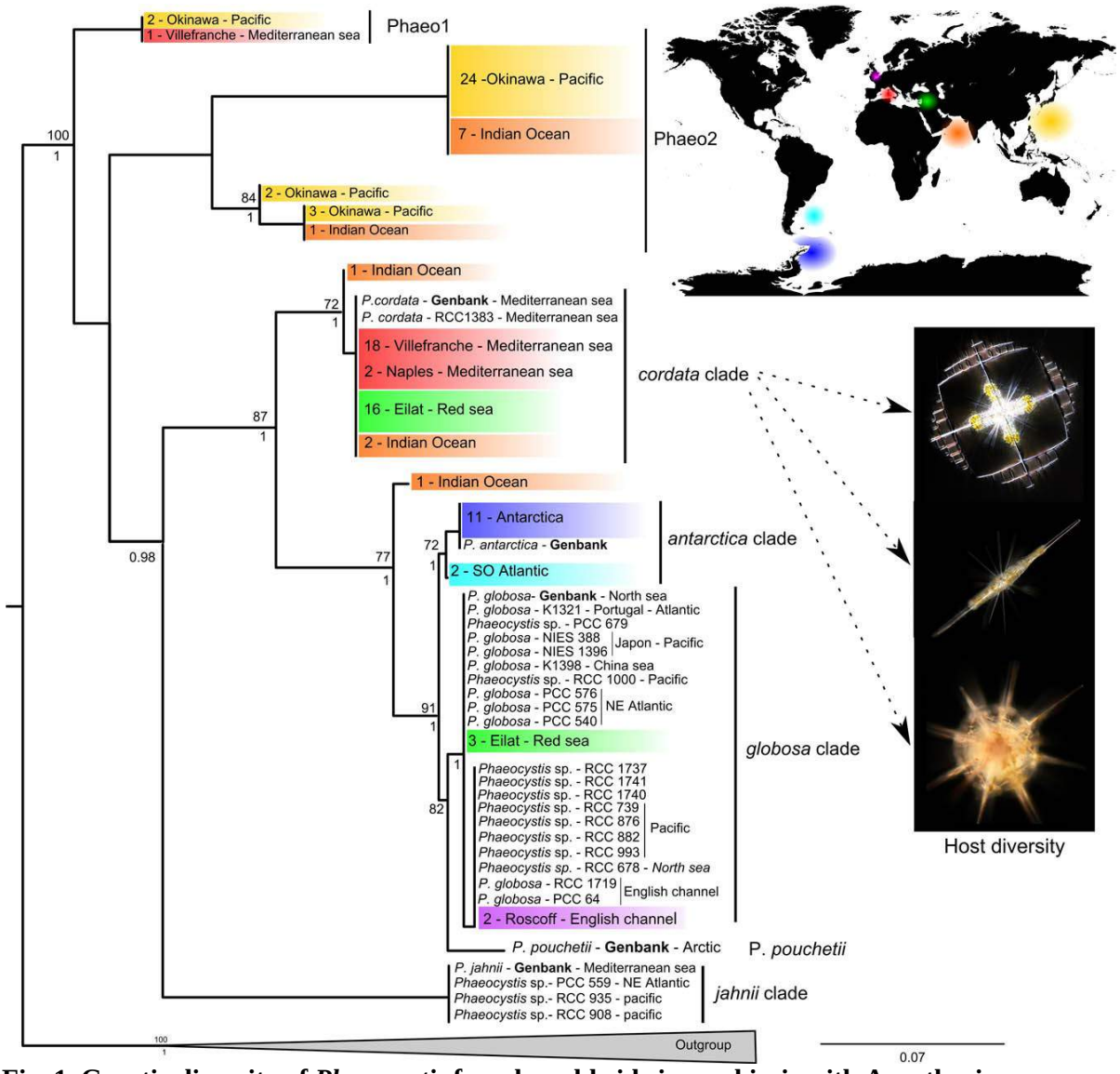


Fig. 1. Genetic diversity of *Phaeocystis* found worldwide in symbiosis with Acantharia.
RAxML phylogenetic reconstruction of the genus *Phaeocystis* based on the concatenation of the ribosomal genes 18S and 28S rDNA and the plastidial genes psbA and rbcL (131 taxa and 3.1 kb aligned positions). Sequences of *Phaeocystis* in symbiosis are colored according to their geographic origin, and the number of host specimens examined (representing different species) is indicated. Cultures of free-living *Phaeocystis* and their corresponding geographic origin are in black. All sequences were produced in this study except for those indicated in bold (from GenBank). Bootstrap values ≥ 60% and Bayesian posterior probability ≥ 0.7 are given above and beneath the node, respectively. The outgroup contains five sequences from other members of the Haptophyta. See SI Materials and Methods for one-gene phylogenies.

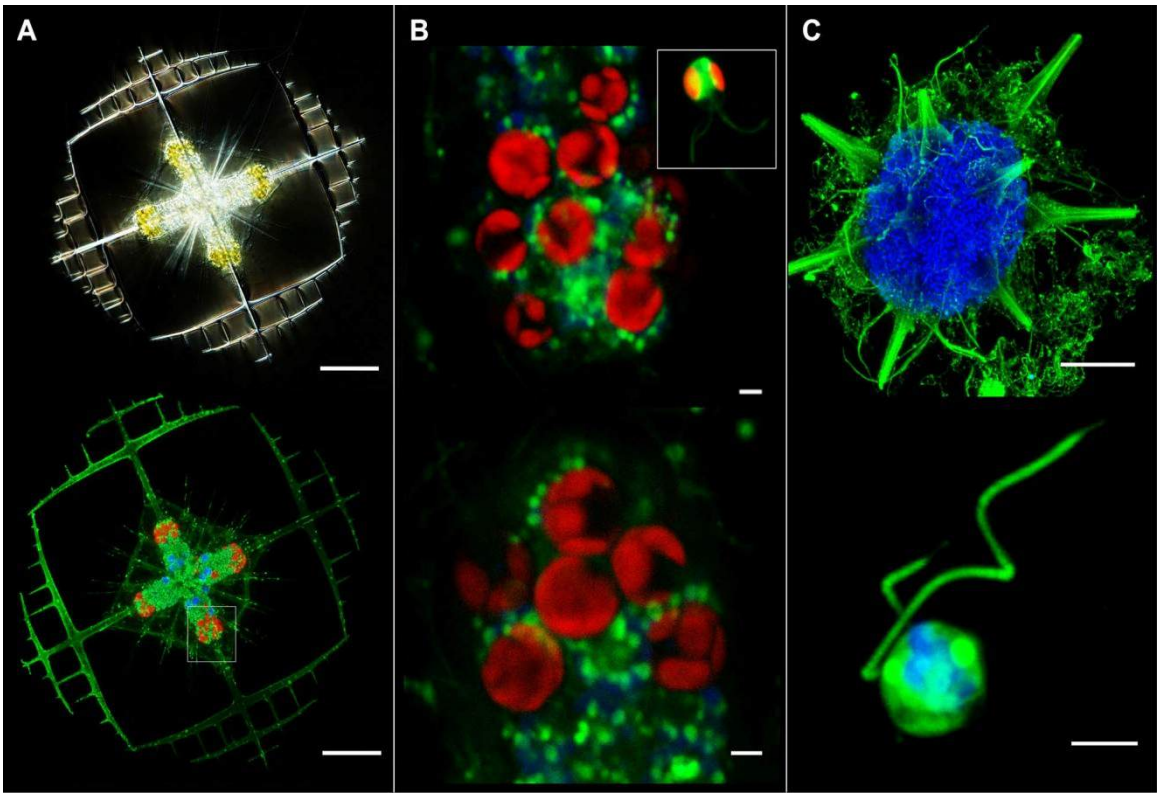


Fig. 2. Acquisition and maintenance of *Phaeocystis* in symbiosis.
(*A*) Adult Acantharia harboring golden endosymbiotic *Phaeocystis* cells in light microscopy (top) and red auto-fluorescence of chlorophyll-containing cells in fluorescence microscopy (bottom). Blue and green staining are the nuclei and membranes, respectively; scale bars: 50µm. (*B*) Detail of the phenotypic transformation of the symbiont *in hospite*: the size and the number of plastids of *Phaeocystis* cells increases from the free-living (inset) to the symbiotic form (top and bottom); scale bars: 2µm. (*C*) The thousands of blue nuclei in the adult Acantharia (top) represent individual "swarmers" (detail bottom) that are released in the environment at the end of the life cycle. The absence of symbionts at this stage indicates that *Phaeocystis* are acquired *de novo* at each host generation (horizontal transmission); scale bars: 50 µm (top) and 2µm (bottom). See SI Materials and Methods for details on fluorescence image acquisition.

Taxonomically unrelated organisms that interact closely in a symbiotic relationship can mutually influence each other's evolution and thus exhibit congruent phylogenies. In the case of the Acantharia-*Phaeocystis* symbiosis, cophylogenetic analyses based on 94 associations showed that symbiont phylogeny does not mirror host phylogeny (Fig S5). In addition, reconstruction of putative coevolutionary events (such as cospeciation or host-switch) onto host phylogeny did not reveal a strong cospeciation pattern (Figs. S6). The absence of a cophylogenetic signal can be attributed to the lack of species-level host specificity in the interaction and to the high level of dispersal of the extensive free-living population of *Phaeocystis* outside the host. Parallel speciation of symbiotic partners is rare in cases such as this where the symbiont is transmitted horizontally from the environment (Fig. 2*C*) through host generations (30). In summary, biogeography, rather than strict taxonomic specificity, is the main determinant in this flexible symbiotic relationship.

Evolutionary origin of the symbiosis

We used the exceptional fossil records of Radiolaria and Haptophyta to calibrate Bayesian relaxed molecular clock analyses on both host and symbiont phylogenies, allowing assessment of the evolutionary origin of the Acantharia-*Phaeocystis* symbiosis. The molecular dating analysis revealed that *Phaeocystis* originated between 288 and 178 Mya (Fig. 3*A*), and that the symbiosis, indicated by the emergence of the monophyletic group of symbiotic Acantharia (clades E and F in Fig. 3*B*), arose once between 175 and 93 Mya. The earliest possible origin of the Acantharia-*Phaeocystis* symbiosis corresponds to the first diversification that led to extant *Phaeocystis* species (~178 Mya). Coemergence has been reported in well-known plant-fungi associations (31) and coral-microalgae symbioses (4), allowing the colonization of terrestrial and reef ecosystems, respectively.

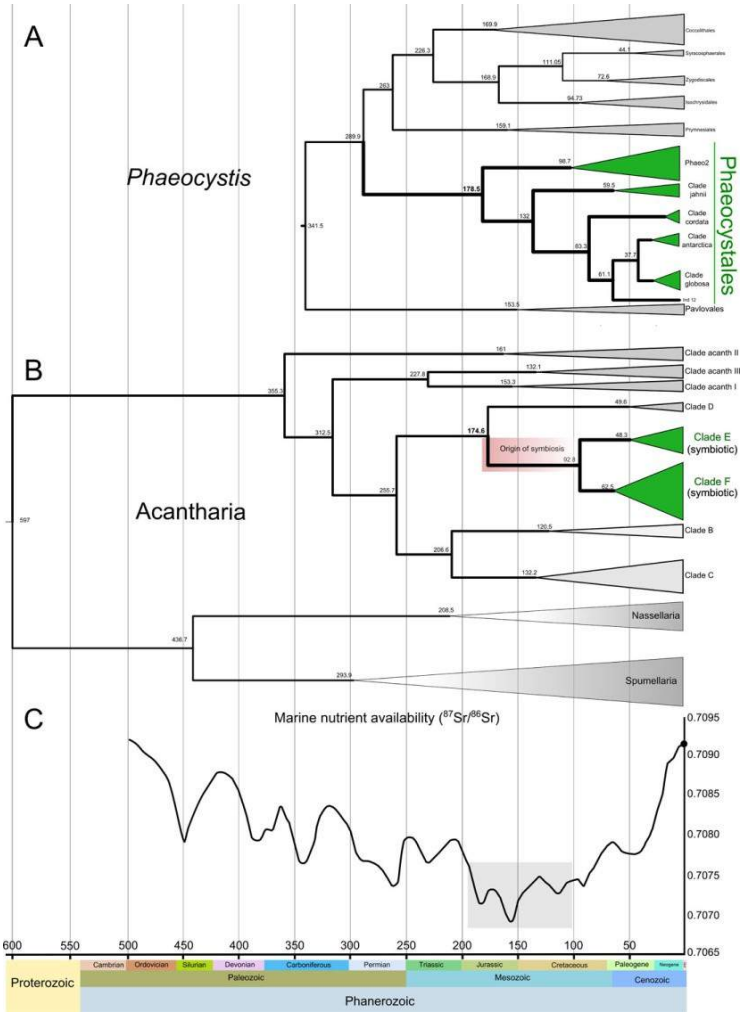


Fig. 3. Fossil-calibrated molecular clock dating the origin of the Acantharia-*Phaeocystis* photosymbiosis. Chronograms resulting from the Bayesian relaxed molecular clock analysis of haptophytes including *Phaeocystis* (*A*), and Acantharia including symbiotic species from clades E and F in green (*B*). Details of the fossil-based calibrations and the 95% credibility intervals (HPD) for the nodes are provided in SI. (*C*) 87Sr/86Sr ratio values throughout the Phanerozoic eon (550 Mya to present), reflecting nutrient availability in the ocean (modified from 32).

Between 190 and 100 Mya, when the Acantharia-*Phaeocystis* partnership arose, continental weathering (as indicated by the $^{87}Sr/^{86}Sr$ ratio) was lower than at any point during the last 550 My (Fig. 3*C*) and markedly lower than at present (32). Oligotrophy was undoubtedly a constant feature in ancient oceanic gyre water masses and the lower continental nutrient flux to the marine systems in the Jurassic may have resulted in generalized oligotrophy in surface waters. We hypothesize that the biological novelty of this symbiosis, resulting from the combination of the metabolic capacities of a heterotroph and an autotroph, was initially selected because it allowed the holobiont to cope with drastically low nutrient settings in transparent open oceans.

DMSP-DMSO measurements in the holobiont

Symbiotic Acantharia are abundant and widespread in the sunlit oligotrophic waters of modern oceans (18). The ecological success of this photosymbiosis implies the existence of physiological mechanisms to overcome constraints inherent to intracellular photosynthesis in high irradiance environments. *Phaeocystis* produces relatively high amounts of DMSP, which can represent up to 10% of total cell carbon (28). In order to assess the potential role of DMSP, its catabolite DMS, and its oxidation product dimethylsulfoxide (DMSO) in the symbiosis, we measured total DMSP + DMS (referred to here as DMSPt; see SI) and DMSO (DMSOt) in cultured free-living *P. cordata* cells and in field sampled Acantharia-*P. cordata* holobionts. In cultured *P. cordata*, DMSPt concentrations were 12 ± 1 fmol $cell^{-1}$ (358 ± 28 mM on a cell volume basis), within the range reported for other *Phaeocystis* species (28). The holobionts had a DMSPt per cell content three orders of magnitude higher (Table 1), yet they harbored only 29 ± 27 symbionts each. If all holobiont DMSPt were contained in the symbionts, it would represent an inconceivable proportion (ca. 300%) of *Phaeocystis* cell carbon. Thus, the major part of DMSP and DMS must be transferred from the endosymbiont to the host. Furthermore, we found that holobionts have a much larger proportion of DMSO than cultured *Phaeocystis*, as shown by 6-fold lower DMSPt/DMSOt ratios (Table 1). DMSP, DMS and DMSO have been suggested to serve as antioxidants in phytoplankton cells, with DMS and DMSO having, respectively, 60- and 20-fold greater oxidant scavenging efficiencies than DMSP (33). We therefore suggest that the transfer of these three dimethylated sulfur compounds from the symbiont to the host cell confer antioxidant protection to the holobiont in order to cope with oxidants produced by photosynthesis and high irradiance. This might also explain the high DMSP concentrations that have been reported in various hosts involved in other photosymbiotic associations, such as corals, anemones, clams and flatworms (34, 35). In each of these cases, DMSP is produced by symbiotic dinoflagellates that, like haptophytes, are important DMSP producers in their free-living state (28). The overall picture suggests that in UV-transparent oligotrophic waters, like those where the Acantharia-*Phaeocystis* symbiosis likely originated and is still abundant, antioxidant mechanisms are important factors for photosymbiotic associations.

Mutualism or enslavement?

Whereas the benefits of this association for the acantharian host are manifest, for the symbiont it is not clear where the cursor lies along the gradient of symbiotic interactions from mutualism to parasitism. In order to gain benefit from the association over an evolutionary time scale (i.e. $\geq$ 1 generation), *Phaeocystis* cells would need to be released to the environment in a viable condition following the symbiotic phase. If this is the case, the symbiosis likely plays a key role in the ecology of *Phaeocystis*, for example by protecting cells from grazing or viral attack (36) and/or by providing inocula for establishment of free-living populations. The present day ecological success of the microalga *Phaeocystis* may thus partly result from this ancient and persistent symbiosis with Acantharia.

Alternatively, the association may be an ecological (and hence evolutionary) cul-de-sac for the symbionts. To our knowledge, attempts to isolate the symbionts of Acantharia into culture (using mechanical or chemical disruption of the host cell, or natural disruption during gametogenesis) have never been successful, despite the relative ease with which free-living *Phaeocystis* can be cultured. Within the host, the phenotype of *Phaeocystis* symbionts can be extensively modified. The cell volume of symbionts can increase massively (up to tenfold compared to the free-living form), with formation of large vacuoles, several mitochondria and numerous large plastids (Fig. 2*B* and 24). This suggests that the host is able to exert a high degree of control over the symbiont to block cytokinesis and maximize photosynthetic capacity. Contrary to other photosymbioses where viable symbionts can be released (37-39), Acantharia may thus enslave and exploit *Phaeocystis* cells over a period of time before either digesting them or shedding non-viable cells.

Sequestration of prey organelles, such as plastids (kleptoplastidy) and nuclei (karyoklepty), is well documented (11, 40), including a report that a heterotrophic dinoflagellate from Antarctic waters is capable of temporarily retaining functional *Phaeocystis* plastids (41). Acquisition of phototrophy in Acantharia seems rather to occur by irreversible sequestration of whole cells ("cytoklepty"), a rare biological phenomenon that has been described for a particular strain of the ciliate *Mesodinium rubrum* and the sand-dwelling flagellate *Hatena arenicola* that maintain cells of a cryptophyte (42) and a prasinophyte (43), respectively. The observation that the host *Hatena* can transfer the phenotypically modified symbiont to a daughter cell during cell division led to the suggestion that the interaction is representative of an intermediate evolutionary step in plastid acquisition via secondary endosymbiosis. Vertical transmission of symbionts is presumably impossible for Acantharia, which produce reproductive cells (swarmers) that are significantly smaller than *Phaeocystis* cells (Fig. 2). The Acantharia-*Phaeocystis* association could nevertheless provide an alternative model to kleptoplastidy for investigating the genomic and physiological mechanisms involved in the early stages of permanent plastid acquisition.

Concluding remarks

Our study reveals a highly original mode of symbiosis between unicellular planktonic eukaryotes whereby the symbiont is an ecologically prominent component of the ecosystem not only in the symbiotic association, but also in its free-living phase. To our knowledge, this is the first report that describes a harmful and bloom-forming microalga occurring in a symbiotic relationship. In contrast, the bacterial or fungal symbionts within insects and plants, and the microalgal symbionts of corals and other benthic marine organisms are typically very rare in their free-living phase (44), in each case having first been discovered through the association with their host before being found as independent entities in the environment. Establishment of obligate photosymbiotic relationships that rely on horizontal transmission must be challenging in the vast and barren oceanic realm, making this a potentially risky strategy in evolutionary terms for the host. Acantharia, an important group of protistan zooplankton, have overcome the problem by developing a flexible association with one of the most abundant eukaryotic phytoplankton taxa in

the marine environment. Reports of abundant photosynthetic prokaryotes, including the cyanobacteria *Synechococcus* and *Prochlorococcus* forming symbiotic associations with protists in open oceans (45, 46) indicate that this mode of symbiosis may be well suited for the planktonic realm. This discovery provides further evidence of the remarkably high level of ecological and metabolic inter-dependency of planktonic organisms, and participates to increasing awareness of the significance of symbiosis in structuring a key compartment of the biosphere, the oceanic ecosystem (47).

Materials and Methods

Symbiotic Acantharia were collected at different locations worldwide and individually isolated using a micropipette (Table S1). DNA extraction from single cells was conducted as described previously (23). Partial 18S and 28S rDNA genes from host acantharians were amplified using Radiolaria-specific primers (23).

Table S1. Geographic origins of the Acantharia (host) sampled in this study that live with symbiotic *Phaeocystis*

Oceanic region	Sampling sites	Latitude	Longitude
Mediterranean sea	Naples	40°44'52.75"N	14°14'50.53"E
	Villefranche	43°40'55.20"N	7°18'44.76"E
Red sea	Eilat	29°30'18.15"N	34°57'25.01"E
East Pacific Ocean	Okinawa	26°13'6.12"N	127°16'26.69"E
Indian Ocean	Indian st39	18°41'60.00"N	66°17'60.00"E
	Indian st41	14°35'60.00"N	69°54'0.00"E
English Channel	Roscoff	48°45'5.49"N	3°57'18.58"W
Antarctic	Antarctica st84	60°12'0.85"S	60°30'2.10"W
	Antarctica st85	62° 0'13.40"S	49°16'53.36"W
	Antarctica st86	64°21'.569"S	53°00'.367"W
	Antarctica st88	63°24'.598"S	56°47'.621"W
South Atlantic	Seb st80	40°39'.245"S	52°12'.477"W

In parallel, 23 strains of *Phaeocystis* sp. from different culture collections were harvested in exponential growth phase and concentrated by centrifugation. DNA was extracted using the Nucleospin® RNA II kit. The 18S and 28S rDNA and the plastidial 16S, rbcL and psbA genes from symbiotic and cultured *Phaeocystis* were PCR-amplified. Molecular datasets were aligned with MAFFT v6.818 (48) and the optimal model of evolution was determined with jModelTest (49). Phylogenetic relationships were reconstructed with RAxML (50), and Bayesian inference was conducted using Beast v.1.6.1 and companion software (51). The phylogenetic trees of host (Acantharia) and symbiont (*Phaeocystis*) taxa (each symbiont is associated with one host) were used to perform co-phylogenetic analyses. We used an event-based method, Jane 3 (52) and a global fit method, ParaFit (53) implemented in CopyCat (54). Divergence times of Acantharia and *Phaeocystis* were estimated using Bayesian relaxed-clock methods implemented in BEAST v.1.6.1. BEAUti v. 1.6.1 was used to define parameters such as the fossil-based calibrations. In order to allow some uncertainty, most constraints were assigned a normal prior distribution with means corresponding to the fossil date or the first occurrence of a character, and a relatively broad standard deviation encompassing the minimum and maximum age of this calibration. Molecular clock parameters are detailed in SI Materials and Methods.

Aliquots of an exponentially growing *Phaeocystis cordata* culture and 50-200 acantharian cells were preserved for dimethylsulfoniopropionate (DMSP) and dimethylsulfoxide (DMSO) measurements. DMSP was measured as the dimethylsulfide (DMS) evolved by alkaline hydrolysis, using purge and trap gas chromatography coupled to flame photometric detection (GC-FPD). DMSO was analyzed as DMS using the cobalt-doped borohydride ($NaBH_4$) reduction method (55). Detailed methods are described in SI Materials and Methods.

Acknowledgments

We thank S. Romac, E. Boudoux, Y. Souab, M.J. Garet-Delmas for technical assistance and N. Suzuki for morphological identification of Acantharia. We also thank S. Hourdez, D. Stoecker, D. Vaulot and F. Vavre for critical comments of the manuscript. We are grateful to the institutes and funding agencies which supported the collection of samples: the LOV from Villefranche-sur-Mer; the InterUniversity Institute for Marine Sciences in Eilat (ASSEMBLE EU-FP7-227799); the Stazione Zoologica Anton Dohrn in Naples; the Akajima Marine Science Laboratory (CNRS-JST program); the Station Biologique de Roscoff and the TaraOceans expedition (doi:10.1371/journal.pbio.1001177). This research was supported by the Region Bretagne (DIPHOPE 044763), the French ANR 09-BLAN-0348 POSEIDON, and the French "Investissements d'Avenir" projects *OCEANOMICS* and EMBRC-France.

References

1. **Margulis** L, Fester R (1991) *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis.* MIT Press, Cambridge, MA.

2. **Muscatine** L, Falkowski PG, Porter JW, Dubinsky Z (1984) Fate of photosynthetic fixed carbon in light- and shade- adapted colonies of the symbiotic coral *Stylophora pistillata*. *Proc R Soc Lond Ser B* 222:181-202.

3. **Yellowlees** D, Rees TAV, Leggat W (2008) Metabolic interactions between algal symbionts and invertebrates hosts. *Plant Cell Environ* 31:679-694.

4. **Stanley** GD (2006) Photosymbiosis and the evolution of modern coral reefs. *Science* 312:857–858.

5. **Carlos** A, Baillie B, Kawachi M, Maruyama T (1999) Phylogenetic position of *Symbiodinium* (Dinophyceae) isolates from tridacnids (Bivalvia), cardiids (Bivalvia), a sponge (Porifera), a soft coral (Anthozoa), and a free-living strain. *J Phycol* 35:1054 –1062.

6. **LaJeunesse** TC (2001) Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus *Symbiodinium* using the ITS region: In search of a ‘‘species’’ level marker. *J Phycol* 37:866-880.

7. **Pochon** X, Pawlowski J, Zaninetti L, Rowan R (2001) High genetic diversity and relative specificity among *Symbiodinium*-like endosymbiotic dinoflagellates in soritid foraminiferans. *Mar Biol* 139:1069-1078.

8. **Haeckel** E (1887) Report on Radiolaria Collected by H.M.S. Challenger during the Years 1873-1876. In Murray CWT aJ (ed), The Voyage of the H.M.S.Challenger, vol. 18. Her Majesty’s Stationary Office, London, pp 1803.

9. **Coffroth** MA, Lewis CF, Santos SR, Weaver JL (2006) Environmental populations of symbiotic dinoflagellates in the genus Symbiodinium can initiate symbioses with reef cnidarians. *Curr Biol* 5:16(23).

10. **Littman** RA, van Oppen MJH, Willis BL (2008) Methods for sampling free-living *Symbiodinium* (zooxanthellae) and their distribution and abundance at Lizard Island (Great Barrier Reef). *J Exp Mar Biol Ecol* 364:48-53.

11. **Stoecker** DK, Johnson MD, de Vargas C, Not F (2009) Acquired phototrophy in aquatic protists. *Aquat Microb Ecol* 57:279-310.

12. **Shaked** Y, de Vargas C (2006) Pelagic photosymbiosis : rDNA assessment of diversity and evolution of dinoflagellate symbionts and planktonic formainiferal hosts. *Mar Ecol Prog Ser* 325:59-71.

13. **Gast** RJ, Caron DA (1996) Molecular phylogeny of symbiotic dinoflagellates from planktonic foraminifera and radiolaria. *Mol Biol Evol* 13:1192–1197.

14. **Gast** RJ, Caron DA (2001) Photosymbiotic associations in planktonic foraminifera and radiolaria. *Hydrobiologia* 461:1–7.

15. **Anderson** OR (1983a) The radiolarian symbiosis. In: Goff L (ed) *Algal symbiosis: a continuum of interaction strategies*. Cambridge University Press, Cambridge, p 69–89.

16. **Falkowski** PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR (2004) The evolution of modern eukaryotic phytoplankton. *Science* 305:354–360.

17. **Zas'ko** DN, Vedernikov VI (2003) A comparative analysis of the vertical distribution of epipelagic radiolaria, chlorophyll, and zooplankton in different areas of the North Atlantic Ocean (from June to September 2001). *Oceanol* 43:63–71.

18. **Michaels** AF, Caron DA, Swanberg NR, Howse FA, Michaels CM (1995) Planktonic sarcodines (acantharia, radiolaria, foraminifera) in surface waters near Bermuda: abundance, biomass and vertical flux. *J Plankton Res* 17:131-163.

19. **Swanberg** NR, Caron DA (1991) Patterns of feeding in epipelagic oceanic plankton. *J Plankton Res* 13:287-312.

20. **Michaels** AF (1988) Vertical distribution and abundance of Acantharia and their symbionts. *Mar Biol* 97:559-569.

21. **Schewiakoff** WT (1926) The Acantharia. Fauna e Flora del Golfo di Napoli 37:1-755.

22. **Taylor** FJR (1982) Symbioses in marine microplankton. *Ann Inst Oceanogr* Paris 58(S):61-90.

23. **Decelle** J, Suzuki N, Mahé F, de Vargas C, Not F (2012) Molecular phylogeny and morphological evolution of the Acantharia (Radiolaria) *Protist* 163:435-450.

24. **Febvre** J, Febvre-Chevalier C (1979) Ultrastructural study of zooxanthellae of 3 species of Acantharia (Protozoa: Actinopoda), with details of their taxonomic position in the Prymnesiales (Prymnesiophyceae Hibberd, 1976). *J Mar Biol Assoc UK* 59:215-226.

25. **Schoemann** V, Becquevort S, Stefels J, Rousseau V, Lancelot C (2005) *Phaeocystis* blooms in the global ocean and their controlling mechanisms: a review. *Journal of Sea Research* 53:43-66.

26. **Verity** PG, Smetacek V (1996) Organism life cycles, predation, and the structure of marine pelagic ecosystems. *Mar Ecol Prog Ser* 30:277-293.

27. **Arrigo** KR et al. (1999) Phytoplankton community structure and the drawdown of nutrients and CO_2 in the southern Ocean. *Science* 283:365-367.

28. **Stefels** J, Steinke M, Turner S, Malin G, Belviso S (2007) Environmental constraints on the production and removal of the climatically active gas dimethulsulphide (DMS) and implications for ecosystem modelling. *Biogeochemistry* 83:245-275.

29. **Medlin** L, Zingone A (2007) A taxonomic review of the genus *Phaeocystis*. *Biogeochemistry* 83:3-18.

30. **Sachs** JL, Essenberg CJ, Turcotte MMM (2011) New paradigms for the evolution of beneficial infections. *Trends Ecol Evol* 26 (4):202-209.

31. **Heckman** DS et al. (2001) Molecular evidence for the early colonization of land by fungi and plants. *Science* 293:1129-1133.

32. **Cardenas** AL, Harries PJ (2010) Effect of nutrient availability on marine origination rates throughout the Phanerozoic eon. *Nature geosci* 3:430-434.

33. **Sunda** W, Kieber DJ, Kiene RP, Huntsman S (2002) An antioxidant function for DMSP and DMS in marine algae. *Nature* 418:317-320.

34. **Yost** DM, Mitchelmore CL (2009) Dimethylsulfoniopropionate (DMSP) lyase activity in different strains of the symbiotic alga *Symbiodinium microadriaticum*. *Mar Ecol Prog Ser* 386:61-70.

35. **Van Bergeijk** SA, Stal LJ (2001) Dimethylsulfoniopropionate and dimethylsulfide in the marine flatworm *Convoluta roscoffensis* and its algal symbiont. *Mar Biol* 138:209-216.

36. **Kodama** Y, Fujishima M (2009) Infection of *Paramecium bursaria* by symbiotic *Chlorella* species, in *Endosymbionts in Paramecium*, Microbiology Monographs 12, Fujishima M (ed.), (Springer-Verlag Berlin Heidelberg) pp 31–55.

37. **Fishman** Y, Zlotkin E, Sher D (2008) Expulsion of Symbiotic Algae during Feeding by the Green Hydra – a Mechanism for Regulating Symbiont Density?. PLoS ONE 3(7):e2603. doi:10.1371/journal.pone.0002603.

38. **Gates** RD, Baghdasarian G, Muscatine L (1992) Temperature stress causes host cell detachment in symbiotic cnidarians: implications for coral bleaching. *Biol Bull* 182:324-332.

39. **Castro-Sanguino** C, Sanchez JA (2011) Dispersal of *Symbiodinium* by the stoplight parrotfish *Sparisoma viride*. *Biol Lett* 8(2):282-286.

40. **Johnson** MD, Oldach D, Delwiche CF, Stoecker DK (2007) Retention of transcriptionally active cryptophytes nuclei by the ciliate *Myrionecta rubra*. *Nature* 445:426-428.

41. **Gast** RJ, Moran DM, Denett MR, Caron DA (2007) Kleptoplastidy in an Antarctic dinoflagellate: caught in evolutionary transition? *Environ Microbiol* 9:39-45.

42. **Hansen** PJ, Fenchel T (2006) The bloom-forming ciliate *Mesodinium rubrum* harbours a single permanent endosymbiont. *Marine Biology Research* 2:169-177.

43. **Okamoto** N, Inouye I (2005) A secondary symbiosis in progress? *Science* 310:287.

44. **Nyholm** SV, McFall-Ngai MJ (2004) The winnowing: establishing the squid-vibrio symbiosis. *Nat Rev Microbiol* 2(8):632-642.

45. **Foster** RA, Carpenter EJ, Bergman B (2006) Unicellular cyanobionts in open ocean dinoflagellates, radiolarians, and tintinnids: ultrastructural characterization and immuno-localization of phycoerythrin and nitrogenase. *J Phycol* 42:453-463.

46. **Foster** RA, Collier JL, Carpenter EJ (2006) Reverse transcription PCR amplification of cyanobacterial symbiont 16S rRNA sequences from single non-photosynthetic eukaryotic marine planktonic host cells. *J Phycol* 42:243-250.

47. **Strom** SL (2008) Microbial ecology of ocean biogeochemistry: a community perspective. *Science* 230:1043-1045.

48. **Katoh** K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* 9(4):286-298.

49. **Posada** D (2008) jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol* 25:1253-1256.

50. **Stamatakis** A, Hoover P, Rougemont J (2008) A rapid boostrap algorithm for the RAxML Web-Servers. *Syst Biol* 75(5):758-771.

51. **Drummond** AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.

52. **Conow** C, Fielder D, Ovadia Y, Libeskind-Hadas R (2010) Jane: a new tool for the cophylogeny reconstruction problem. *Algorithms Mol Biol* 5:16.

53. **Legendre** P, Desdevises Y, Bazin E (2002) A statistical test for host-parasite coevolution. *Syst Biol* 51(2):217-234.

54. **Meier-Kolthoff** J, Auch A, Huson D, Goker M (2007) COPYCAT: cophylogenetic analysis tool. *Bioinformatics* 23 (7):898-900.

55. **Simó** R, Vila-Costa M (2006) Ubiquity of algal dimethylsulfoxide in the surface ocean: geographic and temporal distribution patterns. *Mar Chem* 100:136-146.

Supplementary information

SI Materials and Methods

Collection and DNA extraction of symbiotic Acantharia and cultured *Phaeocystis* strains. Symbiotic Acantharia were collected in 2010 and 2011 at different locations worldwide (Table S1). Acantharian cells were harvested at the subsurface by slowly towing nets (64 and 200 µm mesh size), and either directly isolated or preserved within the bulk sample in 70% ethanol. Micropipette isolation of cells using an inverted microscope was conducted, and each single cell sorted was individually cleaned three times in successive 0.2 µm filtered seawater bath. Freshly isolated and cleaned cells were then maintained for several hours in 0.2 µm filtered seawater to allow self-cleaning. Both fresh and ethanol-preserved cells were photographed under an inverted microscope, and finally isolated into a guanidine-containing extraction buffer (GITC). Eppendorf tubes containing the samples were stored at -20 °C until processing. The single-cell DNA extraction of the Acantharia-symbiont holobiont was carried out as described previously (1).

In parallel, 23 strains of *Phaeocystis* sp., isolated in different oceans worldwide, were obtained from various culture collections. Codes for strain sources are: RCC: Roscoff Culture Collection, France; NIES: National Institute for Environmental Studies, Japan; PCC: Plymouth Culture Collection, U.K.; K: Scandinavian Culture Collection of Algae and Protozoa, Copenhagen, Danemark. Cultured cells were harvested in exponential growth phase and concentrated by centrifugation. Total nucleic acids were extracted using the the Nucleospin® RNA II kit (Macherey-Nagel, Hoerdt, France) and quantified using a Nanodrop ND-1000 Spectrophotometer (Labtech International, France).

Gene amplification and phylogenetic methods for the host Acantharia. Polymerase Chain Reactions (PCR) on Acantharia were performed with Radiolaria-specific primers to amplify the partial 18S and 28S (D1/D2 domains) rDNA genes (see 1 for details). Additional sequences of symbiotic Acantharia were acquired from GenBank and added to the sequences obtained in this study. This dataset was then aligned with MAFFT v6.818 (2), concatenated in one partition (94 taxa with 2,4 kb aligned characters), and phylogenetic relationships were reconstructed with RAxML (3) using the GTR + G model. The phylogenetic tree of the Acantharia was used for the cophylogenetic analyses (see below). Sequences obtained from the host Acantharia were deposited in GenBank under accession numbers JQ697697 to JQ6977738.

Gene amplification and phylogenetic methods for *Phaeocystis* in symbiosis and in culture. PCRs on symbiotic and cultured *Phaeocystis* were conducted to amplify the 18S and 28S rDNA and the plastidial rbcL and psbA genes (see 4-11 for detailed information on primers). Amplifications were performed with the Phusion high-fidelity DNA polymerase (Finnzymes) in a 25-µl reaction volume using the following PCR parameters: 30 s at 98 °C, followed by 35 cycles of 10 s denaturation at 98 °C, 30 s annealing at 50 °C for the 18S, 28S and psbA and 55°C for rbcL, and 30 s extension at 72 °C, with a final elongation step of 10 minutes at 72 °C. PCR products were then purified by EXOSAP-IT (GE Healthcare Bio-Sciences Corp.), and bidirectionally sequenced using the ABI-PRISM Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems). Raw sequences were edited and assembled with Chromas Pro v.1.5 (Gene Codes Corporation, Ann Arbor, Michigan).

Single locus analyses were performed for each genetic marker. Sequences of acantharian symbionts and *Phaeocystis* strains obtained in this study were added to GenBank sequences from other *Phaeocystis* strains and related environmental clone libraries. The four datasets were thus aligned with MAFFT v6.818. The default alignment algorithm was used for sequences coding for plastidial genes (rbcL and psbA), whereas for rDNA sequences the Q-INS-*i* alignment option taking into account the secondary structure was selected. Ambiguous regions of the 18S and the 28S rDNA alignment, characterized by a high rate of insertion/deletion, were removed before analyses using the software Gblocks v.0.91b (12). A list of taxa included in these single locus datasets is provided upon request. For phylogenetic analyses, each dataset contained five outgroup sequences from the Prymnesiophyceae (*Emiliania huxleyi*, *Isochrysis galbana*, *Pleurochrysis carterae*, *Prymnesium parvum* and *Imantonia rotunda*). For each dataset, the optimal model of evolution was determined under the AIC, AICc and BIC criteria implemented in jModelTest (13). Maximum Likelihood (ML) analyses and statistical support were assessed by performing 1000 bootstrap replicates as implemented on the RAxML version 7.2.0 blackbox servers. Resulting single-gene phylogenies are shown in Figs. S1, S2 and S3.

After visual checking of the congruence of each single-locus phylogeny, a concatenated dataset was constructed. This dataset constituted 3170 bp obtained by concatenation of the sequences of the ribosomal loci 18S (686 pb) and 28S (827 pb), and the plastidial genes psbA (564 pb) and rbcL (1093 pb) for 131 taxa: 98 symbionts, 28 *Phaeocystis* culture strains, and 5 Haptophyta taxa as the outgroup. Accession numbers of the sequences and details of the dataset are available upon request. All analyses on the concatenated 4-locus alignment were carried out with a model in which the data set was partitioned by marker, thus defining four partitions. Maximum likelihood phylogenies were inferred using RAxML with GTR+I+Γ model parameters estimated independently for each partition by jModelTest. Analyses were performed for each dataset at least four times, with different starting trees. Statistical support was assessed by performing 2000 boostrap replicates.

The Bayesian inference was conducted on the concatenated datasets (in one partition) using Beast v.1.6.1 and companion software (14) under the GTR+I+Γ(4) model. Two Markov Chain Monte Carlo (MCMC) chains were run for 50 million generations, sampling every 1000 generations. The two runs were combined with LogCombiner v1.6.1, and convergence of log-likelihoods and parameter values were assessed in Tracer v1.4.1. 10% of the total trees were discarded as burn-in, and the remaining trees were used by TreeAnnotator 1.5.4 to build the consensus tree and to calculate the posterior probabilities (PP) of each node. The final tree was visualized with FigTree v1.3.1. The multi-gene phylogeny is presented Fig. S4.

Mantel test. Mantel tests were performed to assess whether the distribution of the host (Acantharia) and the symbiont (*Phaeocystis*) are driven by a biogeographical pattern. A geographical distance matrix based on coordinates and a pairwise genetic distance matrix were built for each genetic marker of each partner, using Geographic Distance Matrix Generator v.1.2.3 (15) and MEGA5 (16), respectively. A Pearson's correlation between the matrices was then tested using the R package ape, based on 999 permutations. Correlations were significant (p-value = 0.001) only for the genetic markers of the symbiont *Phaeocystis*: 0.406, 0.425, 0.858 and 0.652 for the 18S rDNA, 28S rDNA, psbA and rbcL, respectively. No significant correlations were found for both the 18S rDNA and 28S rDNA of the host Acantharia ($p > 0.29$).

Cophylogenetic analyses. Several methods have been published to study cophylogenetic patterns between hosts and their symbionts (17-18), which can be classified into event-based methods and global fit methods (19). Event-based methods generally aim at reconciling tree topologies of hosts and symbionts by adequately mixing four or five kinds of coevolutionary events (cospeciation or codivergence, host-switch, duplication, sorting, and failure to diverge – not used in all methods) and find the best reconstructions by minimizing its global cost (each event type is attributed a cost) to produce optimal cophylogenetic scenarios. The significance of the global cost is assessed against a random distribution of costs generated using random trees: if the observed optimal cost is significantly lower than optimal costs computed from randomly generated trees, then a global cospeciation signal is present. Global fit methods do not rely on events but assess the congruence between the two trees taking the pattern of host specificity into account; again the observed level of congruence is tested against a random distribution. No scenario is produced but the computational burden is much lighter than for event-based methods. The phylogenetic trees of 94 host taxa (Acantharia) and 94 symbiont taxa (each symbiont is associated with one host), inferred as explained above, have been used in these cophylogenetic analyses. We used an event-based method, Jane 3 (20) and a global fit method, ParaFit (21) implemented in CopyCat (22). Jane was used with the following event-cost scheme (Cospeciation= 0, Duplication = 1, Host switch = 2, Sorting = 1, Failure = 1), a number of generations of 500 and a population size of 50. Statistical tests for tree congruence in ParaFit and Jane were carried out with 999 permutations, and symbiont trees instead of tip mappings were randomized in Jane. Results of these different cophylogenetic analyses are shown on Figs S5 and S6.

Molecular clock analyses. Divergence times of Acantharia and *Phaeocystis* was estimated using Bayesian relaxed-clock methods implemented in BEAST v.1.6.1. For host Acantharia, 18S rDNA sequences of several representatives from each acantharian clade were retrieved from GenBank: clades B, C, D, E, F and Acanth I, II and III. Since Acantharia do not have a fossil record, we used an indirect calibration with 5 fossil-based calibration points from other fossilizable Radiolaria: Spumellaria and Nassellaria (a list of used taxa with their GenBank accession numbers is available upon request). The 18S rDNA sequences of these radiolarian groups, publicly available, were therefore added to the acantharian sequences to build a dataset of 98 taxa with 525 unambiguously aligned positions using Gblocks v.0.91b. The monophyly of the Acantharia (ingroup) was set in BEAST. Different models were analyzed by alternatively adding or removing calibrations. The cross-comparison between these different models allowed us to test the reliability of each calibration and see how the age of each node was affected by various calibrations. The best model was chosen according to the molecular clock statistics as explained below.

Here are the details of the five calibration points used in this study:

Node 1: The calibration for the root of the tree corresponds to the first appearance of radiolarian fossils in the Phanerozoic, which is recognized to be in the early Cambrian, 542 Ma (23, 24). We therefore set a minimum (542 Ma) and maximum (800 Ma) bound with a uniform probability, U (542-800), to allow uncertainty about the origin of Radiolaria.

Node 2: This calibration is based on the first occurrence of fossils of Nassellaria in the Devonian, 316 – 459 Mya (25, 26). The prior was normally distributed with a mean of 370 and a standard deviation of 50: N (370; 50). The Nassellaria-Spumellaria clade was forced to form a monophyletic group as shown in recent multigene phylogenies (27, 28).

Node 3: The genus *Hexacontium* and the family Spongodiscidae (represented by the taxa *Dictyocoryne* spp, *Stylodictya* sp., *Spongodiscus* spp., *Spongaster* sp. and *Euchitonia* sp) appear in the fossil record in the Triassic (26). However, according to (25), *Hexacontium* would occur rather in the Jurassic (150 – 200 Ma) and another study (29) identifies the first members of the Spongodiscidae in the Campanian deposits around 80 Ma. In different molecular phylogenies (27), *Hexacontium* and Spongodiscidae group together within a monophyletic clade. Thus, we decided to constrain the node of this clade with a calibration at 200 Ma and a broad standard deviation of 50: N (200; 50), covering the Triassic and Jurassic due to the uncertainties mentioned above.

Node 4: Fossils of *Actinomma* sp. would first appear in the Triassic but with some taxonomic uncertainty (25), and were observed in the Cretaceous in many places worldwide (26). The family Actinommidae would have a reliable fossil in the Jurassic (175 Ma). The node corresponding to the divergence of *Actinomma* sp. was therefore set at 170 Ma: N (170; 30).

Node 5: This node corresponds to the first fossils of the genus *Cladococcus*, which would be in the Cenozoic era (Paleocene, 60 Ma; (25, 26)). We used N (60; 20) as prior for this calibration.

The Markov chains were run two times for 30 million generations and sampled every 1000 generations. The two runs were combined with LogCombiner v1.6.1, and the first 6000 trees were discarded as burn-in. MCMC chain convergence and stationarity were assessed using Tracer v1.5 by examining effective sample sizes (ESS values > 500) for all parameters, such as posterior, prior, likelihood etc. Node ages and the 95% highest posterior density interval (HPD) for divergence times were calculated using TreeAnnotator v1.6.1 and visualized with FigTree v1.3.1 (Fig. S7).

For *Phaeocystis*, the divergence time analysis was carried out with 4 fossil-based calibrations from coccolithophores used in a previous study (30). The molecular dataset comprises the nuclear 18S and 28S rDNA and plastidial rbcL genes of *Phaeocystis* (from the clades defined in this study) and from other haptophytes: 78 taxa in total with 3773 aligned characters. The three genes were considered into one single partition (30), the Pavlovales were set as the outgroup and *Phaeocystis* as forming a monophyletic group.

We used four calibration points:

Node 1: First occurrence in the fossil record of heterococcoliths: N (220, 4).
Node 2: Divergence of *Umbilicosphaera* sp. and *Coccolithus* sp.: N (65, 2).
Node 3: First occurrence of *Helicosphaera* sp.: N (25, 1).
Node 4: Divergence of *Umbilicosphaera* sp. and *Calcidiscus* sp.: N (24, 2).

The Markov chains were run two times for 50 million generations and sampled every 1000 generations. The two runs were combined with LogCombiner v1.6.1, and the first 10000 trees were discarded as burn-in. The corresponding molecular clock is shown Fig. S8.

DMSP and DMSO analyses and calculations of cell content. Aliquots of an exponentially growing *Phaeocystis cordata* culture were preserved in 12 mL gas-tight glass vials, with no headspace. The vials were immediately capped and crimped after adding two NaOH pellets (0.2 M final concentration, pH > 12). Samples with 0.22 µm filtered seawater containing 50-200 acantharian cells, and blanks containing only filtered seawater were processed and stored in the same way. Dimethylsulfoniopropionate (DMSP) was measured, within 2 months, as the dimethylsulfide (DMS) evolved by alkaline hydrolysis, using purge and trap gas chromatography coupled to flame photometric detection (GC-FPD). Aliquots of 10 to 20 µL, withdrawn with a gas-tight syringe, were analyzed for the *P. cordata* culture, while 20 to 100 µL were analyzed in the case of *Acantharia* samples. The aliquots were injected into a purge vial containing ca. 1-3 mL of milliQ water, and were purged with He (40 mL/min). The stripped DMS was cryogenically trapped, re-volatilzed, injected, analyzed and quantified as described elsewhere (31). Analytical replicates (n = 2 to 4) of each sample had a relative analytical error < 7 % (average 4 %). In turn, the mean DMS+DMSP concentrations of triplicate samples of the *P. cordata* culture showed good agreement (experimental error < 8 %). Since *Phaeocystis* strains are able to cleave DMSP, the culture medium might contain DMS, which was not measured. Hence, the values reported here for the cultures are actually DMS + DMSP concentrations. DMS was assumed to represent around 10 % of the DMS + DMSP pool, as previously reported for *Phaeocystis* cultures with cell densities in the same order of those analyzed here (32, 33). In the natural *Acantharia* samples, conversely, reported values correspond to particulate (cellular) DMSP, since total DMS + DMSP concentrations were corrected by subtraction of dissolved DMS + DMSP concentrations in filtered seawater.

Dimethylsulfoxide (DMSO) was analyzed as DMS using the cobalt-doped borohydride ($NaBH_4$) reduction method (34). After having purged the DMS evolved from alkaline DMSP cleavage from the sample aliquots, a small piece (20-50 mg) of a $NaBH_4$ was added to the purge vial. For 2 minutes, the reducing agent was allowed to react in the capped vial connected to the cryogenic trap, so that the evolved H_2 stripped the formed DMS into the trap. Then, the purge was initiated and held over 6 more minutes. Finally, trapped DMS was injected and analyzed by GC-FPD as described above. In the *P. cordata* culture, reported values correspond to total (particulate + dissolved) DMSO, with dissolved DMSO concentrations potentially being of the same order as those of particulate DMSO (35). In the natural *Acantharia* samples, conversely, reported values correspond to particulate (cellular) DMSO, since total DMSO concentrations were corrected by subtraction of dissolved DMSO concentrations in filtered seawater.

Measured DMSPt and DMSOt concentrations in the *Acantharia* and the *P. cordata* suspensions (pmol/ml) were converted into pmol/cell content by dividing by cell numbers in each sample. Putative intracellular DMSPt concentrations in cultured and symbiont *P. cordata* were estimated by considering that the spherical-equivalent diameter size of the cultured strain was 4.0 µm and that of the symbiont cells (measured by confocal microscopy) was 6.79 ±2.14 µm. Carbon content of *P. cordata* cells was estimated by the equation of $C\ (pg/cell) = 0.216\ V^{0.939}$ (36). Contribution of DMSP to total cell carbon was calculated considering that 1 pmol DMSP = 5 pmol C.

Fluorescence microscopy. Microscopy was conducted using a Leica SP5II confocal laser scanning microscope and HCX PL APO lambda blue 63x1.40 OIL UV objective or HCX PL APO CS 20 x0.70 IMM UV objective, according to specimen size. All pictures but swarmers are from live specimens freshly collected from environment prior confocal scanning laser microscopy imaging (Leica Microsysteme SP5II). Swarmers were fixed with 1% glutaraldehyde prior imaging. Nuclei and membrane structures were stained 30 min with 10µM Hoechst 33342 (Invitrogen) and 17 µM DiOC6 (Invitrogen), respectively. Two sequence steps were designed to collect first DiOC6 signal (ex488nm/em500-520nm) simultaneously with chlorophyll signal (ex633 nm/em670-700nm), and then Hoechst signal (ex405/em420-470). Red, Green, and Blue channels are respectively dedicated to chlorophyll, DiOC6/membranes, and Hoechst33342/nuclei fluorescence. Image handling was performed using Fiji (ImageJ: http://pacific.mpi-cbg.de/). All pictures from figure 2 are maximum projections from suitable z-stacks, but the Acantharia close-up illustrates a single optical slice.

1. Decelle J, Suzuki N, Mahé F, de Vargas C, Not F (2012) Molecular phylogeny and morphological evolution of the Acantharia (Radiolaria) *Protist* 163:435-450.

2. Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* 9(4):286-298.

3. Stamatakis A, Hoover P, Rougemont J (2008) A rapid boostrap algorithm for the RAxML Web-Servers. *Syst Biol* 75(5):758-771.

4. Medlin LK, et al. (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* 71:491-499.

5. Ando H, et al. (2009) Intraspecific variations in the ITS region of recent radiolarians. *Earth Evol Sci* 3:37-44.

6. Fuller N, et al. (2006) Analysis of photosynthetic picoeukaryote diversity at open ocean sites in the Arabian Sea using a PCR biased towards marine algal plastids. *Aquat Microb Ecol* 43:79-93.

7. Lepère C, et al. (2011) Whole-genome amplification (WGA) of marine photosynthetic eukaryote populations. *FEMS Microbiol Ecol* 76 :513-523.

8. Coolen MJL, et al. (2004) Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth Planet Sci Lett* 223:225-239.

9. Simon N et al. (2000) Oligonucleotide Probes for the Identification of Three Algal Groups by Dot Blot and Fluorescent Whole-Cell Hybridization. *J Eukaryot Microbiol* 47:76-84.

10. Liu H, et al. (2009) Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *PNAS* 106: 12803-12808.

11. Fujiwara S, et al. (1994) Molecular phylogenetic analysis of *rbc*L in the Prymnesiophyta. *J Phycol* 30:863–71.

12. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17**:**540-552.

13. Posada D (2008) jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol* 25:1253-1256.

14. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.

15. Ersts PJ Geographic Distance Matrix Generator (ver 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg.

16. Tamura K et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 28:2731-2739.

17. Page RDM *Tangled Trees: Phylogeny, cospeciation, and coevolution*. The University of Chicago Press. ed. 2003.

18. Light J, Hafner M (2008) Codivergence in Heteromyid Rodents (Rodentia: Heteromyidae) and Their Sucking Lice of the Genus Fahrenholzia (Phthiraptera: Anoplura). *Syst Biol* 57(3):449-465.

19. Desdevises Y (2007) Cophylogeny: insights from fish-parasite systems. *Parassitologia* 49(3):125.

20. Conow C, Fielder D, Ovadia Y, Libeskind-Hadas R (2010) Jane: a new tool for the cophylogeny reconstruction problem. *Algorithms Mol Biol* 5:16.

21. Legendre P, Desdevises Y, Bazin E (2002) A statistical test for host-parasite coevolution. *Syst Biol* 51(2):217-234.

22. Meier-Kolthoff J, Auch A, Huson D, Goker M (2007) COPYCAT: cophylogenetic analysis tool. *Bioinformatics* 23 (7):898-900.

23. Anderson OR (1983) Ed., *Radiolaria* (Springer-Verlag, NewYork, 355 pp).

24. Braun A, Chen J, Waloszek D, Maas A (2007) in *First early Cambrian Radiolaria*. (Geol Soc, London, Spec Publ) 286:143-149.

25. De Wever P, Dumitrica P, Caulet JP, Nigrini C, Caridroit M (2001) Ed., *Radiolarians in the Sedimentary Record* (Gordon and Breach Sci. Publ., Amsterdam).

26. PaleoDB Paleobiology Database (http://paleodb.org/)

27. Krabberød AK, et al. (2011) Radiolaria divided into Polycystina and Spasmaria in combined 18S and 28S rDNA phylogeny. *PLoS ONE* 6: e23526 doi:10.1371/journal.pone.0023526.

28. Ishitani Y, et al. (2011) Multigene phylogenetic analyses including diverse radiolarian species support the "Retaria" hypothesis – The sister relationship of Radiolaria and Foraminifera. *Mar Micropaleontol* 81:32-42.

29. O'Dogherty L, et al. (2009) Catalogue of Mesozoic radiolarian genera. *Geodiversitas* 31(2):213-270.

30. Liu H, Aris-Brosou S, Probert I, de Vargas C (2010) A time line of the environmental genetics of the haptophytes. *Mol Biol Evol* 71:161-176 (2010).

31. Saló V, Simó R, Calbet A (2010) Revisiting the dilution technique to quantify the role of microzooplankton in DMS(P) cycling: Laboratory and field tests. *J Plankton Res* 32:1255-1267.

32. Stefels J, Van Boeckel WHM (1993) Production of DMS from dissolved DMSP in axenic cultures of the marine phytoplankton species *Phaeocystis* sp.. *Mar Ecol Prog Ser* 97:11-18.

33. Matrai PA, et al. (1995) Light-dependent production of DMS and carbon incorporation by polar strains of *Phaeocystis* spp. *Mar Biol* 124:157-167.

34. Simó R, Vila-Costa M (2006) Ubiquity of algal dimethylsulfoxide in the surface ocean: geographic and temporal distribution patterns. *Mar Chem* 100:136-146.

35. Hatton AD, Wilson ST (2007) Particulate dimethylsulphoxide and dimethylsulphoniopropionate in phytoplankton cultures and Scottish coastal waters. *Aquat Sci* 69:330- 340.

36. Menden-Deuer S, Lessard EJ (2000) Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol Oceanogr* 45:569-579.

Supplementary Figures

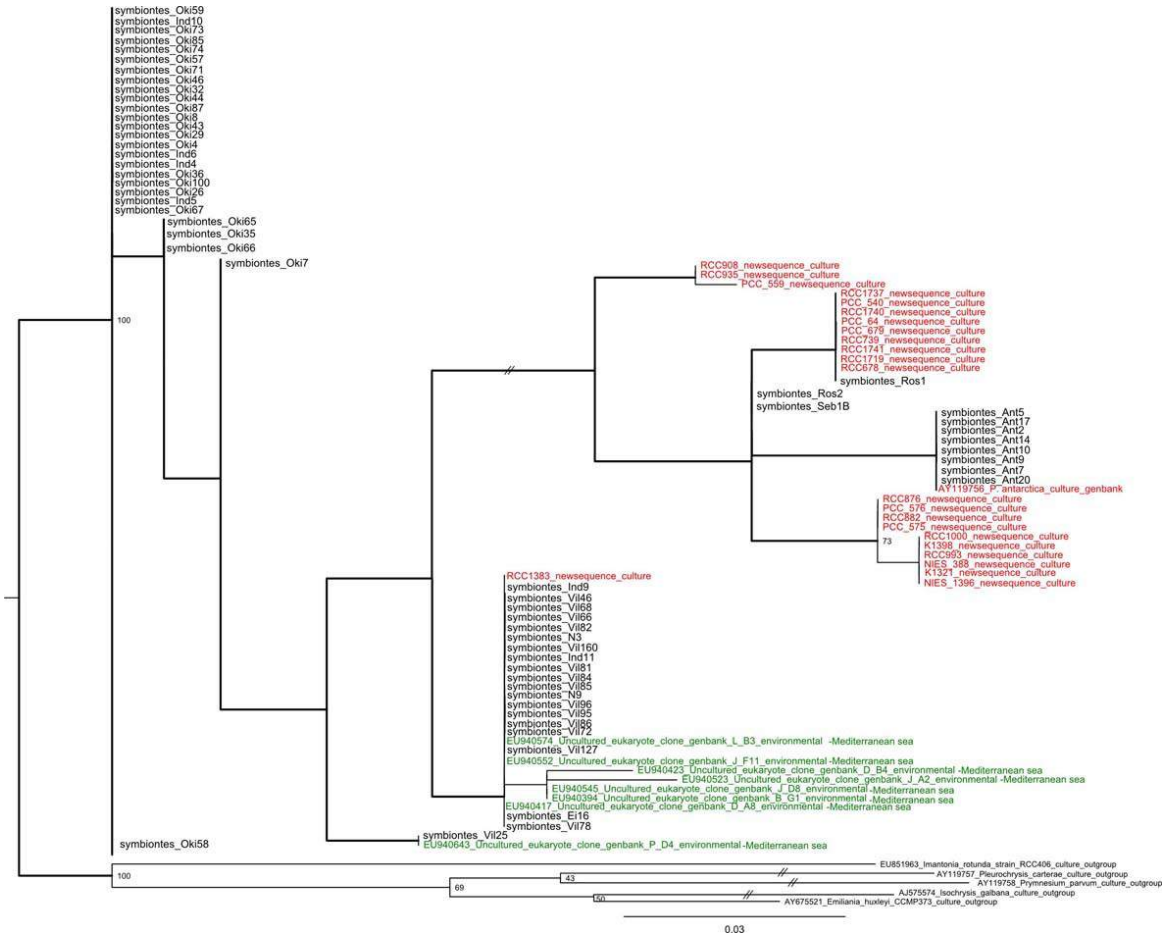


Fig. S1. Maximum Likelihood phylogeny inferred from RAxML with the psbA gene. Black: symbiont sequences; red: culture sequences; green: environmental sequences. Only RAxML bootstrap values ≥ 70% are shown.

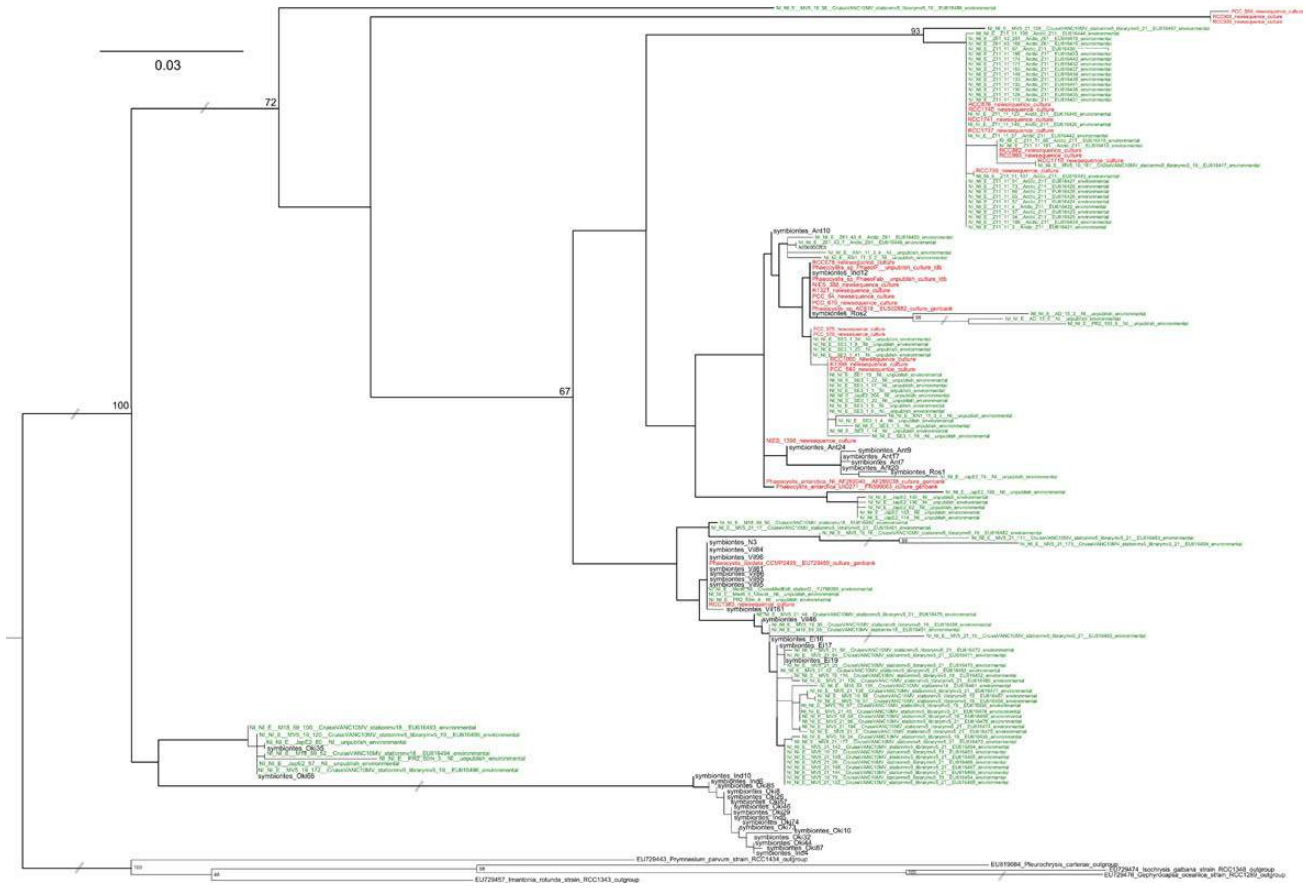


Fig. S2. Maximum Likelihood phylogeny inferred from RAxML with the 28S rDNA marker. Black: symbiont sequences; red: culture sequences; green: environmental sequences. Only RAxML bootstrap values ≥ 70% are shown.

Fig. S3. Maximum Likelihood phylogeny inferred from RAxML with the 18S rDNA marker. Black: symbiont sequences; red: culture sequences; green: environmental sequences. Only RAxML bootstrap values ≥ 70% are shown.

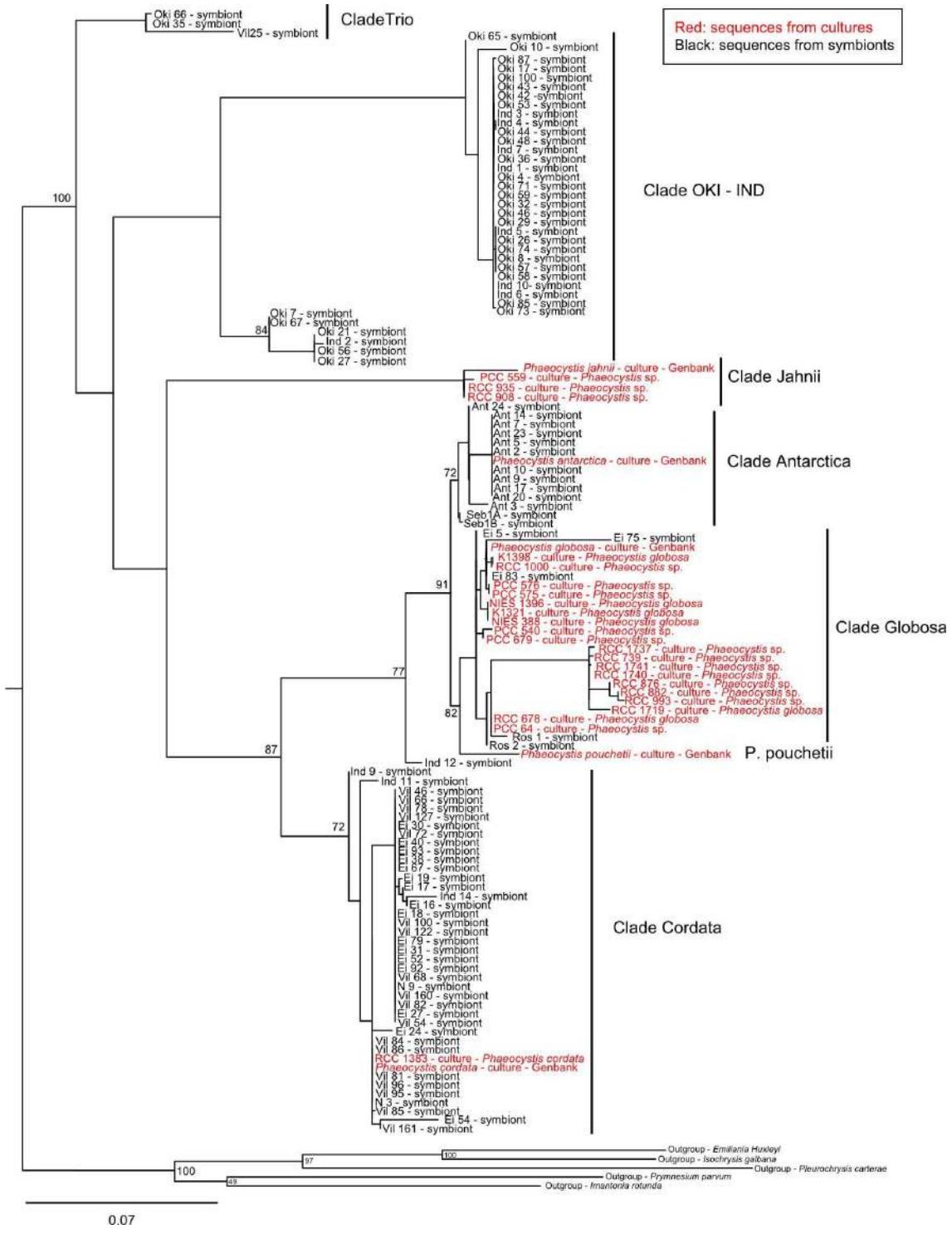


Fig. S4. Maximum Likelihood phylogeny inferred from RAxML with the concatenated genes 18S and 28S rDNA, rbcL and psbA . Only RAxML bootstrap values ≥ 70% are shown.

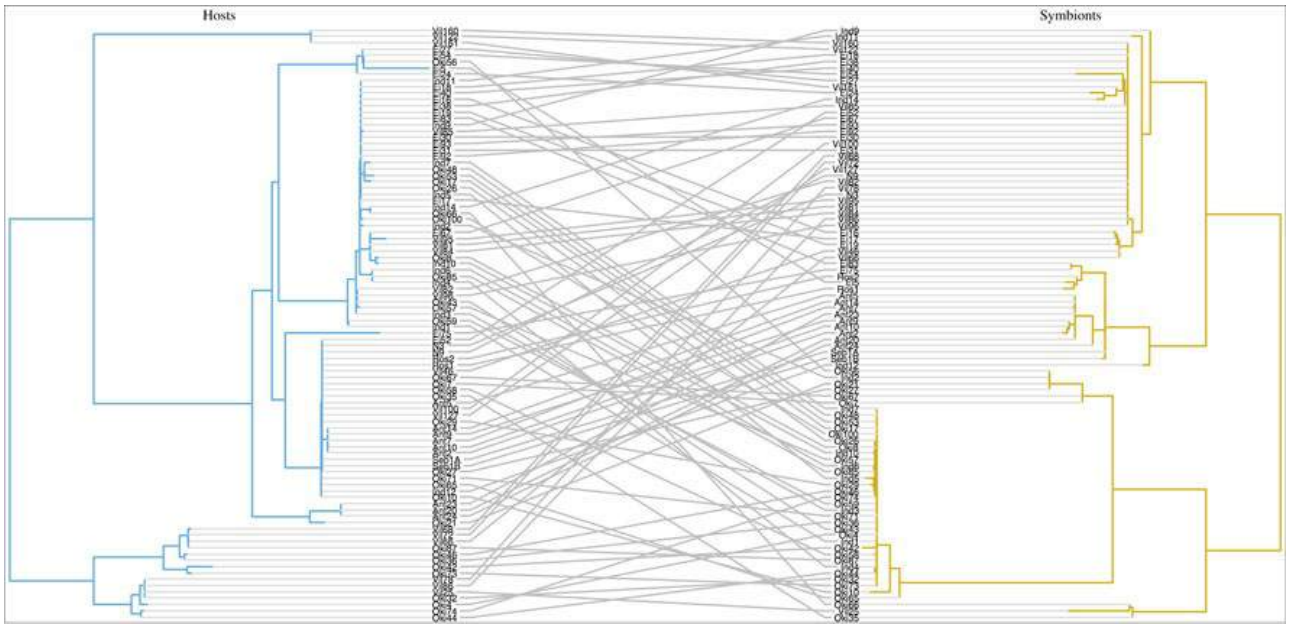


Fig. S5. Pattern of host-symbiont associations, based on the ML phylogenetic trees for 94 acantharian taxa (host) and their symbiotic microalgae (*Phaeocystis* sp.)

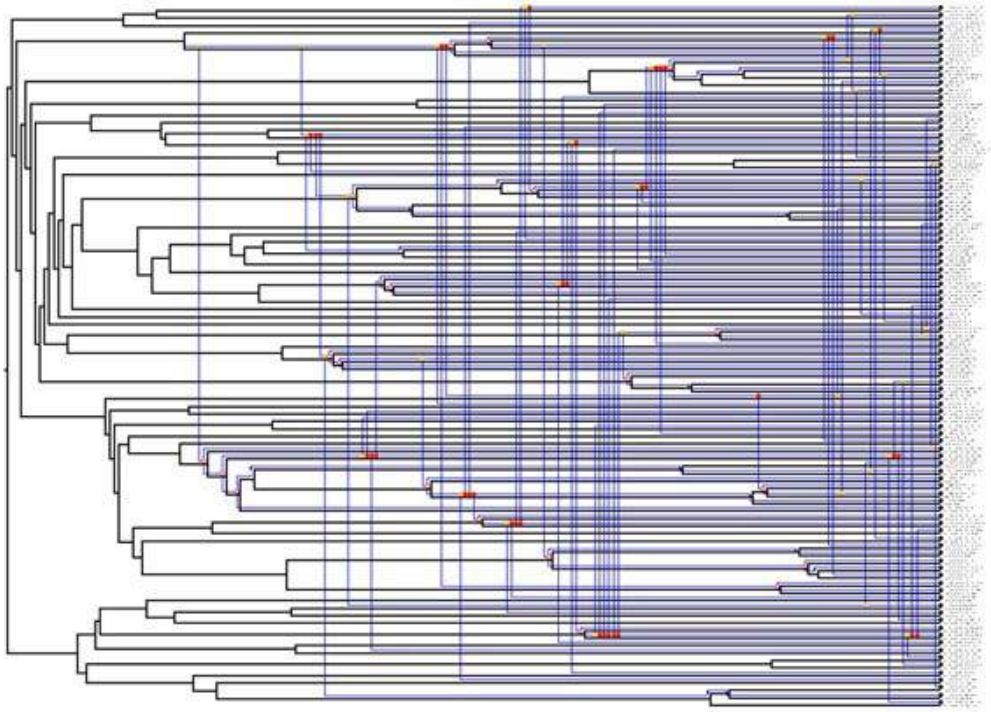


Fig. S6. Optimal and significant tree reconciliation produced by the software Jane. The symbiont tree is mapped onto the host tree by mixing cospeciations (white circles), duplications (colored circles), host-switches (arrows) and sorting events (dashed lines).

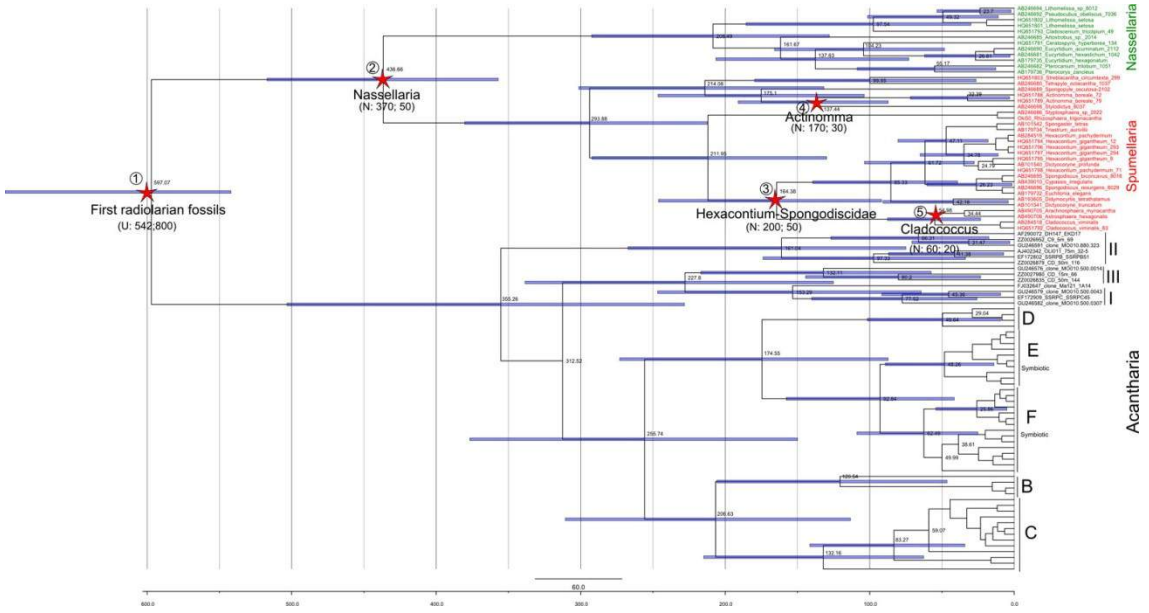


Fig. S7. Time-calibrated tree of the radiolarian groups, Acantharia, Spumellaria and Nassellaria, using 5 fossil-based calibration points (red stars), based on nuclear 18S rDNA. Node divergences were estimated with a Bayesian relaxed clock model and GTR + I + G model, implemented in the software package BEAST. Blue bars indicate the 95% highest posterior density (HPD) intervals of the posterior probability distribution of node ages.

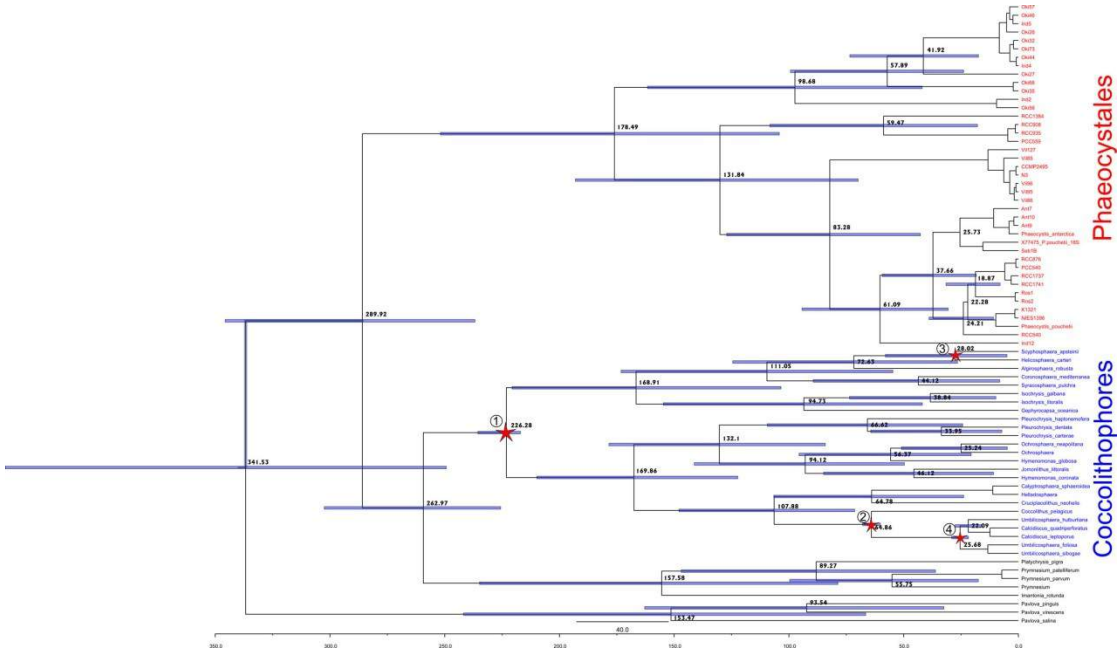
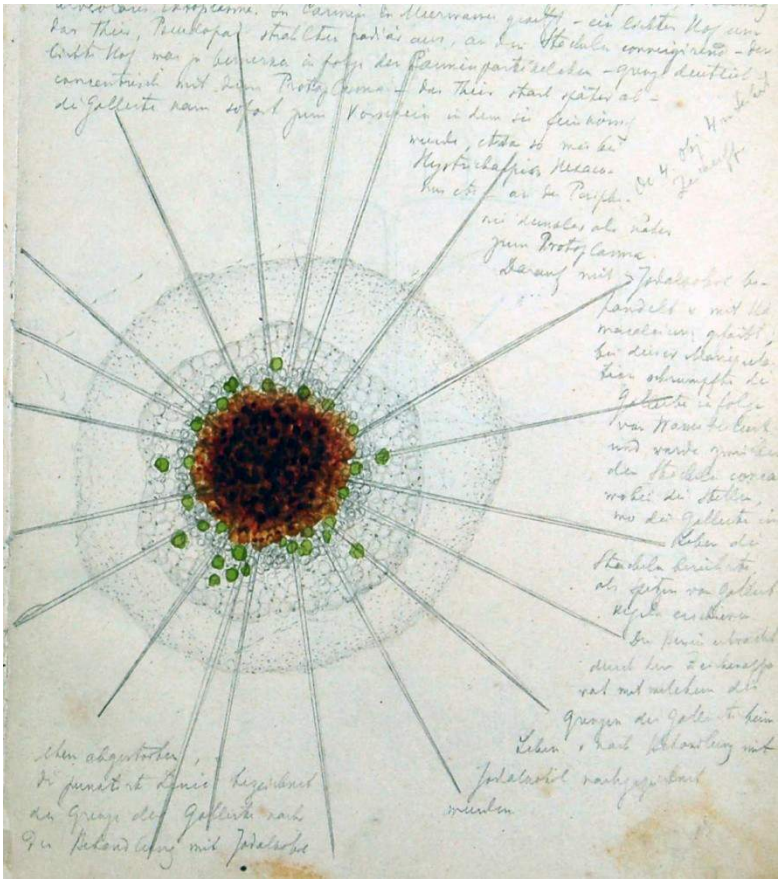


Fig. S8. Time-calibrated tree of haptophytes, including the Phaeocystales, using 4 fossil-based calibration points from coccolithophores (red stars), based on nuclear 18S and 28S rDNA, and plastidial rbcL. Node divergences were estimated with a Bayesian relaxed clock model and the GTR + I + G model, implemented in the software package Beast. Blue bars indicate the 95% highest posterior density (HPD) intervals of the posterior probability distribution of node ages.

B) Multiple microalgal partners in symbiosis with the acantharian *Acanthochiasma* sp. (Radiolaria)



***Acanthochiasma* par Schewiakoff**

Multiple microalgal partners in symbiosis with the acantharian *Acanthochiasma* sp. (Radiolaria)

Johan Decelle[1,2,], Raffaele Siano[3], Ian Probert[4], Camille Poirier[1,2], Fabrice Not[1,2]

1- UPMC Univ. Paris 06, UMR 7144, Station Biologique, 29680 Roscoff, France
2- CNRS, UMR 7144, Station Biologique, 29680 Roscoff, France
3- IFREMER, Centre de Brest, DYNECO/Pelagos, BP 70 29280 Plouzané, France
4- UPMC Univ. Paris 06, FR 2424, Station Biologique, 29680 Roscoff, France

Published in ***Symbiosis*, 2012 vol.58**

Abstract

Acantharia (Radiolaria) are widespread and abundant heterotrophic marine protists, some of which can host endosymbiotic eukaryotic microalgae. Although this photosymbiotic association was first described at the end of the 19th century, the diversity of the symbiotic microalgae remains poorly characterized. Here, we examined the identity of the microalgae associated with the acantharian species *Acanthochiasma* sp. by sequencing partial 18S and internal transcribed spacer (ITS) ribosomal DNA genes from cultured symbionts and directly from isolated holobiont specimens. Single *Acanthochiasma* cells contained multiple symbiotic partners, including distantly related dinoflagellates (*Heterocapsa* sp., *Pelagodinium* sp., *Azadinium* sp. and *Scrippsiella* sp.) as well as a haptophyte (*Chrysochromulina sp.*). This original association of multiple symbiotic microalgae within a single host cell raises questions about the specificity and functioning of the relationship. These microalgae exhibit the common ecological feature of being abundant and widely distributed in coastal and oceanic waters, some occasionally forming extensive blooms. Some of the microalgal genera found in association with *Acanthochiasma* (i.e. *Pelagodinium* and *Chrysochromulina*) are known to occur in symbiosis with other heterotrophic protists such as Foraminifera and other Radiolaria, whereas *Heterocapsa*, *Scrippsiella* and *Azadinium* have never previously been reported to be involved in putative symbiotic relationships. The unusual association unveiled in this study contributes to our understanding of the ecological and evolutionary significance of photosymbiosis in Acantharia and also provides new insights into the nature of such partnerships in the planktonic realm.

Introduction

Symbiosis defines a mode of life whereby distantly related organisms live in close association with each other over multiple generations (de Bary, 1878). Symbiotic relationships encompass a wide range of interaction strategies and mechanisms, from parasitism to mutualism. In the 19th century, the first investigations of symbiosis were carried out on lichens, Cnidaria and Radiolaria. Originally considered parasitic (Cienkowsky 1871), Brandt (1881) used physiological experiments to demonstrate that the symbiotic microalgae of these host types are in fact independent entities that are highly beneficial for the nutrition of the hosts, and formally described the symbionts as *Zooxanthella nutricula*. Heterotroph-phototroph symbioses (photosymbioses) in reef-dwelling invertebrates have since been studied extensively and are typically regarded as mutualistic relationships. The symbiont provides photosynthetically derived products to the host, which in turn maintains a sheltered and nutrient-rich environment for the symbiont (Muscatine et al. 1984; Yellowlees et al. 2008). By comparison, photosymbiotic relationships in planktonic organisms have received little attention despite their abundance and widespread distribution in the photic zone of the world ocean (Stoecker et al. 2009). Certain large heterotrophic unicellular eukaryotes, such as Foraminifera and Radiolaria (Acantharia and Polycystinea), are known to harbor microalgal symbionts, especially in oligotrophic oceanic waters where they significantly contribute to primary production (Michaels 1988). In these cases, the symbiotic association is apparently obligatory for the hosts, and in most cases they must acquire their microalgal symbionts from the surrounding environment at each generation (i.e. horizontal transmission). Microalgae involved in symbiosis are typically considered to be specialized for a symbiotic life-style, but the degree of benefit and the dependence on symbiosis for their survival is less clear (Douglas and Smith 1989; Wooldrige 2010).

The diversity of symbiotic microalgae in the planktonic realm has been poorly characterized, essentially due to the alteration or loss of diagnostic morphological features of the symbiotic cells *in hospite* (e.g. loss of flagella, scales, cell wall, etc.) and the low success rate of culture isolation. In such cases, DNA-based techniques using molecular markers such as the 18S rDNA gene have been proven to be an effective means of obtaining good taxonomic resolution. Molecular studies have demonstrated that the colloquial name "zooxanthellae" (from *Zooxanthella nutricula*) in fact entails a broad diversity of microalgal lineages. In contrast to most invertebrates in benthic-coastal ecosystems that exclusively host the dinoflagellate genus *Symbiodinium* (order Suessiales), Radiolaria do not seem to be specialized for a single group of microalgae, but can live in symbiosis with representatives from distinct eukaryotic lineages, such as haptophytes, dinoflagellates (order Peridiniales) and prasinophytes (Decelle et al. 2012a; Gast and Caron 1996, 2001). Ultrastructural and molecular analyses have shown that these three different lineages can even occur as symbionts in a single polycystine genus, *Spongodrymus* (Anderson 1983; Gast et al. 2000). The cyanobacteria *Synechococcus* sp. and *Prochloroccocus* sp. can also be photosymbionts of radiolarians, as described in the polycystine *Dictyocoryne sp.* (Foster et al. 2006a,b; Yuasa et al. 2012). Thus, Radiolaria-microalga photosymbiotic partnerships seem to be diverse, yet few associations have been examined with molecular

techniques at relevant taxonomic scales, preventing an accurate description of taxon-specific patterns.

The Acantharia are a monophyletic group of radiolarians characterized by their mineral endoskeleton composed of celestite (strontium sulfate). Acantharia generally outnumber their Polycystinea and Foraminifera counterparts in coastal and open ocean waters (Michaels 1995). A recent phylogenetic analysis of these to date uncultivated protists described nine molecular clades (I to III and A to F) and highlighted the need for a revision of the taxonomic framework that dates back to the pioneering works of Haeckel (1888) and Schewiakoff (1926) (Decelle et al. 2012b). Acantharia are known to be active predators using their axopods to capture a wide range of small protists (e.g. ciliates) that are rapidly digested in the outer part of the cytoplasm, called the ectoplasm (Swanberg and Caron 1991). Like other radiolarians, some species also have an indirect photosynthetic capacity by hosting symbiotic microalgal cells in their endoplasm. In the Equatorial Pacific and Gulf of Mexico, about 40% of Acantharia were found to bear algal symbionts, contributing up to 80% of acantharian biomass (Stoecker 1996), and in the Sargasso Sea and Pacific Ocean, the symbiotic forms can occasionally account for up to 20% of the total primary production in surface waters (Michaels 1988). Three decades ago, *in hospite* ultrastructural investigations described the photosymbionts of Acantharia as being either haptophytes or dinoflagellates on the basis of organelle morphology (Febvre and Febvre-Chevalier 1979; Hollande and Carré 1974). Since then, there have been few attempts to provide a more precise taxonomic assignation of both partners. A molecular study recently demonstrated that Acantharia from clades E and F (orders Symphiacanthida and Arthracanthida) live in symbiosis worldwide with the haptophyte genus *Phaeocystis*, a phytoplankton taxon that is very abundant and widespread in its free-living condition (Decelle et al. 2012a). *Acanthochiasma* sp. (order Holacanthida) from the earlier diverging clade B is also known to harbor intracellular microalgae (Schewiakoff 1926).

The objective of this study was to genetically identify the microalgal cells associated with the acantharian species *Acanthochiasma* sp.. To do so, we analyzed symbiont 18S and internal transcribed spacer (ITS) ribosomal DNA gene sequences directly from the holobiont (host and microalgae) and from clonal cultures of microalgae obtained after microdissection and single-cell isolation. Identification of these symbiotic microalgae will contribute to our understanding of the evolution, functioning and ecology of photosymbiosis in Acantharia.

Materials and Methods

Sample collection

Twelve cells of the acantharian species *Acanthochiasma* sp. were collected in the Mediterranean Sea (Villefranche-sur-Mer; 43°40.552 N, 7°18.447 E) in September 2010 and 2011 by gently towing a plankton net (150 µm mesh size) at the subsurface. Morphological identification was performed based on the diagnostic features described in the acantharian taxonomic framework of Schewiakoff (1926). Cells were immediately isolated from the raw plankton sample using glass micropipettes under inverted microscopy, transferred into 0.2-µm-filtered seawater, and incubated for several hours to allow self-cleaning (debris and particles are removed and prey digested). After incubation, each acantharian cell was rinsed several times in 0.2-µm-filtered seawater, and observed through an epifluorescence microscope to verify the red chlorophyll auto-fluorescence indicative of intracellular microalgae. To identify the microalgae, molecular analyses were performed either on cultures of microalgae isolated after microdissection of the host or directly on the holobiont cell. Detailed information on each isolated specimen, such as images and collection data, are available at http://abims.sb-roscoff.fr/renkan.

Culture isolation and molecular analysis

Acanthochiasma sp. cells were microdissected under an inverted microscope and the intracellular microalgae subsequently isolated by micropipette. Individual symbiont cells were transferred into a single well of a multi-well plate containing K/10 (-Si,+Ni) medium (Keller et al. 1987) and maintained at 19°C with illumination provided by daylight neon tubes at an intensity of ca. 30 µEinsteins.m^{-2}.s^{-1} and a photoperiod (L:D) of 14:10 hours. Successful cultures were harvested during exponential growth phase and concentrated by centrifugation. Total nucleic acids were extracted using the Nucleospin® RNA II kit (Macherey-Nagel, Hoerdt, France) and quantified using a Nanodrop ND-1000 Spectrophotometer (Labtech International, France). The eukaryote-specific primers 63F/1818R were used to amplify the 18S rDNA of the different cultures (information on primers is detailed in Online Ressource 1). In addition, the specific primers ITS-Cer/D1R-R were used to amplify the partial Internal Transcribed Spacer (ITS) of the dinoflagellate cultures. Polymerase Chain Reactions (PCRs) were performed using Phusion high-fidelity DNA polymerase (Finnzymes) in a 25-µl reaction volume as follows: an initial denaturation step at 98 °C for 30 sec, followed by 35 cycles at 98 °C for 10 sec, at 50 °C (18S rDNA) or at 53°C (ITS) for 30 sec, and at 72 °C for 30 sec, with a final elongation step of 10 minutes at 72 °C. PCR products were then purified by EXOSAP-IT (GE Healthcare Bio-Sciences Corp.) and bidirectionally sequenced using the ABI-PRISM Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems). Note that the PCR parameters and the associated steps were similar throughout this study, except for the annealing temperature that changed according to the primers (Online Ressource 1). Cultures of symbiotic microalgae were deposited in the Roscoff Culture Collection (RCC - http://www.sb-roscoff.fr/Phyto/RCC/): AR1 (RCC 3001), AR2 (RCC 3002), AC24-2 (RCC3003), AC24-0 (RCC3004, culture lost), AC24-1 (RCC3005), AC24-3 (RCC 3006).

Molecular analysis on the holobiont cell

Each holobiont isolated was photographed prior to transfer into an extraction buffer (GITC) and DNA extraction was performed as explained in Decelle et al. (2012b). Molecular identification of the host *Acanthochiasma* sp. was first performed in order to check whether it entails cryptic species and thus to assess the specificity of the symbiosis. To do so, the 28S rDNA (D1/D2 domains) and the ITS (including the ITS1, 5,8S and the ITS2) of different host individuals were PCR amplified with Radiolaria-specific primers (Online Ressource 1). Since the symbiotic microalgae of Acantharia have previously been described as belonging to either haptophytes or dinoflagellates (Febvre and Febvre-Chevalier 1979; Hollande and Carré 1974), specific primers for these two lineages were selected to specifically amplify their partial 18S rDNA and ITS ribosomal genes from each holobiont cell. For the 18S rDNA genes of haptophytes and dinoflagellates, the specific primers Prym429F/Prym02R and DIN464F/S69 were used respectively (Online Ressource 1). The dinoflagellate-specific ITS primers mentioned above (ITS-Cer/D1R-R) were also used for the holobiont PCR. A nested PCR was sometimes necessary to obtain a detectable quantity of amplicons, especially for the 18S rDNA genes of the haptophytes. In such cases, the first amplification was performed with the general primers 63F and 1818R (25 cycles), and 1µl of 10X diluted amplicon was used for re-amplification with the internal specific primers of haptophytes or dinoflagellates. Sequences obtained in this study were deposited in GenBank under accession numbers JX661018 to JX661041.

Table S1: Primers used in this study

Analysis	Gene targeted	Forward primer	Reverse primer	Specificity	Tm (°C)	Reference
host identity	28S	28S_Rad2: 5'- TAA GCG GAG GAA AAG AAA-3'	ITS_a3_R: 5'-TCA CCA TCT TTC GGG TCC CAA CA-3'	Radiolaria	50	Ando et al 2009
host identity	ITS	ITS_18S_Rad1: 5'-ACC GCC CGT CGC TCC TAC CG-3'	ITS_28S_Rad1: 5'-CCC TCA CGG TAC TTG TTC GC-3'	Radiolaria	59	Ando et al 2009
symbiont identity	18S	63F: 5'- ACG CTT GTC TCA AAG ATT A-3'	1818R: 5'-ACG GAA ACC TTG TTA CGA-3'	Eukaryotes	50	Lepère et al 2010
symbiont identity	18S	DIN464F: 5'-TAACAATACAGGGCATCCAT-3'	S69: 5'- CCGTCADTTCCTTTRAGDTT-3'	Dinoflagellates	53	Gomez et al 2009
symbiont identity	18S	Prym429F: 5'- GCG CGT AAA TTG CCC GAA-3'	Prym 02R: 5'- GGA ATA CGA GTG CCC CTG AC-3'	Haptophytes	55	Coolen et al 2004; Simon et al 2000
symbiont identity	ITS	ITS-Cer : 5-'GTCGCTCCTACCGA TGAGT-3'	D1R-R : 5'-TATGCTTAAATTCAGCAGGT-3'	Dinoflagellates	53	Simon et al (unpub.); Scholin et al, 2004

Phylogenetic analyses

Four distinct matrices (18S rDNA dinoflagellates, ITS dinoflagellate Peridiniales, ITS dinoflagellate Suessiales and 18S rDNA haptophytes) were built from the microalgal sequences obtained in this study and additional sequences retrieved from GenBank. Separate analyses on the ITS of Peridiniales and the ITS of Suessiales were required to improve the alignment quality. The four matrices were aligned with the program MUSCLE, implemented in Seaview v.4.0 (Gouy et al. 2010), and the optimal model of evolution was determined under the AIC, AICc and BIC criteria using MEGA v5.05 (Tamura et al. 2011). Phylogenetic reconstructions were

performed using Maximum Likelihood (ML) and statistical support assessed by performing 100 bootstrap replicates with MEGA v5.05.

Results

Sampling and molecular identification of host cells

We collected twelve *Acanthochiasma* sp. host cells that always contained microalgae in the endoplasm even after at least 1 hour of incubation in 0.2-µm filtered seawater (Fig. 1; Online Resource 2). Microalgal cells were of different sizes and shapes, from ca. 4 to 15 µm, suggesting the occurrence of distinct eukaryotic lineages. Based on epifluorescence microscopy, no orange signal characteristic of cyanobacteria such as *Synechococcus* was detected in the specimens observed (Fig. 1).

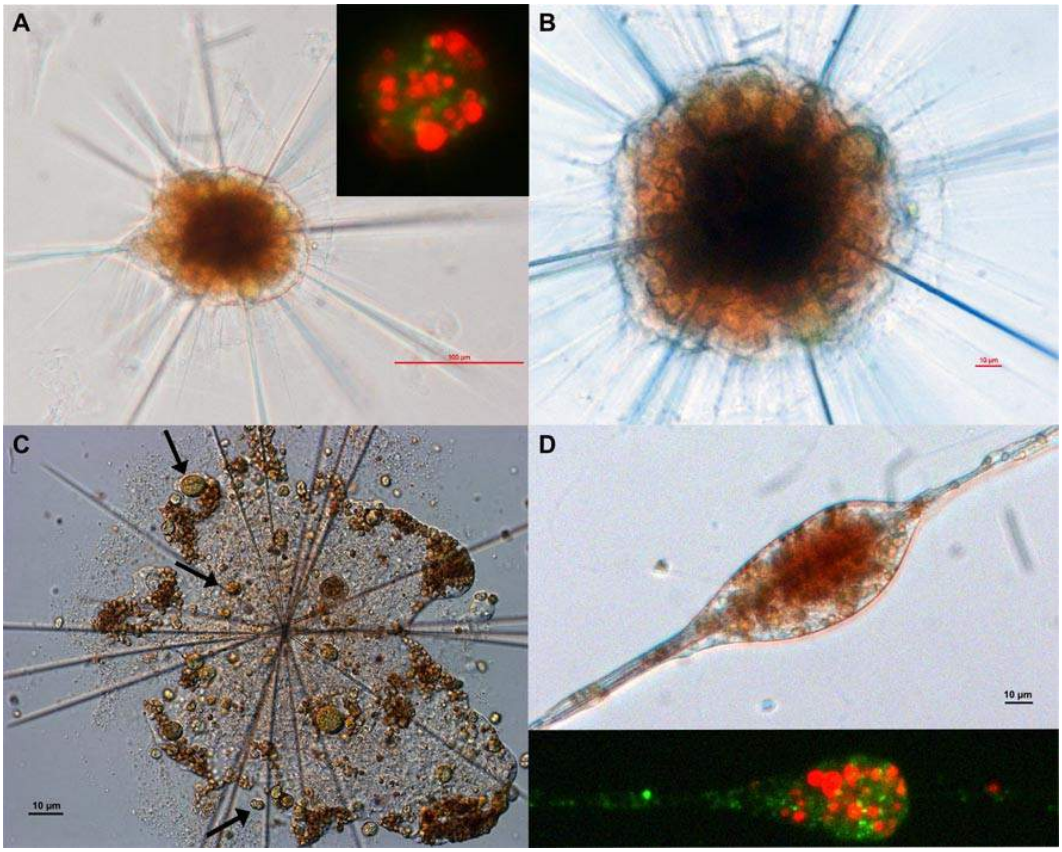


Fig. 1 Microscope observations of the host acantharian species *Acanthochiasma* sp. that harbors microalgae of different sizes and shapes (e.g. small gymnoid cells and large round cells) in its endoplasm, as indicated by the black arrows in C. The insets of A and D highlight the red auto-fluorescence of chlorophyll-containing microalgae where different cell sizes are also observed. In C and D, *Acanthochiasma* sp. triggers encystment upon intense manipulation under the microscope

We sequenced partial 28S rDNA (~700 bp), and entire ITS (ITS1 + 5.8S rDNA + ITS2) ribosomal genes (~930 bp) for eight and three host specimens, respectively. For both of these phylogenetic markers, the sequences obtained were 100% identical (Table 1). Having identical genetic sequences, including the ITS2 that provides resolution at the species level for many protistan lineages (Pawlowski et al. 2012, Coleman 2007), all of the host specimens collected in this study can reasonably be considered to belong to a single taxonomic entity. Note that the acantharian specimens from clade B in Decelle et al (2012b) were erroneously identified as *Acanthocyrta haeckeli*, but correspond in fact to *Acanthochiasma* sp..

Molecular identification of the intracellular microalgae

We retrieved sequences of dinoflagellates from both cultures and holobiont cells, whereas haptophyte sequences were only obtained from holobiont cells (Table 1, Online Resource 2). Sequences obtained from cultures and holobiont cells are indicated hereafter with the suffixes -c and -h, respectively.

Table 1 Information about the twelve host cells of *Acanthochiasma* sp. collected in September 2010 and 2011, including the genetic identity of their intracellular microalgae

Host	Methods ID	Collection Date	Host 28S	Host ITS	*Chrysochromulina* clade B2	*P. beii*	*Heterocapsa* sp.	*Scrippsiella* sp.	*Azadinium* sp.
Vil20	holobiont	18th sept 2010	identical	identical	X				
Vil45	holobiont	18th sept 2010	identical	identical	X				
Vil51	holobiont	18th sept 2010	identical	identical	X				
Vil64	holobiont	19th sept 2010	identical		X		X		
PEC14	holobiont	21st sept 2011	identical		X		X		
Vil39	holobiont	18th sept 2010	identical				X		
Acanth23	holobiont	22nd sept 2011	identical				X		
PEC18	holobiont	22nd sept 2011	identical						X
Vil60	holobiont	19th sept 2010				X			
AR1	culture	22nd sept 2011				AR1 (RCC 3001)			
AR2	culture	22nd sept 2011				AR2 (RCC 3002)			
AC24	culture	22nd sept 2011				Ac 24-0 (RCC 3004) Ac 24-2 (RCC 3003)	Ac 24-1 (RCC 3005)	Ac 24-3 (RCC 3006)	

Dinoflagellates: The twelve 18S rDNA sequences recovered from nine host individuals belonged to the dinoflagellate orders Peridiniales and Suessiales (Fig. 2). In the former, five sequences fell into a monophyletic group representing the genus *Heterocapsa* (bootstrap values (BV) = 56%), and three of these (Vil 39-h, Vil 64-h and Acanth 23-h) grouped more specifically with reference sequences of *Heterocapsa rotundata* (BV = 94%), sharing up to 99% identity. In addition, two symbiont sequences, AC 24-3-c and PEC 18-h, were related to the genera *Scrippsiella* (BV = 100%) and *Azadinium* (BV = 100%), respectively. None of the sequences retrieved from *Acanthochiasma* grouped with the Peridiniales clade named "*Scrippsiella nutricula*", which encompasses several sequences of symbiotic microalgae living within the jellyfish *Vellela vellela* and other Radiolaria (Collodaria and Spumellaria).

Five 18S rDNA sequences (AC 24-0-c, AC 24-2-c, Vil 60-h, AR1-c and AR2-c) from four distinct host cells fell into the order Suessiales, and were closely related to the reference sequence of *Pelagodinium béii* (99% identity), previously known as *Gymnodinium béii* (Siano et al. 2010; Spero 1987). The sequences of *P. béii* formed a monophyletic group (BV = 74%) and included the sequence AC 24-0-c. The sequences AC 24-2-c, Vil 60-h, AR1-c and AR2-c appeared to constitute a distinct subgroup that branched with the *P. béii* group with low support (BV < 50%). Microalgal sequences from the same host cell were genetically different for the host AC 24, indicating that a single host cell can live with different genotypes or strains of *P. béii*. Remarkably, the host individual AC 24 simultaneously possessed microalgae of the genera *Scrippsiella*, *Pelagodinium* and *Heterocapsa* (Table 1).

The 18S rDNA phylogenetic tree of dinoflagellates was weakly resolved as previously recognized (Saldarriaga et al. 2004). In order to confirm and improve the phylogenetic placement of the dinoflagellates associated with *Acanthochiasma* sp., we amplified and sequenced the more resolutive locus ITS that has been shown to be a good marker for dinoflagellate phylogeny at the species level (LaJeunesse 2001; Stern et al. 2012). Phylogenetic analyses were performed separately for the orders Suessiales and Peridiniales.

Fig. 2 18S rDNA phylogeny of dinoflagellates including sequences of symbiotic microalgae of different host taxa (blue) and microalgae associated with the host *Acanthochiasma* sp. (bold red font). The phylogenetic tree was built by a Maximum Likelihood (ML) analysis based on 1534 aligned positions, 86 taxa and the GTR+G+I model. The tree was rooted with 3 sequences of *Alexandrium* sp. from the order Gonyaulacales as the outgroup. ML bootstrap values > 50% are shown at nodes (100 pseudo-replicates)

ITS of Peridiniales: We obtained four ITS sequences from the Peridiniales identified with the 18S rDNA gene (Fig. 3). The sequence AC 24-1-c branched within a clade including two sequences of the species *Heterocapsa pygmaea* and one sequence of *H. triquetra* (BV = 100%). The sequence Vil 39-h, which was very close to *H. rotundata* according to the 18S rDNA phylogeny, was located in a clade between an undescribed *Heterocapsa* taxon and several sequences of *H. arctica* (BV = 88%). Note that the ITS sequence of *H. rotundata* is not available in public databases. The phylogenetic placement of PEC 18-h and AC 24-3-c was consistent with that obtained with the 18S rDNA gene, being included in the genera *Azadinium* sp. (BV = 100%) and *Scrippsiella* sp. (BV = 83%), respectively.

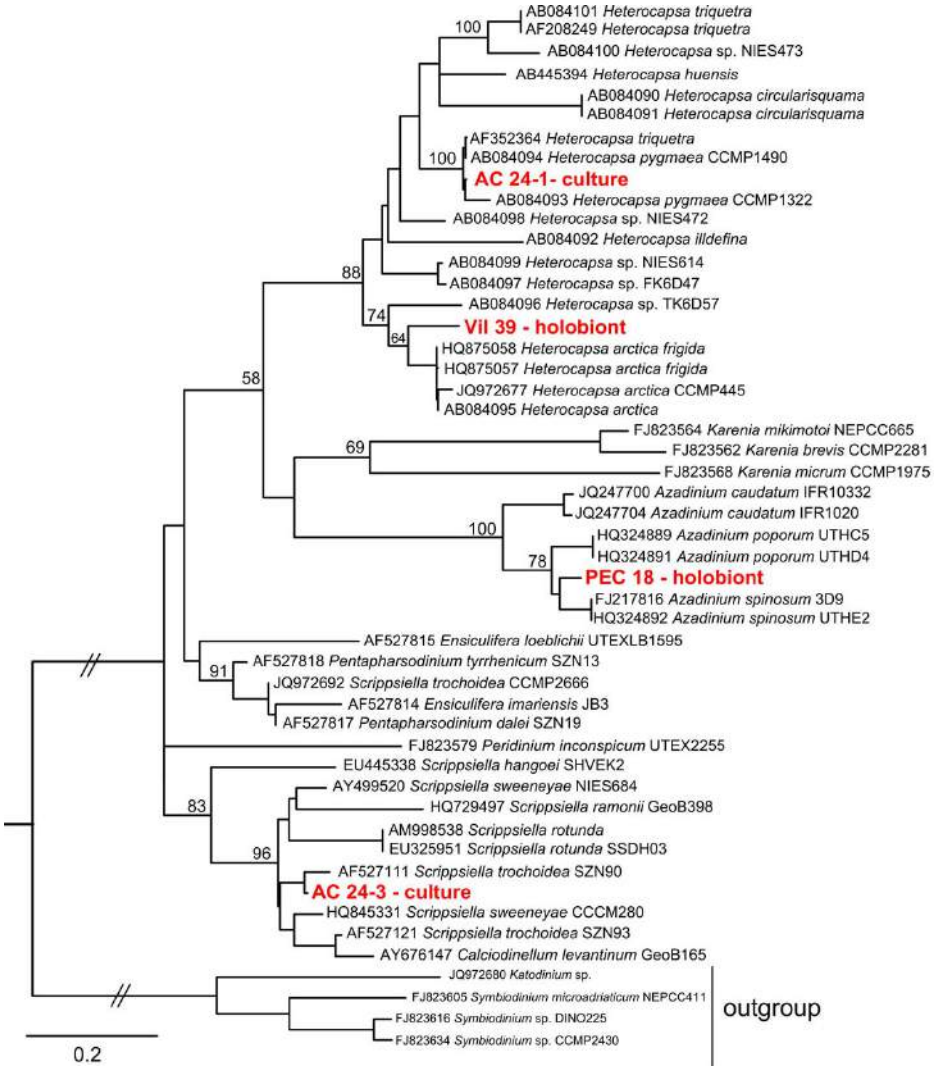


Fig. 3 ITS rDNA phylogeny of the dinoflagellate order Peridiniales including sequences of microalgae associated with the host *Acanthochiasma* sp. (bold red font). The phylogenetic tree was built by a Maximum Likelihood (ML) analysis based on 731 aligned positions, 50 taxa and the GTR+G model. The tree was rooted with 4 sequences belonging to the sister order Sussiales as an outgroup. ML bootstrap values > 50% are shown at nodes (100 pseudo-replicates)

ITS of Suessiales: The phylogenetic tree constructed with the ITS rDNA marker confirmed the close affiliation of the Suessiales symbionts to *Pelagodinium béii* (Fig. 4). This fast-evolving gene resolves four sub-clades within *Pelagodinium* (P1a, P1b and P2a, P2b) that were defined by Shaked and de Vargas (2006). The sequence AC 24-0-c clustered with sub-clade P1b with relatively high support (BV = 85%). The ITS sequences AC 24-2-c and AR1-c that formed a distinct sub-group as in the 18S rDNA phylogeny, were included with stronger support in the monophyletic *P. béii* group (BV = 64%), with no clear affinity with sub-clades P1 and P2 (BV = 51%). PCR amplifications of ITS from Vil 60-h and AR2-c were not successful, but it can reasonably be supposed that their sequences would branch with AC24-2-c and AR1-c, as in the 18S rDNA phylogeny (Fig. 2).

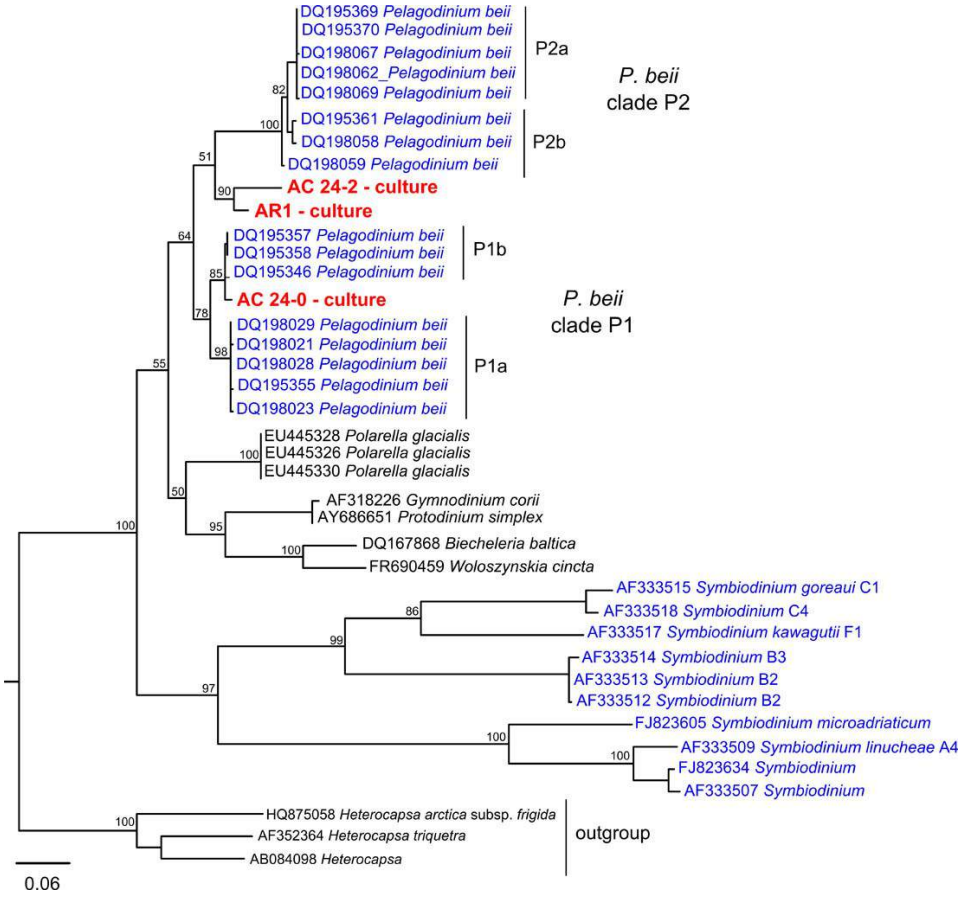


Fig. 4 ITS rDNA phylogeny of the dinoflagellate order Suessiales including sequences of symbiotic microalgae of different host taxa (blue) and microalgae associated with the host *Acanthochiasma* sp. (bold red font). The phylogenetic tree was built by a Maximum Likelihood (ML) analysis based on 813 aligned positions, 39 taxa and the HKY+G model. The tree was rooted with 3 sequences as outgroup of *Heterocapsa* sp. belonging to the sister order Peridiniales. ML bootstrap values > 50%. are shown at nodes (100 pseudo-replicates). Sub-clades P1a, P1b, P2a and P2b of the species *Pelagodinium béii* are labeled according to Shaked and de Vargas (2006)

Haptophytes: Five 18S rDNA sequences from five *Acanthochiasma* sp. cells were related to the haptophytes, and more specifically to the order Prymnesiales (Fig. 5). The phylogenetic reconstruction indicated that the five sequences grouped within the well-supported clade B2 (BV = 100%), which corresponds to the core *Chrysochromulina* species (Edvardsen et al. 2011). While four sequences grouped together (Vil 20-h, PEC 14-h, Vil 51-h and Vil 64-h), Vil 45-h appeared to belong to another sub-clade. Prymnesiales clade B2 also includes three sequences (named symbiont of Foraminifera 1, 3 and 4) that correspond to microalgae that were found in symbiosis with planktonic Foraminifera (Gast and Caron 2001).

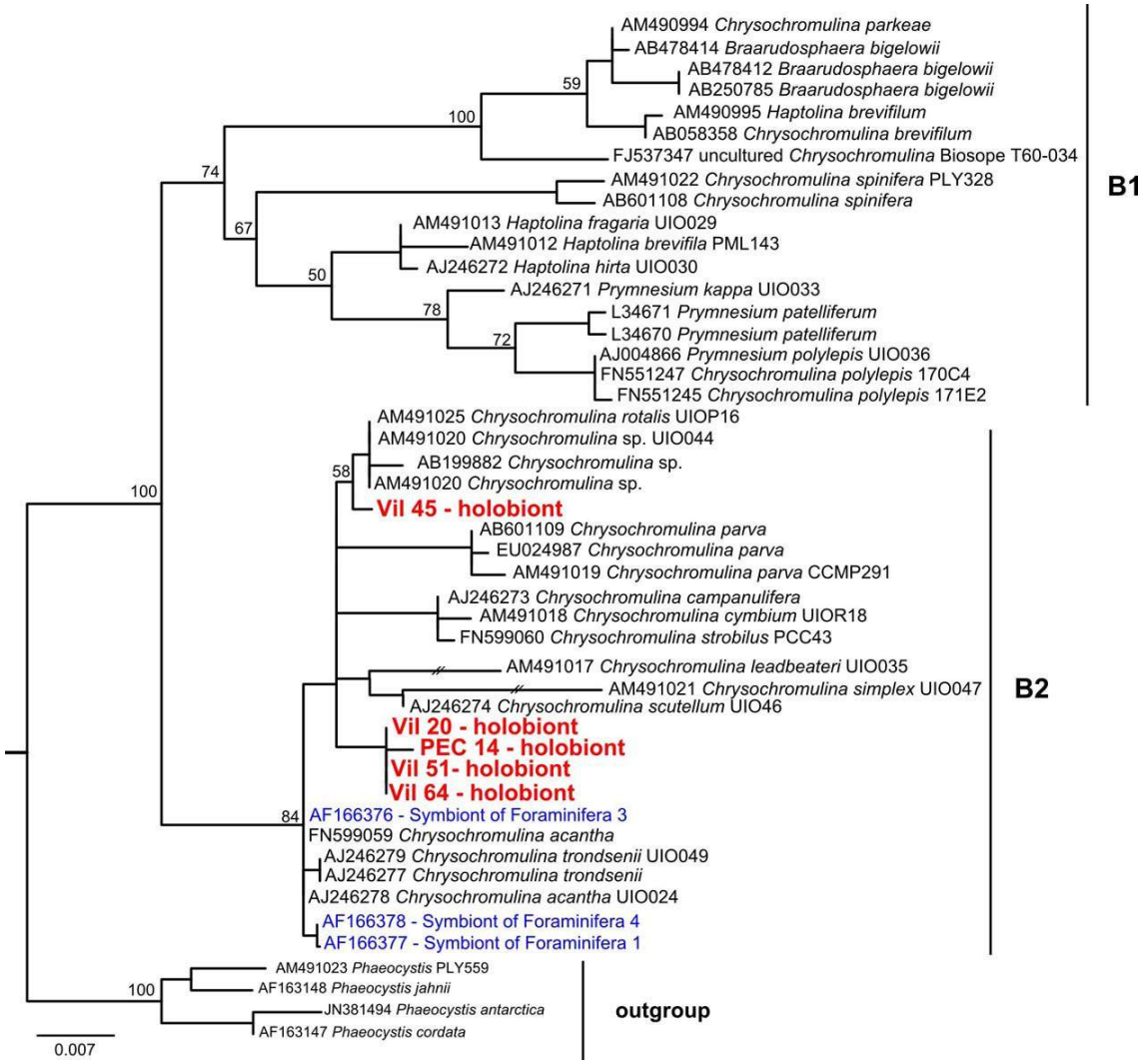


Fig. 5 18S rDNA phylogeny of haptophytes including the sequences of microalgae associated with the host *Acanthochiasma* sp. (bold red font). The phylogenetic tree was built by a Maximum Likelihood (ML) analysis based on 717 aligned positions, 47 taxa and the K2+G model. The tree was rooted with 4 sequences of *Phaeocystis* sp. as an outgroup. ML bootstrap values > 50% are shown at nodes (100 pseudo-replicates). Clades B1 and B2 are labeled according to Edvardsen et al. (2011). In clade B2, 3 sequences of symbiotic microalgae (highlighted in blue) obtained from planktonic Foraminifera in Gast et al (2000) are shown

Discussion

Prey or true endosymbiotic microalgae?

We unveiled an unexpected diversity of microalgae, including the haptophyte *Chrysochromulina* and the dinoflagellate genera *Pelagodinium* (Suessiales) and *Heterocapsa*, *Scrippsiella* and *Azadinium* (Peridiniales), closely associated with a single host species of Acantharia. The possibility that these microalgae correspond to prey cells ingested by *Acanthochiasma* prior to collection cannot be ruled out. These taxa are known to be an important dietary source for primary consumers and metazoan predators in marine ecosystems (Jeong et al. 2002; Kamiyama 2000; Lundgren and Granéli 2011). However, several lines of evidence argue in favor of the endosymbiotic nature of the microalgae identified. Firstly, light and epifluorescence microscopy observations revealed that the microalgae were located in the host endoplasm (Fig. 1), which in non-symbiotic acantharians is typically devoid of prey particles (Schewiakoff 1926). In addition, Acantharia have been observed to digest their prey extremely rapidly (within minutes) in the ectoplasm where enzyme-containing vacuoles are released (Schewiakoff 1926; our observations). With this in mind, we paid particular attention in our sampling strategy to maintaining the acantharian cells for 1-3 hours in 0.2-µm-filtered seawater before rinsing, microdissection and isolation. With this procedure, we were able to obtain microalgal sequences from holobiont cells and sequences of free-living cultures of symbiotic microalgae. Another line of evidence resides in the fact that the host specimens studied (12 in total) were sampled in consecutive years but exhibited common microalgal taxa, indicating a degree of specificity in the relationship (Table 1). Overall, this suggests that there is an intimate and long-lasting relationship between these different microalgae and the acantharian species *Acanthochiasma* sp.. Finally, some of the sequences obtained in this study were closely related to microalgae that are known to be symbionts of Foraminifera and other Radiolaria.

The dinoflagellate *P. béii* that we found associated with *Acanthochiasma* sp. is known to live in symbiosis worldwide with different species of planktonic Foraminifera (*Orbulina universa*, *Globigerinoides ruber*, *G. sacculifer* and *G. conglobatus*, Gast and Caron 2001; Shacked and de Vargas 2006). Likewise, members of the haptophyte order Prymnesiales have been found living as permanent endosymbionts not only in Radiolaria (*Spongodrymus* sp., Anderson et al. 1983; Gast and Caron 2001), but also in planktonic Foraminifera. More particularly, the *Chrysochromulina* species from clade B2 were described to occur within the Foraminifera species *Globigerinella siphonifera*, where they reside along the spines in the endoplasmic rhizopodial network that can be extended out of the shell during the day (Gast and Caron 2001; Faber et al. 1988).

To our knowledge, the Peridiniales dinoflagellate genera *Heterocapsa*, *Scrippsiella* and *Azadinium* have never previously been reported to be involved in symbiotic relationships. They are very likely not parasites of *Acanthochiasma* sp. since they do not belong to known groups of heterotrophic radiolarian parasites, such as the alveolates MALV I and MALV II (Bråt et al. 2012; Guillou et al. 2008). In addition, the proportion of hosts infected with these known parasites is generally about 1-10% (Siano et al. 2011), compared to the much higher prevalence

of these three genera found in *Acanthochiasma*. Overall, the weight of evidence indicates that the microalgal sequences retrieved correspond to photosymbionts hosted within the endoplasm of *Acanthochiasma* sp., but the mechanisms underlying the relationships remain to be elucidated.

Multiple microalgal partners

A combination of different techniques allowed us to identify microalgae of different sizes and shapes (Fig. 1) belonging to highly distant lineages, not only in a single acantharian species, but also within a single host cell (e.g. Vil 64, PEC 14, AC 24; Table 1). Occurrences of multiple symbiont taxa within a single host are considered to be rare because antagonistic interactions, such as competition for space and resources, can emerge between symbionts, and may threaten host fitness and the stability of the relationship (Douglas 1998; Franck 1996). Around 11% of cells from populations of planktonic Foraminifera have been found to simultaneously harbor different *P. béii* genotypes (Shaked and de Vargas 2006). Some benthic foraminiferal cells can occasionally live with two or three species of endosymbiotic diatoms (Lee and Coreira 2005). Corals, especially scleractinian species, and giant clams (*Tridacna* sp.) are commonly found with a heterogeneous assemblage of *Symbiodinium* sp., comprising up to four different genotypes (Baker and Romanski 2007; Carlos et al. 2000). Compared to these cases, *Acanthochiasma* is remarkable in that it associates with a much more diverse spectrum of photosynthetic partners, including several distantly related dinoflagellates as well as a haptophyte. In contrast, Acantharia from the more recently diverging clades E and F have established an exclusive symbiotic relationship with members of the haptophyte genus *Phaeocystis*, irrespective of host species and geographic location (Decelle et al. 2012a). Given the relatively distant and more basal position of *Acanthochiasma* sp. within acantharian phylogeny (clade B, Decelle et al. 2012b), we conclude that the multiple-partner photosymbiosis described here was established independently and before that of Acantharia from clades E and F, and that it represents a relatively primitive mode of symbiosis that is not specialized on a single photosynthetic partner. In this context, it is interesting to note that the skeleton organization is less complex in *Acanthochiasma* (simple long spicules), and that known extant diversity within the *Acanthochiasma* lineage is extremely low compared with that within the monophyletic lineage of Acantharia (clades E and F) living with *Phaeocystis*.

The two approaches adopted in this study, single-holobiont PCR and culturing, did not systematically recover the same microalgal sequences in all host cells (Table 1). This could be due to limitations in single-holobiont PCR inherent to the low number of intracellular microalgae and the variable abundance of symbiont taxa within each host. In addition, the success rate for phytoplankton cultures strongly depends on the taxa isolated and the choice of medium used. For instance, *Chrysochromulina* species from clade B2 are known to be difficult to grow in culture (Liu et al. 2009), and this may explain why we only obtained sequences of this taxon from holobionts. Nevertheless, combining the two approaches appears to be a valuable strategy for identification of photosymbionts from uncultivable unicellular hosts.

Insights into planktonic photosymbiosis

In coastal benthic ecosystems, the dinoflagellate *Symbiodinium* (Suessiales) is a common photosymbiont of marine invertebrates and protists, such as sponges, cnidarians, mollusks and benthic Foraminifera (Pochon et al. 2006). Here, we provide evidence that the genus *Pelagodinium*, a close relative of *Symbiodinium*, is also a generic symbiont in different taxa of planktonic protists (Acantharia and Foraminifera). Like *Symbiodinium*, this genus represents a broad phylogenetic entity, which likely contains several undescribed sub-clades and cryptic species (Fig. 4). A recent study reported that other members of the Suessiales were found in symbiosis with common freshwater sponges in the oligotrophic Lake Baikal (Annenkova et al. 2011). The Suessiales evidently hold a prevalent role in photosymbiosis in aquatic ecosystems and seem to be particularly well suited for diverse hosts, putatively possessing a fundamental symbiotic competence. However, photosymbiotic relationships between other genera of the Suessiales (e.g. *Biecheleria*, *Biecheleriopsis*, *Protodinium*, *Polarella*) and invertebrates and/or protists still need to be demonstrated.

Current knowledge suggests that free-living *Symbiodinium* are specific to shallow coastal waters, which may explain why they have not been found to be involved in planktonic photosymbioses. In contrast, with the exception of *Pelagodinium* for which little information is available, the other microalgae found closely associated with Acantharia (i.e. *Phaeocystis*, *Heterocapsa*, *Azadinium*, *Scrippsiella*, *Chrysochromulina)* all have extensive and ubiquitous populations thriving in coastal and oceanic waters (Gottschling et al. 2005; Litaker et al. 2002; Liu et al. 2009; McDonald et al. 2007; Tillman et al. 2009). Some of these microalgae even form blooms that are occasionally harmful to the trophic chain and detrimental for human activities (e.g. *Heterocapsa*, *Azadinium*, *Chrysochromulina*, *Phaeocystis*; Edvardsen and Imai 2006). Likewise, the cyanobacteria *Synechococcus* from clade II, found in symbiotic association with other Radiolaria (Yuasa et al. 2012), are also very abundant in their free-living phase worldwide, particularly in coastal/continental shelf areas from low latitudes (Zwirglmaier et al. 2008). Evidence is therefore mounting that the abundance of the free-living phase of microalgae in pelagic ecosystems, and hence their availability for hosts that must acquire new symbionts during each generation, is an important determining factor for involvement in a symbiotic relationship. In this context, it is difficult to see a net ecological benefit for the microalgae involved in asymmetric associations with Acantharia. Together with assessment of the physiological roles of each symbiont, investigation of the biogeography and temporal variation of this peculiar multiple-partner relationship would help to better understand potential ecological and evolutionary impacts for the host and symbiont partners.

Acknowledgments

We thank the Laboratoire d'Océanographie de Villefranche-sur-Mer (LOV) for providing sampling facilities and Morgan Perennou and Gwen Tanguy from the GENOMER platform at the Station Biologique de Roscoff. This project was supported by the network "Bibliothèque du Vivant" funded by the CNRS, the Muséum National d'Histoire Naturelle, the INRA and the CEA (Centre National de Séquençage). This research was supported by the Region Bretagne (DIPHOPE 044763), the JST-CNRS program, and the EU FP7 I3 project ASSEMBLE and KBBE project MACUMBA.

References

Anderson OR (1983) Radiolaria. Springer-Verlag, New York

Annenkova NV, Lavrov DV, Belikov SI (2011) Dinoflagellates associated with freshwater sponges from the ancient Lake Baikal. Protist 162:22-236

Baker AC, Romanski AM (2007) Multiple symbiotic partnerships are common in scleractinian corals, but not in octocorals: Comment on Goulet (2006). Mar Ecol Prog Ser 335:237-242

Banaszak AT, Iglesias-Prieto R, Trench RK (1993) *Scrippsiella velellae* sp. nov. (Peridiniales) and *Gloeodinium viscum* sp. nov (Phytodiniales), dinoflagellate symbionts of hydrozoans (Cnidaria). J Phycol 29:517-528

Brandt K (1881) Ueber das Zusammenleben von Thieren und Algen. Verh Physiol Ges Berlin 1881(1882):22-26

Bråt J et al (2012) Radiolaria Associated with Large Diversity of Marine Alveolates. Protist 163:767-777

Carlos AA, Baillie BK, Maruyama T (2000) Diversity of dinoflagellate symbionts (Zooxanthellae) in a host individual. Mar Ecol Prog Ser 195:93-100

Cienkowski L (1871) Über Schwarmer-Bildung bei Radiolarien. Archiv für Mikroskopische Anatomie 7:372-381

Coleman AW (2007) Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. Nucleic Acids Res 35:3322-3329

Decelle J et al (2012a) An original mode of symbiosis in open ocean plankton. PNAS. Early edition 15 oct.

Decelle J, Suzuki N, Mahé F, de Vargas C, Not F (2012b) Molecular phylogeny and morphological evolution of the Acantharia (Radiolaria) Protist 163:435-450

de Bary, AH (1878) Vortrag: Uber Symbiose. Tagblatt der 51. Versammlung Deutscher Naturforscher und Aerzte in Cassel. Baier & Lewalter, Kassel, pp. 121–126

Douglas AE (1998) Evolution of mutualistic symbiosis without vertical transmission in symbiosis. Heredity 81:599-603

Douglas AE, Smith DC (1989) Are endosymbioses mutualistic? TREE 4 :350-352

Edvardsen B et al. (2011) Ribosomal DNA phylogenies and a morphological revision provide the basis for a revised taxonomy of the Prymnesiales (Haptophyta). Eur J Phycol 46:202-228

Edvardsen B, Imai I (2006) The ecology of harmful flagellates within Prymnesiophyceae and Raphidophyceae. In Ecology of Harmful Algae, Ecological Studies, Vol. 189 (Granéli, E. & Turner, J.T., editors), 67-79. Springer, Berlin

Faber WW Jr, Anderson OR, Lindsey JL, Caron DA (1988) Algal-foraminiferal symbiosis in the planktonic foraminifer *Globigerinella aequilateralis*: I. Occurrence and stability of two mutually exclusive chrysophyte endosymbionts and their ultrastructure. J Foram Res 18:334-343

Febvre J, Febvre-Chevalier C (1979) Ultrastructural study of zooxanthellae of three species of Acantharia (Protozoa: Actinopoda), with details of their taxonomic position in the prymnesiales (prymnesiophyceae, Hibberd, 1976). J Mar Biol Ass UK 59:215-226

Foster RA, Collier JL, Carpenter EJ (2006a) Reverse transcription PCR amplification of cyanobacterial symbiont 16S rRNA sequences from single non-photosynthetic eukaryotic marine planktonic host cells. J Phycol 42:243-250

Foster RA, Carpenter EJ, Bergman B (2006b) Unicellular cyanobionts in open ocean dinoflagellates, radiolarians, and tintinnids: ultrastructural characterization and immuno-localization of phycoerythrin and nitrogenase. J Phycol 42:453-463

Franck SA (1996) Host symbiont conflict over the mixing of symbiotic lineages. Proc R Soc Lond B 263:339-344

Gast RJ, Caron DA (2001) Photosymbiotic associations in planktonic foraminifera and radiolaria. Hydrobiologia 461:1-7

Gast RJ, McDonnell TA, Caron DA (2000) srDNA-based taxonomic affinities of algal symbionts from a planktonic foraminifer and a solitary radiolarian. J.Phycol. 36:172-177

Gast RJ, Caron DA (1996) Molecular phylogeny of symbiotic dinoflagellates from planktonic foraminifera and radiolaria. Mol Biol Evol 13:1192-1197

Gottschling M et al (2005) A molecular phylogeny of *Scrippsiella* sensu lato (Calciodinellaceae, Dinophyta) with interpretations on morphology and distribution. Eur J Phycol 40:207-220

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27(2):221-224

Guillou L et al (2008) Widespread occurrence and genetic diversity of marine parasitoids belonging to syndiniales (Alveolata). Env Microbiol 10:3349-3365

Haeckel E (1888) Die Radiolarien (Rhizopoda Radiolaria). Eine Monographie, Dritter Teil, Die Acantharien oder Actinopyleen-Radiolarien. Verlag von Georg Reimer, Berlin, pp 33

Hollande A, Carré D (1974) Les xanthelles des radiolaires sphaerocollides, des acanthaires et de *Velella velella*: infrastructure – cytochimie – taxonomie. Protistologica 4:573-601

Jeong HJ, Yoon JY, Kim JS, Yoo1 YD, Seong KA (2002) Growth and grazing rates of the prostomatid ciliate *Tiarina fusus* on red-tide and toxic algae. Aquat Microb Ecol 28:289-297

Kamiyama T (2000) Application of a vital staining method to measure feeding rates of field ciliate assemblages on a harmful alga. Mar Ecol Prog Ser 197:299-303

Keller MD, Selvin RC, Claus W, Guillard RRL (1987) Media for the culture of oceanic ultraphytoplankton. J Phycol 23:633-638

LaJeunesse TC (2001) Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus *Symbiodinium* using the ITS region: in search of a "species" level marker. J Phycol 37:866-880

Lee JJ, Correia M (2005) Endosymbiotic diatoms from previously unsampled habitats. Symbiosis 38:251–260

Litaker RW et al (2002) Seasonal niche strategy of the bloom-forming dinoflagellate *Heterocapsa triquetra*. Mar Ecol Prog Ser 232:45-62

Liu H et al (2009) Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. PNAS 106:12803–12808

Lundgren V, Granéli E (2011) Influence of altered light conditions and grazers on *Scrippsiella trochoidea* (Dinophyceae) cyst formation. Aquat Microb Ecol 63:231-243

McDonald SM, Sarno D, Scanlan DJ, Zingone, A (2007) Genetic diversity of eukaryotic ultraphytoplankton in the Gulf of Naples during an annual cycle. Aquat Microb Ecol 50:75-89

Michaels AF (1988) Vertical distribution and abundance of Acantharia and their symbionts. Marine Biology 97:559–569

Michaels AF, Caron DA, Swanberg NR, Howse FA, Michaels CM (1995) Planktonic sarcodines (acantharia, radiolaria, foraminifera) in surface waters near Bermuda - abundance, biomass and vertical flux. J Plankton Res 17:131–163

Muscatine L, Falkowski PG, Porter JW, Dubinsky Z (1984) Fate of photosynthetic fixed carbon in light- and shade- adapted colonies of the symbiotic coral *Stylophora pistillata*. Proc R Soc Lond Ser B 222:181-202

Pawlowski et al (2012) CBOL Protist Working Group: Barcoding eukaryotic richness beyond the animal, plant and fungal kingdoms. PLos Biology. In press

Pochon X, Montoya-Burgos JI, Stadelmann B, Pawlowski J (2006) Molecular phylogeny, evolutionary rates, and divergence timing of the symbiotic dinoflagellate genus *Symbiodinium*. Mol Phylogenet Evol 38:20-30

Saldarriaga JF, Taylor FJR, Cavalier-Smith T, Menden-Deuer S, Keeling PJ (2004) Molecular data and the evolutionary history of dinoflagellates. Eur J Protist 40:85-111

Schewiakoff WT (1926) The Acantharia. Fauna e Flora del Golfo di Napoli 37:1-755

Shaked Y, de Vargas C (2006) Pelagic photosymbiosis: rDNA assessment of diversity and evolution of dinoflagellate symbionts and planktonic foraminiferal hosts. Mar Ecol Prog Ser 325:59-71

Siano R et al (2011) Distribution and host diversity of Amoebophryidae parasites across oligotrophic waters of the Mediterranean Sea. Biogeosciences 8:267-268

Siano R, Montresor M, Probert I, Not F, de Vargas C (2010) *Pelagodinium* gen. nov and *P-béii* comb. nov., a dinoflagellate symbiont of planktonic Foraminifera. Protist 161**:**385-399

Spero HJ (1987) Symbiosis in the planktonic foraminifer, *Orbulina universa*, and the isolation of its symbiotic dinoflagellate, *Gymnodinium béii* sp. nov. J Phycol 23:307-317

Stern et al (2012) The ribosomal internal transcribed spacer (ITS) as a candidate dinoflagellate barcode marker. PLoS ONE 7(8): e42780. doi:10.1371/journal.pone.0042780

Stoecker DK, Gustafson DE, Verity PG (1996) Micro- and mesoprotozooplankton at 140°W in the equatorial Pacific: heterotrophs and micotrophs. Aquat Microb Ecol 10:273-282

Swanberg NR, Caron DA (1991) Patterns of feeding in epipelagic oceanic plankton. J Plankton Res 13:287-312

Tamura K et al (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 28:2731-2739

Tillman U, Elbrächter M, Krock B, John U, Cembella A (2009) *Azadinium spinosum* gen. et sp. nov. (Dinophyceae) identified as a primary producer of azaspiracid toxins. Eur J Phycol 44:63-79

Wooldridge SA (2010) Is the coral-algae symbiosis really “mutually beneficial” for the partners? Bioessays 32:615-625

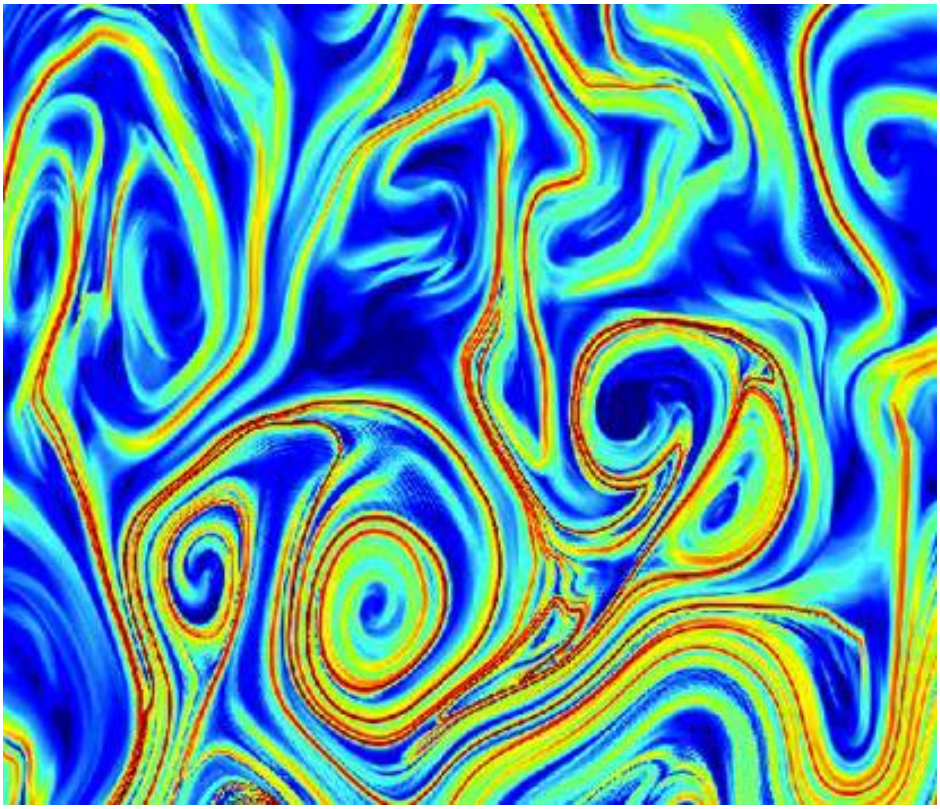
Yellowlees D, Rees TAV, Leggat W (2008) Metabolic interactions between algal symbionts and invertebrates hosts. Plant Cell Environ 31:679-694

Yuasa T, Horiguchi T, Mayama S, Matsuoka A, Takahashi O (2012) Ultrastructural and molecular characterization of cyanobacterial symbionts in *Dictyocoryne profunda* (polycystine radiolaria). Symbiosis DOI 10.1007/s13199-012-0174-2

Zwirglmaier K et al (2008) Global phylogeography of marine S*ynechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. Env Microbiol 10(1):147-161

Chapitre 3

Biogeography and abundance of the Acantharia in the world's oceans



A 1000-km-wide patch of the ocean with rapid currents colored in red (C. Harrison)

Worldwide distribution of Acantharia in the photic zone

Johan Decelle, Fréderic Mahé, Samuel Chaffron, Fabrice Not, Colomban de Vargas, Sarah Romac

In PREPARATION

Introduction

The open ocean is a vast and highly oligotrophic environment characterized by a low biomass but with a considerable impact in global biogeochemical cycles. Ecological interactions are of major importance to fully understand the evolution of organisms but also the functioning of the ecosystems (Thompson, 2005; Strom, 2008; Margulis and Fester, 1991). Yet, symbiotic consortia between planktonic microorganisms remain poorly characterized. A partnership with a photosynthesizing microalga (photosymbiosis) and a nitrogen-fixing cyanobacterium have been shown to be prevalent strategies in unicellular hosts in order to thrive in the remote oceanic regions (Foster *et al.*, 2011; Stoecker *et al.*, 2009; Thomspon *et al.*, 2012; Caron, 2000).

Among planktonic protists, Acantharia (Radiolaria) that build a mineral skeleton in strontium sulfate (celestite) hold a key position in the community as they typically outnumber the other large skeleton-building protists Foraminifera and Polycystinea in oligotrophic oceans (Michaels, 1988). Although Acantharia seem ubiquitous in the oceans, the complete dissolution of their celestite skeleton in classical fixatives used in plankton studies led to an incomplete knowledge of their geographic distribution, as well as, the ecology of the different taxonomic groups. Because of a highly efficient phagotrophy, these uncultured protists feed on a wide diversity of prey, from bacteria to small metazoans (Caron and Swanberg, 1991), but some acantharian species are also able to establish a photosymbiosis with dinoflagellates and haptophytes (Febvre and Febvre-Chevalier, 1979; Decelle *et al.*, 2012a, chapter II). In the upper layer of oligotrophic waters, symbiotic Acantharia were found to contribute up to 80% of acantharian biomass (Stoecker, 1996), and can occasionally account for up to 20% of the total primary production (Michaels, 1988). These acantharians host 10 to 100 cells of the haptophyte *Phaeocystis*, an abundant and widely distributed microalga in its free-living phase, which is considered as a key species in marine pelagic ecosystems (Decelle et *al.*, 2012b (chapter I);

Schoemann *et al.*, 2005; Verity and Smetacek, 1996). This photosymbiotic relationship has been observed in different oceanic regions worldwide and involves consistently the genus *Phaeocystis*. Yet, this association appears flexible according to the geography, whereby the Acantharia live with the local *Phaeocystis* species (Decelle *et al.*, 2012a, chapter II). Although the exact nature of the relationship remains to be demonstrated, weight of evidence suggests that this photosymbiosis would not represent a fully mutually beneficial partnership. Acantharia seem to acquire phototrophy by irreversibly sequestering and transforming the morphology of *Phaeocystis* cells. Examination of the biogeography and abundance of this peculiar Acantharia-*Phaeocystis* association is required to better understand its modes and functioning in the oceans.

Explanations for the distribution of marine microbial taxa mainly focused on abiotic variables, such as the temperature or nutrient concentrations, but more rarely on biotic factors. The absence or loss of compatible mutualistic partners may constrain the dispersal of organisms (Richardson *et al.*, 2000; Pringle *et al.*, 2009). Like for most photosymbiotic hosts, Acantharia have to obligatory acquire their microalgal symbionts from the environment at each generation (horizontal transmission). This recurrent process appears to be very challenging in the microbia-diluted open ocean, more particularly for unicellular hosts that have short generation times (3-4 weeks). Effective establishment of the photosymbiosis therefore requires that a free-living pool of *Phaeocystis* is available in the surrounding environment at the right time. In this context, this microalga is likely one of the most important biotic parameter that shape the biogeography and ecology of the symbiotic Acantharia. Similarly, if this photosymbiosis turns out to be also mandatory for the symbiont, *Phaeocystis* distribution might also be controlled by biogeography of its host.

The free-living phase of terrestrial and marine symbiont populations is generally unknown (Douglas, 1996, 1998), mainly because most of them are elusive in the environment and can be members of the rare biosphere (Pedrós Alió, 2012; Nyholm and McFall-Ngai, 2004). Yet, the ecology of the symbiont outside its host is essential to understand the mechanisms by which the partners associate in the environment and maintain their intimate interaction through time. Our current knowledge of the distribution of *Phaeocystis* is relatively well-known but mostly relies on microscopic observations and culture isolates from coastal nutrient-rich regions where this haptophyte can form extensive blooms (Schoemann *et al.*, 2005). In contrast, the ecology of *Phaeocystis* is less understood in the open ocean. Culture-independent DNA-based molecular surveys have found only few sequences related to *Phaeocystis* in the South Pacific (Shi *et al.*, 2009; Moon-van der Staay *et al.*, 2001) and Indian Ocean (Liu *et al.*, 2009). These sparse studies provide a limited understanding of the dispersal capacity of *Phaeocystis* in oceanic waters, hence giving only a partial picture of the free-living reservoir of symbiont for Acantharia.

A recent molecular phylogeny of Acantharia described at least nine molecular clades (I, III and A to F; Figure 1) (Decelle *et al.*, 2012b, chapter I). The photosymbiosis with *Phaeocystis* specifically occurs in the most recently diverging clades E and F, which form a monophyletic lineage presenting a high morphological disparity between taxa. The fact that other clades (except B2) do not live with endosymbiotic microalgae allows comparing the distribution and

abundance of the symbiotic and non-symbiotic Acantharia, hence highlighting the putative ecological importance of this photosymbiosis in the ocean. Nowadays, the advent of the environmental metabarcoding offers a unique opportunity to have a comprehensive and rapid view of the community composition at very large scales of uncultivable microorganisms (Soggin *et al.*, 2006). The wealth of short ribosomal sequences (e.g. V4 or V9 region of the ribosomal 18S rDNA gene) can nevertheless be reliably interpreted through reference databases with a broad taxonomic coverage. Such global environmental survey can be now envisioned for Acantharia due to a comprehensive morpho-genetic reference database of single cells.

Here, we explored the community structure of the Acantharia in the photic zone of the worldwide oceans through a metabarcoding analysis. We particularly focused on the taxa in symbiosis with *Phaeocystis* (clades E and F). To better understand the ecology and the functioning of this photosymbiotic interaction, the spatial distribution of *Phaeocystis* was examined at the same sampling sites, and putative correlations were investigated with different oceanographic parameters.

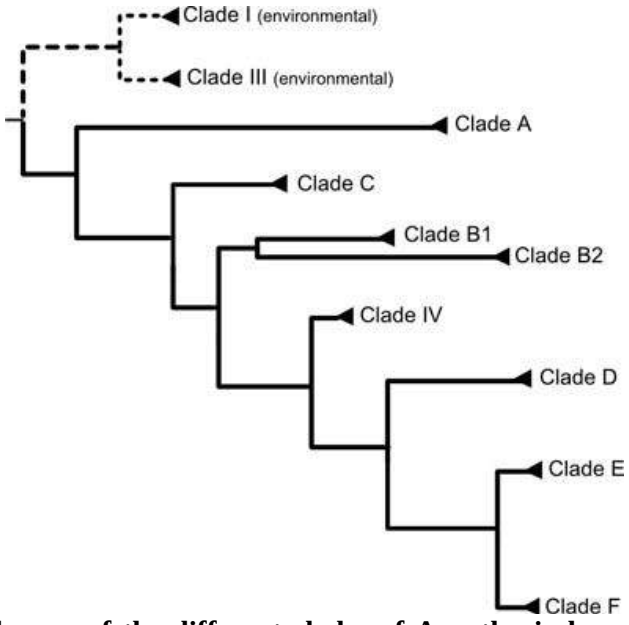


Figure 1 Schematic phylogeny of the different clades of Acantharia based on Decelle et al 2012b (chapter I). This study particularly focuses on clades E and F, representing the Acantharia living in photosymbiosis with the microalga *Phaeocystis*

Results

Oceanographic context of the 35 stations

The 35 stations sampled in this study represent 7 oceanic regions and 12 Longhurst provinces from 42° N to 62° S of latitude (Table S1). Clustering analyses based on Euclidian distance were performed with oceanographic parameters (Table S2) to compare the environmental similarities between stations (Figure 2). The southernmost stations 82, 84 and 85 are highly distant from all the other stations. These stations exhibit the most important day length (15-19 hours), the coldest waters (0-7 °C) and high concentrations of silicates, nitrates and phosphates (Table S2). Another cluster composed of Atlantic Ocean and South Pacific Gyre stations (4, 72, 76, 78, and 98) is characterized by high oligotrophic settings whereas stations 7 (Mediterranean Sea), 36 and 65 (Indian Ocean), 66, 67 and 70 (Atlantic Ocean) and 102 (Pacific Ocean) represent nutrient-rich coastal waters. Other stations, mostly located between tropical and subtropical latitudes, form a large cluster with similar physico-chemical conditions.

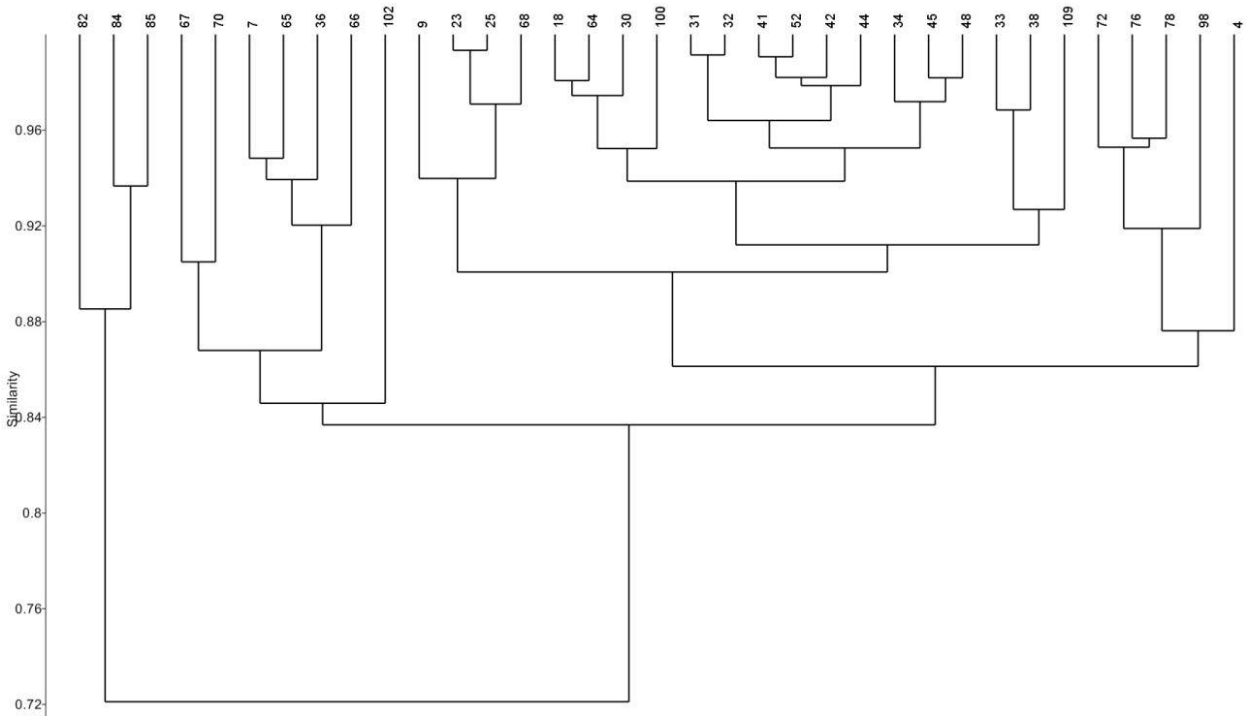


Figure 2: Cluster analysis of the 35 stations sampled in this study based on the Euclidian distances of the environmental parameters (Tables S1 and S2).

General data on Acantharia and their contribution to the protistan community

In the photic zone, a total of 1,288,018 reads (representing 74,270 unique reads) of the variable region V9 of the 18S rDNA gene were assigned to Acantharia. At the surface and Deep Chlorophyll Maximum (DCM), 818,108 and 869,907 reads were obtained, respectively. The

number of acantharian reads obtained in each station varied from 169 (station 67) to 247,887 (station 109), with an average of 46,206 reads (Table S3). For most acantharian clades, the majority of reads were assigned with a high similarity to reference sequences (982,643 and 947,277 assigned at ≥ 95 % and ≥ 97%, respectively). Yet, assignation level appeared lower in clades B1, I, and III, varying from 85 to 95% (Figure S1).

To compare the read numbers between Acantharia and the whole protistan community, we consider hereafter protist reads as all but reads assigned to metazoans. Total acantharian reads in the world photic zone (surface and DCM) contributed on average 0.95 % to the total protistan reads: 1.14% and 0.78% in the 0.8-5 µm and 20-2000 µm size fractions, respectively (Figure S2). Notably, acantharian reads accounted for between 3 and 9 % at stations 4, 7, 30, 32, 34, 38 and 84. One can notice that at several stations reads were significantly more numerous in the small size fraction (e.g. stations 34, 42, 52, 64, 65 etc). Read abundances from the two size fractions are not correlated between each other and across stations (Figure S2). For clades A, B2, C, D, E, the average number of reads in the two size fractions were not statistically different (t-test, $p > 0.05$) but for clades I, B1 and III reads were more numerous in the small size fraction while for clade F, it was in the big size fractions ($p < 0.05$).

Saturation and rarefaction curves for each acantharian clades will be performed with the program Mothur (Schloss et al., 2009),as well as, diversity indices (e.g. Chao1, ACE).

Biogeography of Acantharia: significance of clades E and F in the photic zone

All of the 10 acantharian clades were present at every station but in different relative contributions (Figure 3). Reads of clade F contributed on average 45% to the total acantharian reads in the photic zone and up to 80 % in high (station 84), subtropical, and tropical latitudes (stations 4, 38, 45, 48, 70). Clade C was the second most represented acantharian clade in terms of read number, encompassing up to 50 % of the reads at stations 7, 32 and 67. Clade D represented between 3-10 % at most locations, but appeared to reach 30 % (stations 23) and 61 % (station 25) in the Mediterranean Sea. Finally, clades B2 and E contributed on average to 2.5% and 8.6% in the photic zone, respectively.

Diversity of Acantharia, assessed by the Shannon index (Figure 3), did not show any marked pattern along a latitudinal gradient or according to the distance from coastal areas. While stations 31, 33, 34, 52, 65, 72 and 76 exhibited a broad diversity (Shannon index > 1.8), lower diversity was found at stations 45, 48, 70 and 84 (Shannon index < 1). A clustering based on the Bray-Curtis dissimilarity index with 1000 permutations showed that the Acantharia communities in stations 25, 67, 82 and 85 are distinct each other and significantly different from other stations (Figure S3). This could be influenced by the lower relative contribution of reads from clades E and F at most of these stations. Assemblages from other stations group together with no supported sub-clusters.

Phylogenetic diversity (Faith) of the Acantharia community will be calculated for each station with MOTHUR or the Phylogenetic diversity Analyser program (Version 0.5.1).

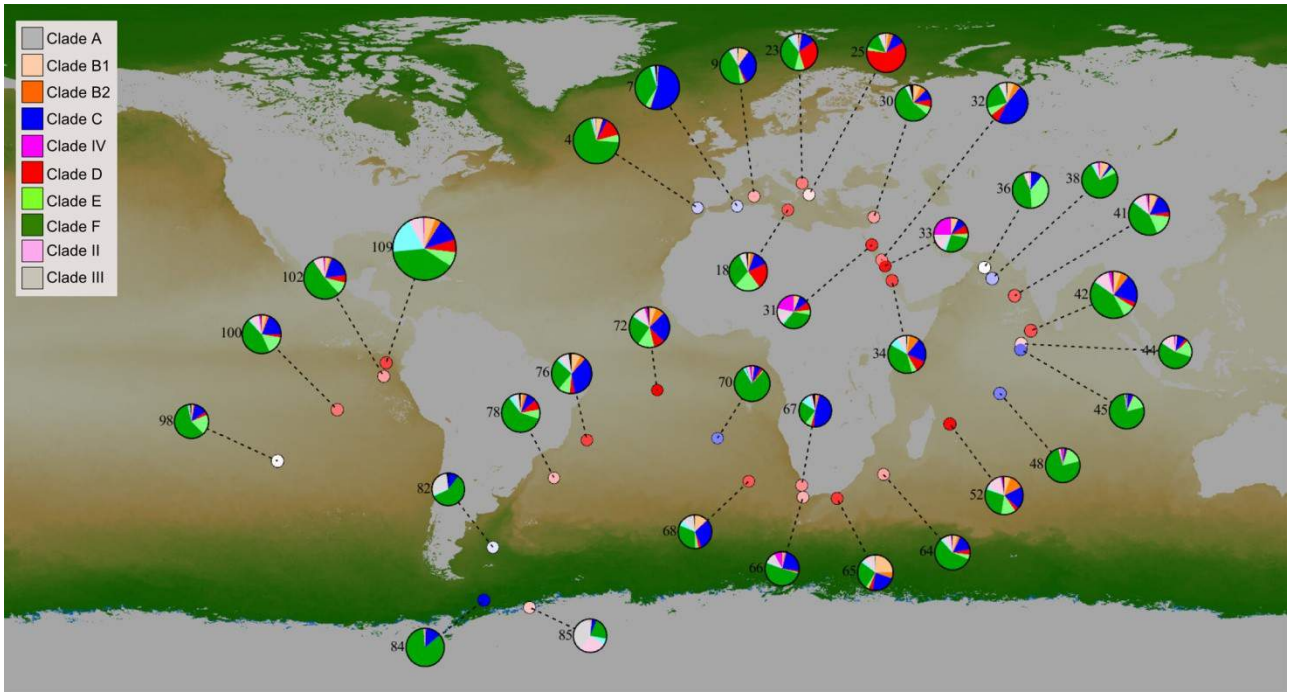


Figure 3: Biogeographic distribution of acantharian clades at different stations. The color of the dots indicating the sampling locations changes according to the Shannon index along a gradient, from blue (low diversity < 1) to red (high diversity > 1). The size of the pie-charts scales with the total number of acantharian sequences acquired at the given station (from 69 at station 67 to 247,887 at station 109).

Beta diversity of symbiotic Acantharia

We considered hereafter the principal reads of each clade (> 85% assignation with a reference sequence and present in > 80 copies). A plot presenting the abundance of reads versus the number of stations where these reads were seen (occupancy) was built (Figures 4). The presence of one read in a station was taken into account when it represented at least 1% of all the acantharian reads at this given station. The reads assigned to symbiotic clades E and F showed a higher dispersal compared to the ones from other clades that had a more restricted distribution (Figures 4A and 4B). Many reads of symbiotic Acantharia are present in more than 12 stations whereas the ones of non-symbiotic clades are typically contained below 10 stations out of the 35 sampled in this study. Given the abundance-occupancy positive relationship generally recognized in macroecology (Gaston *et al.*, 2000; Verberk *et al.*, 2010), the wider dispersal of clades E and F can be explained by their high abundance seen previously (Figure 3). Yet, Pearson's correlations (r) that describe the abundance-occupancy relationships for each clade indicated no or even negative relationship for clades E (r = 0.004) and F (r = -0.27), whereas they tended to be positive in other clades varying from 0.14 to 0.93 (B1: 0.93; B2: 0.49; C: 0.14; D:

0.62; I: 0.54; A: 0.60 and III: 0.16). The absence of positive abundance-occupancy relationship in clades E and F could be explained by the reads that were in low copy numbers in a given station but widely distributed over stations (bottom right zone of the graph; Figure 4A).

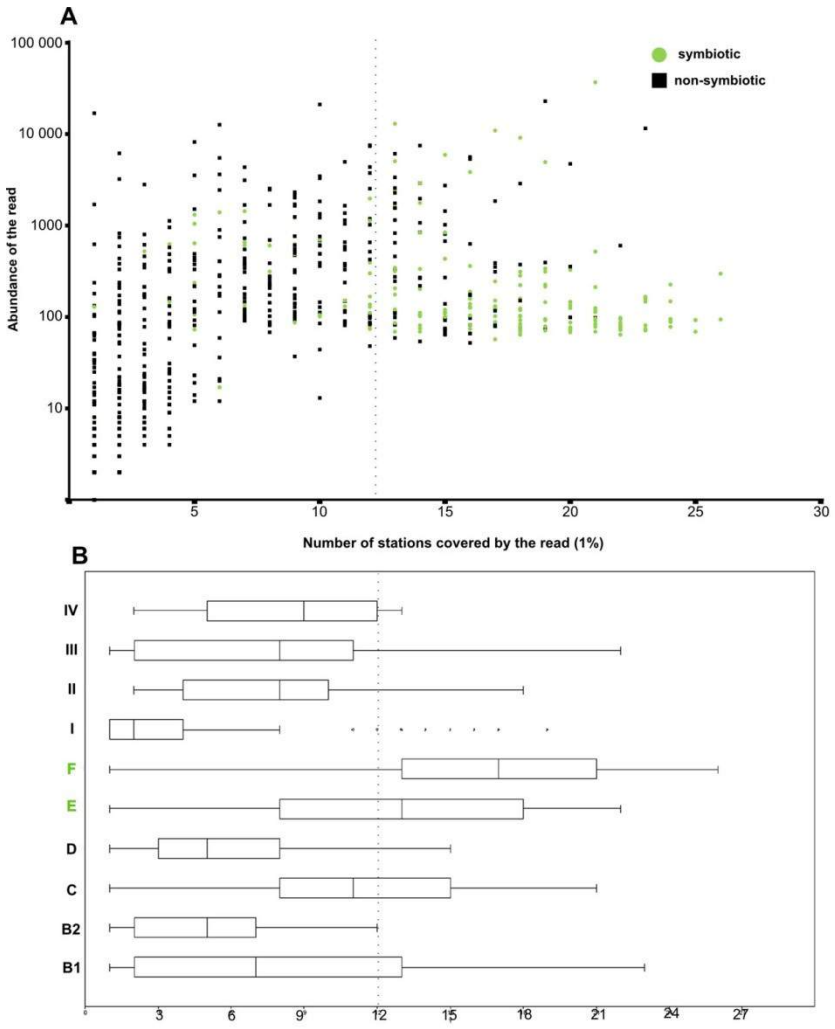


Figure 4: Dispersal of principal reads (>85% of similarity, > 80 copies) belonging to different clades of symbiotic (clade E and F) and non-symbiotic Acantharia. A) A read represented by a dot is plotted according to its abundance (copy number) versus the number of stations where it has been detected (occupancy); B) A Box plot shows the median of the occupancy for each clade.

To examine whether the phylogenetic diversity of clades E and F varies according to the size fractions, depths or oceanic regions (*beta* diversity), analyses were carried out with the Unifrac distance metric. The same reads considered previously were used (assignation > 85% and abundance > 80 copies). Principal coordinates analysis (PCoA) of pairwise unweighted Unifrac values showed no clear partitioning for clades E and F according to the size fractions, sampling depth or oceanic regions (Figure 5): the phylogenetic diversity of these two clades at the V9 resolution was homogeneous whatever these conditions. Nevertheless, we can note that the

phylogenetic diversity of clade F seemed different between the Antarctic and the Mediterranean Sea (red and green points, respectively; Figure 5, right).

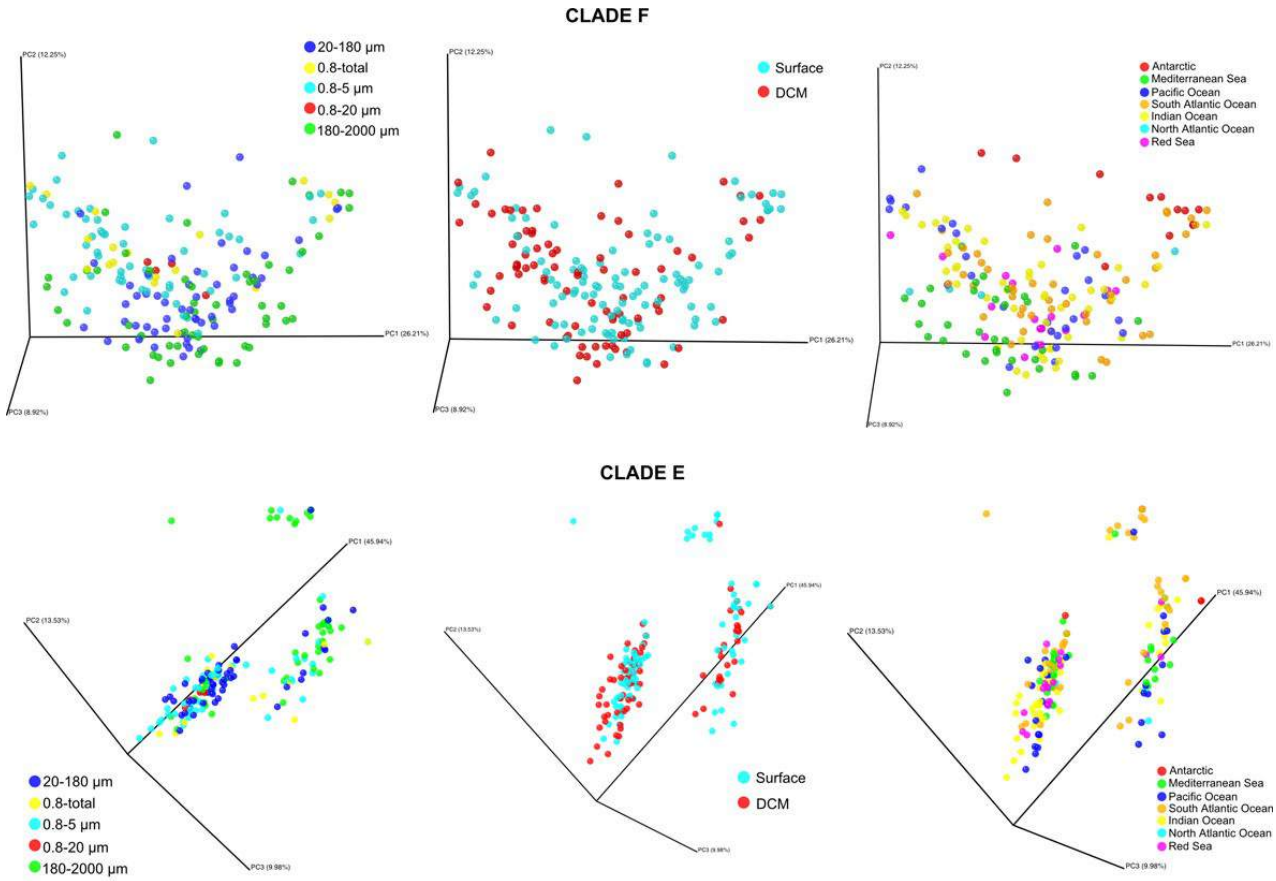


Figure 5: Principal coordinates analysis (PCoA) of unweighted pairwise UniFrac distances of the phylogenetic diversity of clades F (top) and E (bottom) in different size fractions, depths and oceanic regions.

Distribution of *Phaeocytis* reads

We then looked at the distribution and the relative abundance of the haptophyte *Phaeocystis* to characterize the free-living environmental pool of symbionts of clades E and F. Reads of *Phaeocystis* were found in the photic zone of all of the 35 stations sampled, with peaks of abundance at high latitudes in coastal stations 67, 68, 82 (South Atlantic) and 84, 85 (Antarctic) (Figure 6). These higher relative abundances (normalized by the total number of reads in the sample) corresponded to an increased day length (higher photoperiod), and important nutrient concentrations, such as silicates and nitrates (Table S2). For instance, at stations 84 and 85 in Antarctic during the summer, the photoperiod lasted 19 hours, which is the highest value in the dataset (Figure 6). The lowest abundances of *Phaeocystis* reads were reported for stations 34, and 45 and 48 in the Red Sea and Indian Ocean, respectively.

No clear correlation was highlighted between the prevalence of symbiotic Acantharia and the relative abundance of *Phaeocystis*, or the photoperiod and chlorophyll concentration at the DCM (Figure 6). For instance, in station 48, the read number of *Phaeocystis* is very low whereas the prevalence of symbiotic Acantharia is high among the community. Yet, we can note that in station 84, both partners *Phaeocystis* and symbiotic Acantharia display a high prevalence. Symbiotic Acantharia were not necessarily dominant in severe oligotrophic settings, such as in the eastern basin of the Mediterranean Sea (stations 18 to 30), and turned to be also a major component of the acantharian community in more productive waters (station 82).

Abiotic and biotic correlations

We carried out co-occurences analyses to detect more carefully abiotic and biotic correlations that putatively shape the distribution and abundance of clades E and F. Analyses were performed twice, with and without the stations 82, 84 and 85 that were clearly different from the others (Figure 2). Without antarctic stations, co-occurrences were found to be statistically significant between clades E and F, but not between these clades and *Phaeocystis* (Figure S4A). Including the antarctic stations, clades E and F correlated each other again, but also with other clades. As for *Phaeocystis*, only one correlation was detected with the silicates.

More co-occurrence analyses will be performed with all the oceanographic parameters

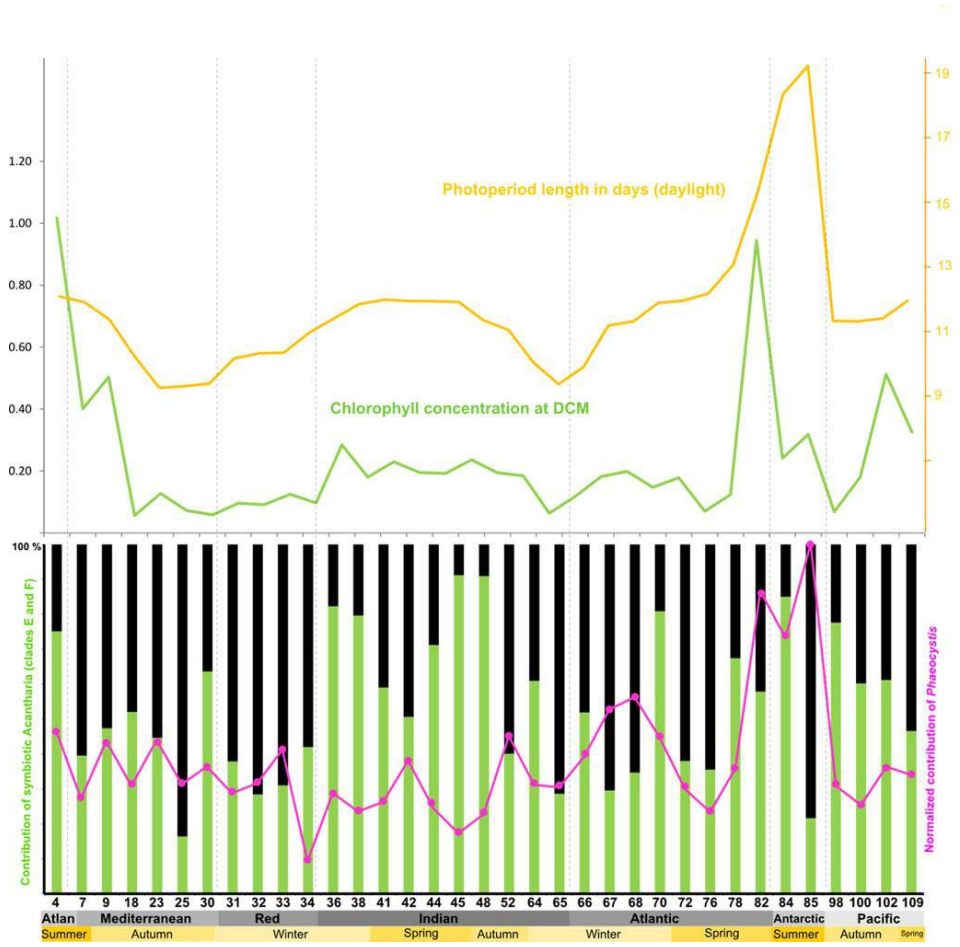


Figure 6: Putative biotic (abundance of *Phaeocystis*) and abiotic parameters (chlorophyll concentration and photoperiod) that shape the contribution of symbiotic Acantharia (clade E and F; green bars) in the photic zone. The curve in pink shows the normalized contribution of *Phaeocystis* reads to the total reads sampled at a given station

Discussion

A morpho-genetic reference database of the uncultured protist Acantharia allowed us to explore its worldwide distribution in the oceans and get insights into the ecology and functioning of the photosymbiosis with *Phaeocystis*.

Contribution of Acantharia among protists and occurence in the small size fraction

Acantharian reads appeared to significantly contribute to protistan reads in different oceanic regions (up to 9%), taking into account the high number of protistan taxa in marine ecosystems. Although the read number only gives a relative abundance, this study confirms previous observations on freshly collected zooplankton and environmental molecular surveys. Acantharia can contribute up to 40% of the zooplankton biomass (Zas'ko and Vedernikov 2003)

and are generally more abundant than their rhizarian counterparts, Foraminifera and Polycystinea (Michaels *et al.,* 1995). In some environmental clone libraries, up to 20% of the sequences can belong to Acantharia (Countway *et al.,* 2007; Lovejoy 2011). We cannot rule out that the multiple nuclei found in Acantharia might amplify their relative contribution in the microbial community (Suzuki *et al.,* 2009), assuming that the number of rDNA copies scales with the number of nuclei. But this bias is mitigated if we focus specifically on the Acantharia, since all the species are polynucleated (except *Haliomatidium* sp.; Febvre *et al.,* 2000).

Together with other studies (Not *et al.,* 2007; Quaiser *et al.,* 2010; Lovejoy *et al.,* 2006, 2011), this survey found acantharian sequences in great numbers in the small size fraction (0.8-5 µm). According to the phylogenetic analyses in these studies and what we observed for clades E and F (Figure 5), we argue that these particular sequences do not belong to new lineages of Acantharia that would have a small cell size. At least three reasons could explain their occurrence in that fraction. Dissolved DNA from acantharian origin floating in the water column can be the first explanation (Paul *et al.,* 1990). Second, these very fragile protists can be easily damaged by the sampling procedure (net towing, pumping and filtration steps). Detritus can therefore end up in the smallest size fraction and be detected by DNA amplification. Note that even a gentle sampling with niskin bottles recovers acantharian sequences in the small size fraction (Lovejoy *et al.,* 2011). Finally, these reads can also correspond to the thousands of 2-3-µm-size "swarmer" cells that can be released in the environment from a single adult specimen (Caron and Swanberg 1990). This could represent an important source of DNA in this type of study since it is a common reproductive stage in large pelagic rhizarian cells, such as Radiolaria and Foraminifera (Bé 1982; Hollande 1965). Reads of clades I, III and B1 are statistically more numerous in the small size fractions, which is the opposite for clade F. This can be related to their respective phylogenetic position. Compared to the recently diverging clades (such as F), basal clades like I, III and B1 tend to have a very simple and fragile skeleton (simple spicules without central junction), which would be more susceptible to be destroyed through the sampling steps (Decelle et al 2012b, chapter I).

The fact that sample collections were performed at a given time (i.e. different seasons) across the oceans, it is difficult to highlight the major environmental drivers and the temporal dynamic of the Acantharia community. Still, many of the 35 stations explored in this study belong to biogeochemical provinces that have a low seasonality (westerlies and trades biomes) and a similar annual environmental variability (Vichi *et al.,* 2011). We therefore suppose that the Acantharia community might not vary drastically over the year in these stable environmental settings, contrary to coastal and high-latitudes waters. Overall, in order to fully elucidate the ecology and seasonal dynamics of Acantharia, and more specifically the symbiotic forms, it is highly required to investigate the community structure at a given location over several seasons or years.

Successful strategy of the photosymbiosis with *Phaeocystis*

Along with the reference database, the metabarcoding approach allowed us to compare the relative contribution of each group of Acantharia in different oceanic regions. This study

demonstrates that the photosymbiosis with the haptophyte *Phaeocystis* is a successful strategy in the photic zone of the oceans. The symbiotic Acantharia (clades E and F) clearly dominated the community in the photic zone of nearly all the locations sampled. This is in agreement with previous microscopic observations where symbiotic specimens were found to represent most of the acantharian biomass at the surface (Michaels, 1988; Stoecker, 1996; Schewiakoff, 1926; Taylor, 1982). Furthermore, compared to the non-symbiotic forms, symbiotic Acantharia exhibit a broader dispersal in the photic zone, a dichotomy also observed in pelagic Foraminifera (Norris, 1996). This can be explained by considering metapopulation dynamic whereby locally abundant species have higher chances of becoming widespread due to a lower extinction rate and/or a higher colonisation rate (Freckleton *et al.,* 2005).

As expected, the Acantharia-*Phaeocystis* photosymbiosis was found in warm and nutrient-poor tropical and subtropical waters, but the association was also detected in high latitudes in more productive and cold waters (e.g. stations 82, 84, and 85). Previous studies also found symbiotic Acantharia in the Iceland Basin and Arctic waters (Bernstein *et al.,* 1999; chapter III). This demonstrates that the association is not restricted to oligotrophic conditions but is rather adapted to different environmental settings. The lack of clear correlations with abiotic parameters that we found in this study for clades E and F can be explained by this large tolerance (Figures 6 and 7). One may question the functioning and ecological advantage of establishing a photosymbiosis in these nutrient-rich waters. Kleptoplastidy, another type of phototrophy acquisition, has been previously reported in antarctic dinoflagellates (Gast *et al.,* 2007), showing that this strategy also occurs in unconstrained nutrient availability. Photoperiod could be another environmental parameter that favours opportunistic physical associations with microalgae.

The higher abundance and wider dispersal lead to the suggestion that symbiotic Acantharia are generalists in the photic zone, i.e. the "jack-of-all-trades" is master of all (Brown *et al.,* 1995; Gaston *et al.,* 1997). Generalists are typically able to tolerate a broad range of environmental conditions and to rely on diverse and/or widely-distributed resources (Verberk *et al.,* 2010). For symbiotic Acantharia, a partnership with a photosynthetic symbiont would allow them to face periods of severe oligotrophy and to enlarge their optimal growth zone (Lombard *et al.,* 2009). In addition to this unlimited nutritional source, Acantharia are also known to maintain predation on a wide array of prey (Swanberg and Caron 1991), which relaxes constrains for range expansion. But most importantly, the main required resource for photosymbiotic hosts is the free-living reservoir of their obligate symbionts.

Given the obligatory and highly dynamic reset of the relationship for Acantharia, the geographic distribution of free-living *Phaeocystis* should be a major constrain of the distribution of its host. *Phaeocystis* is known to be cosmopolite from polar to tropical and subtropical waters, more particularly in coastal regions (Schoemann *et al.,* 2005). This study confirmed and further highlighted the ubiquitous distribution of this microalga since reads were found in every station from low to high latitudes in coastal but also in oceanic waters. This widespread occurrence of *Phaeocystis* could explain the high dispersal of the symbiotic Acantharia in the world ocean. The interaction specificity has a strong biogeographic pattern whereby Acantharia is able to live with

the indigenous *Phaeocystis* species (Decelle *et al.,* 2012a, chapter II). Like invertebrates-*Symbiodinium* photosymbioses (Finney *et al.,* 2010), this flexibility allows Acantharia to interact with the well-adapted local symbiont that very likely provides a high fitness for the holobiont. For instance, Acantharia in Antarctic regions live with the species *Phaeocystis antarctica* that can withstand and remain metabolically active during long periods of darkness and freezing temperatures in austral winter (Tang *et al.,* 2009). Because of the low resolution of the V9 region for the genus *Phaeocystis*, it was not possible here to distinguish the different symbiont species (identical V9 sequence for the species *P. globosa*, *P. antarctica* and *P. pouchetii*). Future studies should examine whether Acantharia accommodate different *Phaeocystis* according to the season at a given geographic location. Furthermore, *Phaeocystis* has a very complex life-cycle with at least 6 stages (Peperzak and Gäebler-Schwarz 2012). The symbiotic phase may be restricted to few of these, therefore having a direct influence on the availability and acquisition of symbionts for Acantharia.

The broad distribution but also the high abundance of *Phaeocystis* is evidently advantageous for Acantharia. This microalga that is known to be a key player in pelagic ecosystems (Verity and Smetacek, 1996) is a highly suitable partner since its abundance theoretically increases encounter rate with the host, hence facilitating the initiation of the symbiosis in the vast, turbulent and microbial-diluted milieu. In this study, no co-occurrences between the two partners were highlighted with the V9 reads. An increased abundance of *Phaeocystis* does not trigger necessarily a higher prevalence of symbiotic Acantharia (Figure 6). As a consequence, this reasonably well-characterized symbiosis (e.g. specificity, biogeography) which could be a test case, may not be detected by co-occurrence network analyses that investigate biological interactions in the microbial community (Chaffron *et al.,* 2010; Steele *et al.,* 2011).

Noteworthy, reads of both Acantharia (clade F) and *Phaeocystis* seemed to increase at station 84. This mirrors a previous study of iron fertilization experiments in the Southern Ocean whereby *Phaeocystis* positively responded, and was the most abundant phytoplankton taxon during the experiment after the diatoms (Hoffmann *et al.,* 2006). Besides, the microalga can form annual blooms in these polar regions during the summer (the period of our sampling) that have a major impact on the sulfur cycle and carbon dioxide sequestration (DiTullio, 2000; Arrigo *et al.,* 1999). In the microzooplancton, an increased population by twofold was detected during the experiment only for the Acantharia, and not for the other Radiolaria and Foraminifera (Henjes *et al.,* 2009). This parallel population change may be interpreted as a symbiont-host response and show the putative impact of the symbiont on the ecology of its host.

The worldwide distribution and the homogeneous diversity of symbiotic Acantharia in the photic zone (lack of distance-decay relationship) raise questions about whether there are geographical/physical barriers in the upper layer of the oceans. Is everything everywhere, but the symbiont selects? Another gene marker with a higher resolution than V9 should be used in future studies to carefully address the biogeographic partitioning of Acantharia and *Phaeocystis*.

Is Acantharia-*Phaeocystis* a true mutualism?

Photosymbioses are considered as mutualistic partnerships, meaning that the life cycle and a higher fitness depend on the interaction. This implies an important mutual dependency and symmetry of the relationship: the range distribution of one partner is constrained by the other's. Unfortunately, the omnipresence of *Phaeocystis* reads in each station provided us from examining the association strength of the relationship in this study. It would have been highly informative to observe whether symbiotic Acantharia occur when free-living *Phaeocystis* is absent, and vice versa if *Phaeocystis* can survive without its host. However, because of its extensive population and lack of marked co-occurrences, it is unlikely that *Phaeocystis* strongly depends on Acantharia for its survival and to gain a higher fitness.

Materials and Methods

Study sites and general sampling.

Biological samples were collected in the photic zone (surface and Deep Chlorophyll maximum, DCM) at different stations worldwide during the TaraOceans expedition in 2009-2011 (Karsenti et al., 2011; details in Table 1). The big size fractions (20-2000 µm) were obtained by towing two plankton nets of 20 and 180-µm-mesh-size, from which the samples once onboard passed through a sieve of 180-µm and 2000-µm sieve, respectively. The bulk samples were then filtered on a 10-µm membrane filter (diameter: 47 mm). For the small size fractions, 100 liters of seawater were gently pumped out on the deck with a peristaltic pump (TechPompes, France) and prefiltered through consecutive 20-µm and 5-µm nets prior to fill a tank. The pump that is equipped with a 40-mm-diameter hose has the advantage to present minimal shear damage for planktonic cells. The seawater in the tank was then filtered on a 0.8-µm membrane filter (diameter: 142 mm). All filters were flash-frozen and preserved in liquid nitrogen onboard, and stored at – 80°C.

Genomic DNA extraction, purification and sequencing

DNA was extracted with the Nucleospin® DNA II kit (Macherey-Nagel, Hoerdt, France). The V9 fragment of the ribosomal 18S rDNA was amplified with the eukaryotic primers 1389f 5'-GTACACACCGCCCGTC-3' and 1510r 5'-CCTTCYGCAGGTTCACCTAC-3'. Amplifications were conducted with the Phusion® High-Fidelity DNA Polymerase (Finnzymes). The PCR mixture (25 µL final volume) contained 5 ng of template with 0.35 µM final concentration of each primer, 3 % of DMSO and 2X of GC buffer Phusion Master Mix (Finnzymes). Amplifications were done following the PCR program: initial denaturation step at 98°C for 30 sec, followed by 25 cycles of 10 sec at 98°C, 30 sec at 57°C, 30 sec at 72°C, and final elongation step at 72°C for 10 min. Each sample was amplified in triplicate to get enough amounts of amplicons. Products of the reactions were run on a 1.5 % agarose gel to check for successful amplification products of the expected length. Amplicons were then pooled, purified using the

NucleoSpin® Extract II kit (Macherey-Nagel, Hoerdt, France), and sent to the CEA Genoscope in Evry. A bridge amplification and pyrosequencing were performed using a Genome Analyser IIx system (Illumina, San Diego, CA, USA).

The quality of the sequences was screened, and only sequences having exact forward and reverse primer match were considered. The V9 reads were assigned with a reference database of V9 sequences of Acantharia (Decelle et al 2012b, chapter I) and eukaryotes (Guillou *et al.*, 2012). Each environmental sequence was compared to all reference sequences using an exact global pairwise alignment algorithm (Needleman and Wunsch 1970) and received the taxonomic assignation of its nearest-neighbor in the reference database (or of the last common ancestor in case of a tie). Pairwise alignments were computed with ggsearch, a tool from the version 36 of the FASTA program package (Pearson and Lipman 1988), using default parameters, and results were stored in a sqlite database. Only reads assigned with more than 85 % of identity with a reference sequence were considered in this study. *Sequence data will be deposited in the MG-RAST public database…*

For some analyses, only reads assigned with more than 85 % of identity to a reference sequence and present at least 80 times in the dataset were considered ("principal reads"). This threshold has been selected based on the alignment of the V9 reads along with reference sequences. Sequences in low copy numbers (generally below 80 copies) turned out to contain random mutations and probably constitute artificial microdiversity (i.e. sequencing errors).

Statistical analyses

Community composition of Acantharia was plotted as pie charts for each sampling station on a map with the program GenGiS (Parks et al 2009). *Alpha* diversity was assessed by the Shannon index and represented by a small red dot at the precise sampling location (Figure 3).

For *beta*-diversity measurements, Unifrac distance metric was used to test whether the phylogenetic lineages of Acantharia between samples (e.g. stations, size fractions, oceanic regions etc.) are significantly different or homogeneous. To do so, the principal reads of clades E and F (assigned ≥85% and in ≥ 80 copie numbers) were first aligned with MAFFT v6.818b (options: maxiterate 1000, global pair). Phylogenetic trees were then built with FastTree 2 (Price et al., 2010) with the GTR model and sequences of other radiolaria as outgroup. These trees were used for inferring the Unifrac values. Principal coordinates analyses (PCoA) were performed in the FastUniFrac web interface (http://bmf2.colorado.edu/ fastunifrac/; Lozupone et al 2006; Hamady *et al.*, 2010) and visualized with the java application King 2.2.

Abiotic and biotic parameters and co-occurrences analyses

Chlorophyll *a* (Chl *a*), temperature, salinity, oxygen concentration (monthly means) were acquired at each station with a CTD rosette throughout the water column. Concentrations of silicates, nitrates and phosphates were obtained from the World Ocean Atlas 2009 (WOA09) at http://www.nodc.noaa.gov/OC5/WOA09/pr_woa09.html. The Day length (hours of light) was retrieved for each station at the same date of the sampling, on the website of the United States Naval Observatory (USNO) at http://aa.usno.navy.mil/data/docs/Dur_OneYear.php.

Co-occurrence analyses were performed using the tool CoNet (http://psbweb05.psb.ugent.be/conet/) that allows combining various metrics in order to detect significant non-random associations among features (e.g. species, genes, physico-chemical parameters, etc) in incidence or abundance data. Four different metrics were combined to detect significant associations between protists and environmental parameters: the Spearman correlation, the Steinhaus similarity as well as the Kullbackleibler and Bray-Curtis distances. The presented networks are combined networks (obtained from each measure, p-values were merged using the Simes method) in which each node is supported by at least two significant measures (after adjustement of p-values using the Benjamini & Hochberg step-up FDR-controlling procedure, FDR=0.05). Networks were explored and visualized with the Cytoscape platform (Shannon et al., 2003).

References

Arrigo KR, et al. (1999). Phytoplankton community structure and the drawdown of nutrients and CO2 in the southern ocean. *Science* **283**: 365–367.

Brown JH, Mehlman, DW, Stevens GC. (1995). Spatial variation in abundance. *Ecology* **76**: 2028-2043.

Brown MV, Philip GK, Bunge JA, Smith MC, Bissett A, Lauro FM, Fuhrman JA, Donachie SP. (2009). Microbial community structure in the North Pacific ocean. *ISMEJ* **3**: 1374-86.

Caron DA. (2000). Symbiosis and mixotrophy among pelagic microorganisms. In: Kirchman, D.L. (Ed.), *Microbial ecology of the oceans*, John Wiley & Sons, Inc., New York, pp. 495-523.

Countway PD, Caron DA, Gast RJ, Savai P. (2007). Comparison of protistan diversity in deep (2500m) vs euphotic zone assemblages in the Sargasso Sea and Gulf Stream (N.Atlantic). *Environ Microbiol* **9**: 1219-1232.

Decelle J *et al.*, (2012a). An original mode of symbiosis in open ocean plankton. *PNAS* **44**-18000-18005.

Decelle J, Suzuki N, Mahé F, de Vargas C, Not F. (2012b). Molecular phylogeny and morphological evolution of the Acantharia (Radiolaria) *Protist* **163**: 435-450.

DiTullio GR, Grebmeier JM, Arrigo KR, Lizotte MP, Robinson DH, Leventer A, Barry JP, VanWoert ML, Dunbar RB. (2000). Rapid and early export of *Phaeocystis antarctica* blooms in the Ross Sea, Antarctica. *Nature* **404**: 595-8.

Douglas AE. (1998). Host benefit and the evolution of specialization in symbiosis. *Heredity* **81**: 599-603.

Febvre J, Febvre C, Michaels A (2000) Acantharia Haeckel. (1881). In Lee JJ, Leedale GF, Bradbury P (Eds) An Illustrated Guide to the Protozoa (2nd edition). Organisms Traditionally Referred to as Protozoa, or Newly Discovered Groups, Society of Protozoologists, Kansas, 783–803.

Febvre J, Febvre-Chevalier C. (1979). Ultrastructural study of zooxanthellae of three species of Acantharia (Protozoa: Actinopoda), with details of their taxonomic position in the prymnesiales (prymnesiophyceae, Hibberd, 1976). *J Mar Biol Ass UK* **59**: 215-226.

Finney CJ, Pettay DT, Sampayo EM, Warner ME, Oxenford HA, LaJeunesse TC. (2010). The relative significance of host-habitat, depth, and geography on the ecology, endemism, and speciation of coral endosymbionts in the genus *Symbiodinium*. *Microb Ecol* **60**: 250-263.

Foster RA, Kuypers MM, Vagner T, Paerl RW, Musat N, Zehr JP. (2011). Nitrogen fixation and transfer in open ocean diatom-cyanobacterial symbioses. *ISMEJ* **5**: 1484-1493.

Foster RA, Collier JL, Carpenter EJ. (2006a). Reverse transcription PCR amplification of cyanobacterial symbiont 16S rRNA sequences from single non-photosynthetic eukaryotic marine planktonic host cells. *J Phycol* **42**: 243-250.

Foster RA, Carpenter EJ, Bergman B. (2006b). Unicellular cyanobionts in open ocean dinoflagellates, radiolarians, and tintinnids: ultrastructural characterization and immunolocalization of phycoerythrin and nitrogenase. *J Phycol* **42**: 453-463.

Freckleton RP, Gill JA, Noble D, Watkinson AR. (2005). Large-scale population dynamics, abundance–occupancy relationships and the scaling from local to regional population size. *Journal of Animal Ecology* **74**: 353-364.

Gast RJ, Moran DM, Dennett MR, Caron DA. (2007). Kleptoplasty in an Antarctic dinoflagellate: Caught in evolutionary transition? *Environ Microbiol* **9**: 39-45.

Gaston KJ, Blackburn TM, Greewood JJD, Gregory RD, Quinn RM, Lawton JH. (2000). Abundance –occupancy relationships. *J. Applied Ecology*. **37**: 39-59.

Gaston KJ, Blackburn TM, Lawton JH. (1997). Interspecific abundance-range-size relationships: an appraisal of mechanisms. *Journal of Animal Ecology* **66**: 579-601.

Gilg IC, Amaral-Zettler LA, Countway PD, Moorthi S, Schnetzer A, Caron D. (2010). Phylogenetic affiliations of mesopelagic acantharia and acantharian-like environmental 18S rRNA genes off the Southern California coast. *Protist* **161**: 197-211.

Gómez, F., Furuya, K., and Takeda, S. (2005). Distribution of the cyanobacterium *Richelia intracellularis* as an epiphyte of the diatom *Chaetoceros compressus* in the western Pacific Ocean. *J Plankton Res* **27:** 323–330.

Haeckel E. (1888). Die Radiolarien (Rhizopoda Radiolaria). Eine Monographie, Dritter Teil, Die Acantharien oder Actinopyleen-Radiolarien. Verlag von Georg Reimer, Berlin, pp 33

Henjes J, Assmy P, Klaas C, Smetacek V. (2007). Response of the larger protozooplankton to an iron-induced phytoplankton bloom in the Polar Frontal Zone of the Southern Ocean (EisenEx). Deep Sea Res I. **54**: 774-791.

Hoffmann LJ, Peeken I, Lochte K, Assmy P, Veldhuis M. (2006). Different reactions of Southern Ocean phytoplankton size classes to iron fertilization. *Limnol Oceanogr* **51**: 1217-1229.

Hollande A, Cachon J, Cachon-Enjumet M (1965) Les modalités de l'enkystement présporogénétique chez les acanthaires. *Protistologica* **1**: 91-112 [In French].

Janson, S, Wouters J, Bergman B, Carpenter EJ. (1999a). Host specificity in the *Richelia*–diatom symbiosis revealed by *hetR* gene sequence analysis. *Environ Microbiol* **1:** 431**-** 438.

Karsenti E, Acinas SG, Bork P, Bowler C, De Vargas C, *et al.* (2011). A Holistic Approach to Marine Eco-Systems Biology. *PLoS Biol* **9**: e1001177. doi:10.1371/journal.pbio.1001177.

Liu H *et al.*, (2009). Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *PNAS* **106**: 12803-12808.

Margulis L, Fester R. (1991). Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis (MIT Press, Cambridge, MA).

Michaels AF, Caron DA, Swanberg NR, Howse FA, Michaels CM. (1995). Planktonic sarcodines (acantharia, radiolaria, foraminifera) in surface waters near Bermuda: abundance, biomass and vertical flux. *J Plankton Res* **17**: 131-163.

Michaels AF. (1988). Vertical distribution and abundance of Acantharia and their symbionts. *Marine Biology* **97**: 559-569.

Needleman SB, Wunsch, CD. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* **48**: 443-453.

Norris RD. (1996). Symbiosis as an evolutionary innovation in the radiation of Paleocene planktic foraminifera. *Paleobiology* **22**: 461-480.

Not F, Gausling R, Azam F, Heidelberg JF, Worden AZ. (2007). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ Microbiol* **9**: 1233-1252.

Nyholm SV, McFall-Ngai MJ. (2004). The winnowing: Establishing the squid-vibrio symbiosis. *Nat Rev Microbiol* **2**: 632-642.

Parks DH, Porter M, Churcher S, Wang S, Blouin C, Whalley J, Brooks S and Beiko RG. (2009). GenGIS: A geospatial information system for genomic data. *Genome Research*, **19**: 1896-1904.

Pearson WR, Lipman DJ. (1988). Improved tools for biological sequence comparison. PNAS **85**: 2444-2448.

Pedrós-Alió C (2012) The rare bacterial biosphere. *Annu Rev Mar Sci* 4: 449-466.

Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* 5(3): e9490. doi:10.1371/journal.pone.0009490

Pringle A, Bever JD, Gardes M, Parrent JL, Rillig MC, Klironomos JN. (2009). Mycorrhizal symbioses and plant invasions. *Ann Rev Ecol Evol Syst* **40**: 699–715.

Prokopowich CD, Gregory TR, Crease TJ. (2003). The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **46**: 48-50.

Quaiser A, Zivanovic Y, Moreira D, López-García P. (2010). Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. *ISMEJ* **5**: 285-304.

Rai AN, Söderbäch E, Bergman B. (2000). Cyanobacterium symbioses. *New Phytol* **147:** 449–481.

Richardson DM, Allsopp N, D'Antonio CM, Milton SJ, Rejmanek M. (2000). Plant invasions - the role of mutualisms. *Biol Rev Camb Philos Soc* **75**: 65–93.

Schewiakoff WT. (1926). The Acantharia. Fauna e Flora del Golfo di Napoli **37**: 1-755.

Schoemann V, Becquevort S, Stefels J, Rousseau V, Lancelot C. (2005). *Phaeocystis* blooms in the global ocean and their controlling mechanisms: a review. *Journal of Sea Research* **53**: 43-66.

Shah Salani F, Arndt H, Hausmann K, Nitsche F, Scheckenbach F. (2012). Analysis of the community structure of abyssal kinetoplastids revealed similar communities at larger spatial scales. *ISMEJ* **6**: 713-723.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated modesl of biomolecular interaction networks. *Genome Res* **13**: 2498-2504.

Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow CE, Sachdeva R, Jones AC, Schwalbach MS, Rose JM, Hewson I, Patel A, Sun F, Caron DA, Fuhrman JA. (2011). Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISMEJ* **5**: 1414-1425.

Stoecker DK, Gustafson DE, Verity PG. (1996). Micro- and mesoprotozooplankton at 140°W in the equatorial Pacific: heterotrophs and micotrophs. *Aquat Microb Ecol* **10**: 273-282.

Strom SL. (2008). Microbial ecology of ocean biogeochemistry: A community perspective. *Science* **320**: 1043–1045.

Suzuki N *et al.*, (2009). Distribution Patterns of the Radiolarian Nuclei and Symbionts Using DAPI-Fluorescence. *Bull Natl Mus Nat Sci, Ser B* **35**: 169-182.

Swanberg NR, Caron DA. (1991). Patterns of feeding in epipelagic oceanic plankton. *J Plankton Res* **13**: 287-312.

Tang KW, Smith WO Jr, Shields AR, Elliott DT. (2009). Survival and recovery of *Phaeocystis antarctica* (Prymnesiophyceae) from prolonged darkness and freezing. *Proc Biol Sci*. **276**: 81-90.

Thompson AW, Foster RA, Krupke A, Carter BJ, Musat N, Vaulot D, Kuypers MM, Zehr JP. (2012). Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science* **337**: 1546-1550.

Thompson JN. (2005). The Geographic Mosaic of Coevolution (Univ. Chicago Press).

Verberk WC, van der Velde G, Esselink H. (2010) Explaining abundance-occupancy relationships in specialists and generalists: a case study on aquatic macroinvertebrates in standing waters. *J Anim Ecol* **79**: 589-601.

Verity PG, Smetacek V. (1996). Organism life cycles, predation, and the structure of marine pelagic ecosystems. *Mar Ecol Prog Ser* **30**: 277-293.

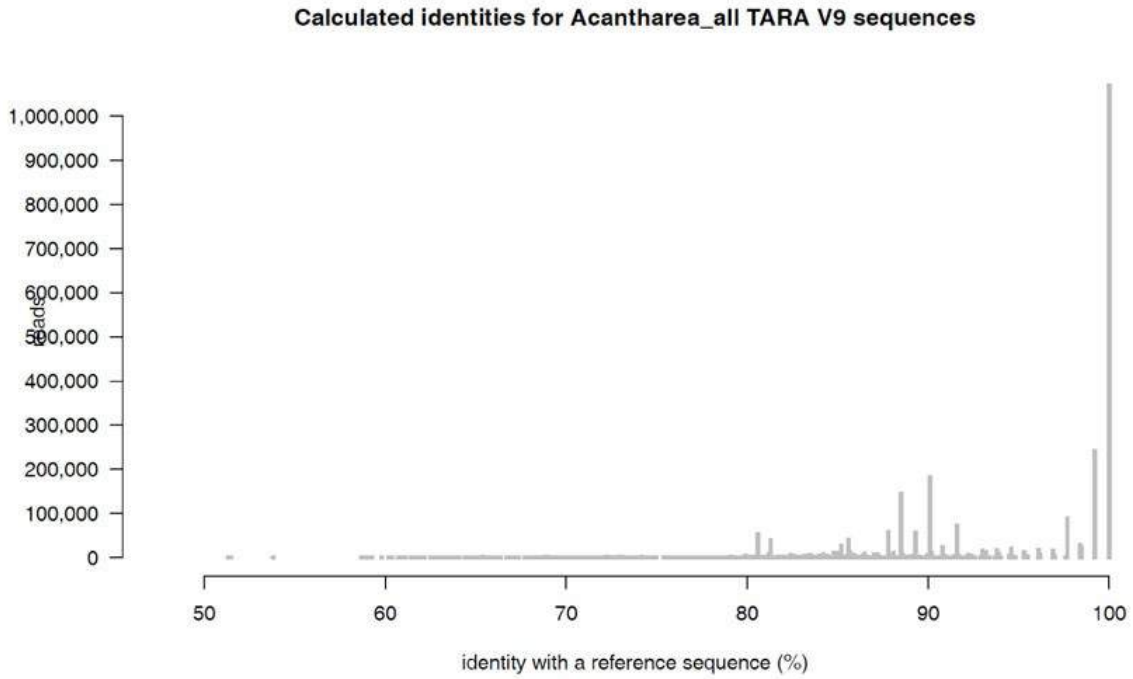
Vichi M, Allen JI, Masina S, Hardman-Mountford NJ. (2011). The emergence of ocean biogeochemical provinces: A quantitative assessment and a diagnostic for model evaluation. *Global Biogeochem Cycles* **25**: doi:10.1029/2010GB003867.

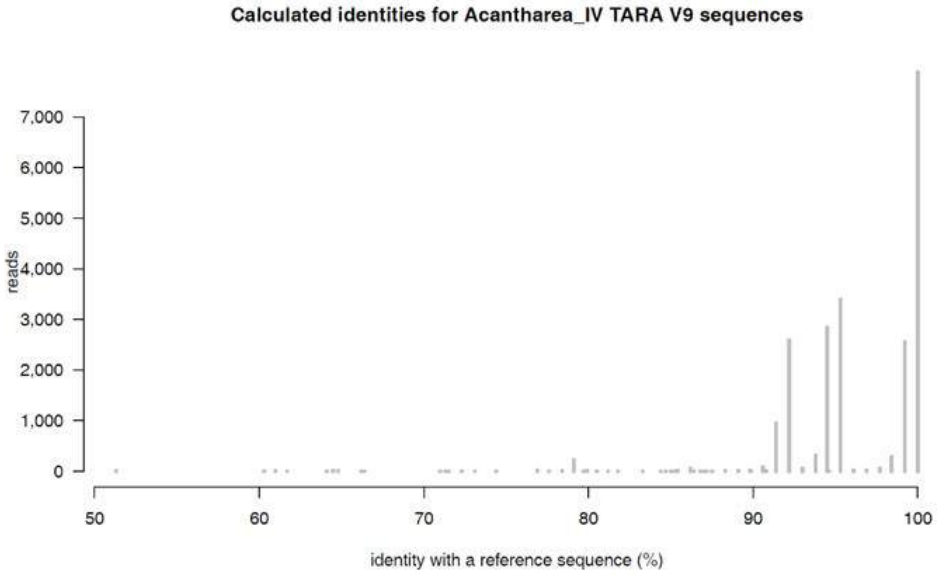
Zas'ko DN, Vedernikov VI. (2003). A comparative analysis of the vertical distribution of epipelagic radiolaria, chlorophyll, and zooplankton in different areas of the North Atlantic Ocean (from June to September 2001). *Oceanol* **43**: 63-71.

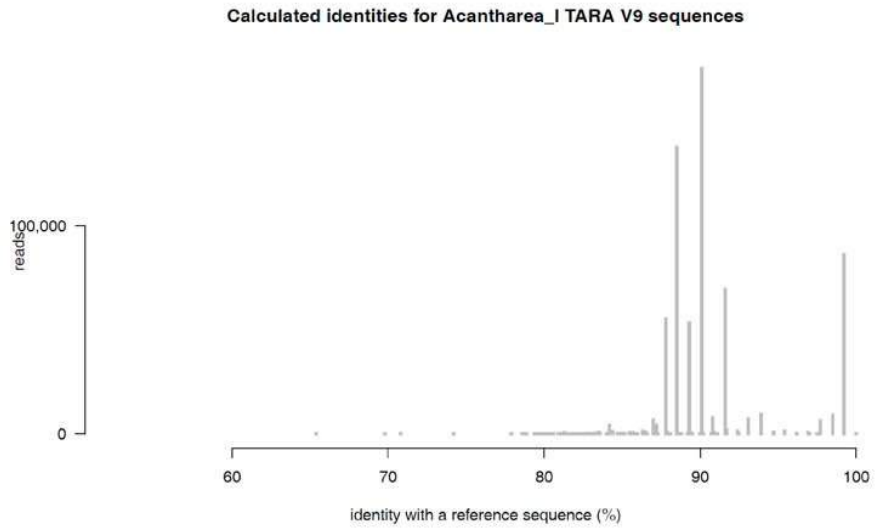
Zwirglmaier K *et al* (2008). Global phylogeography of marine S*ynechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. E*nv Microbiol* **10**: 147-161.

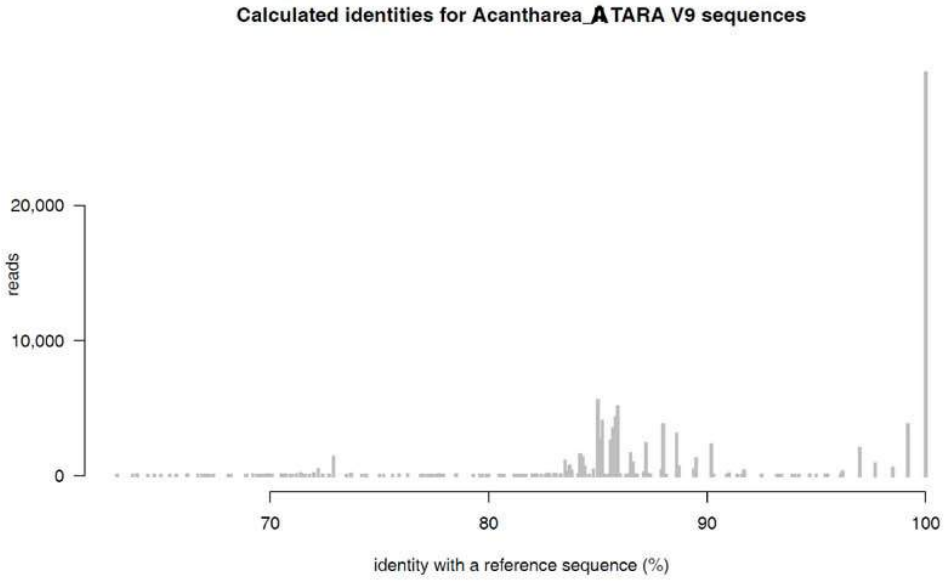
Supplementary information

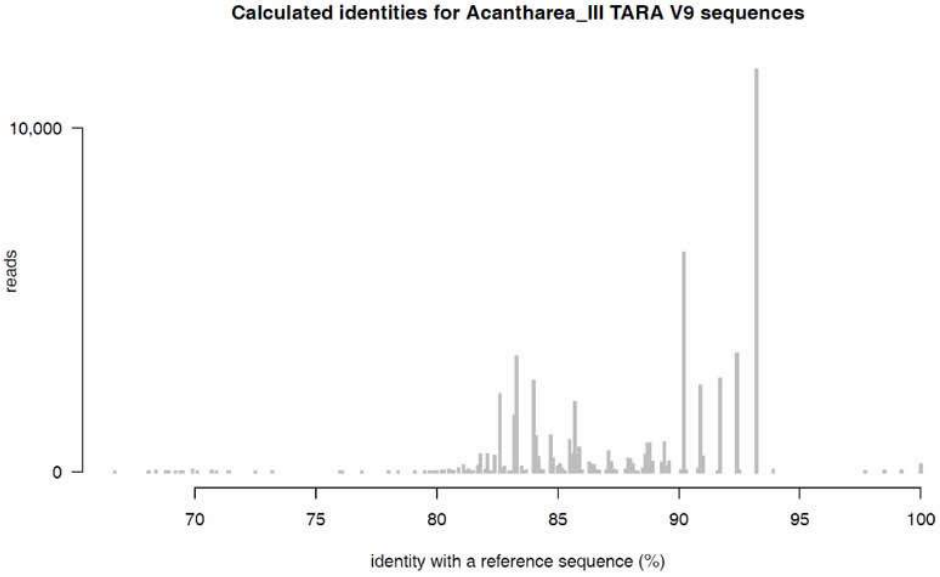
FIGURES

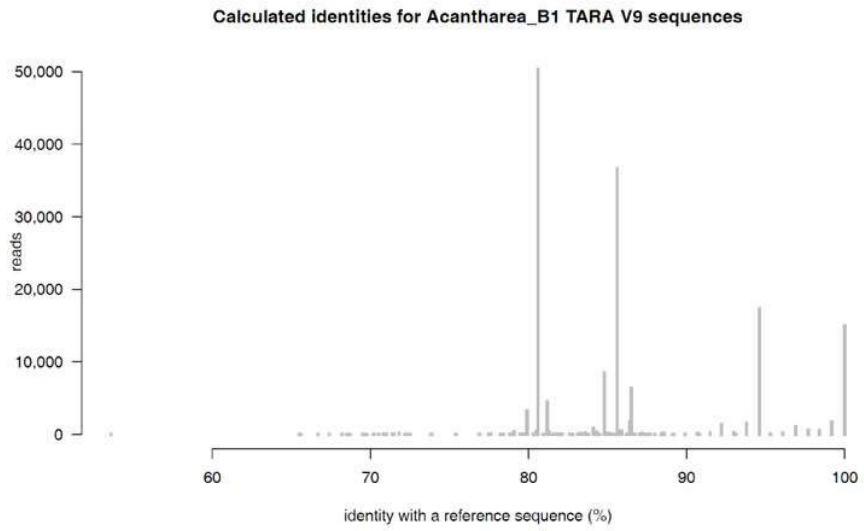


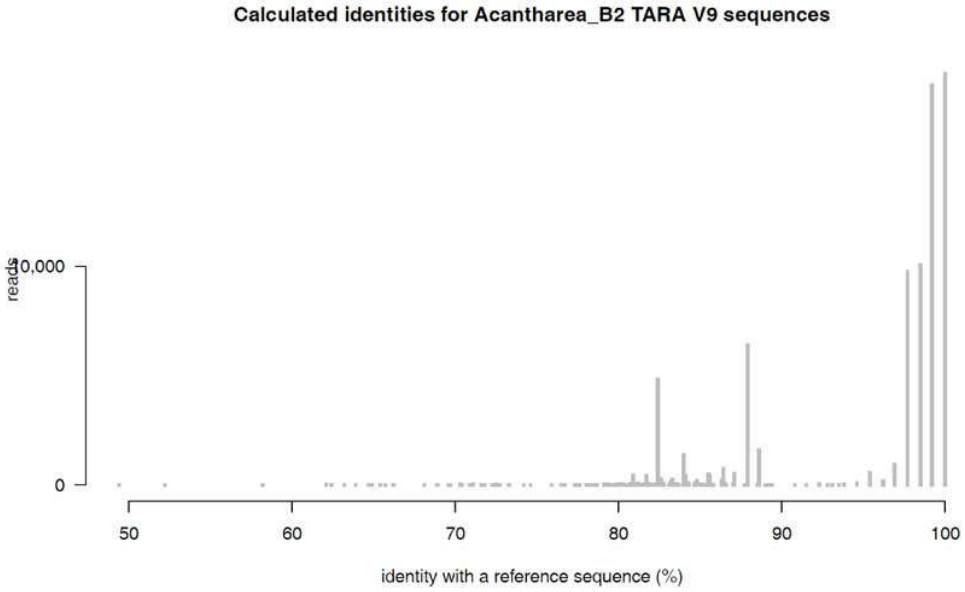


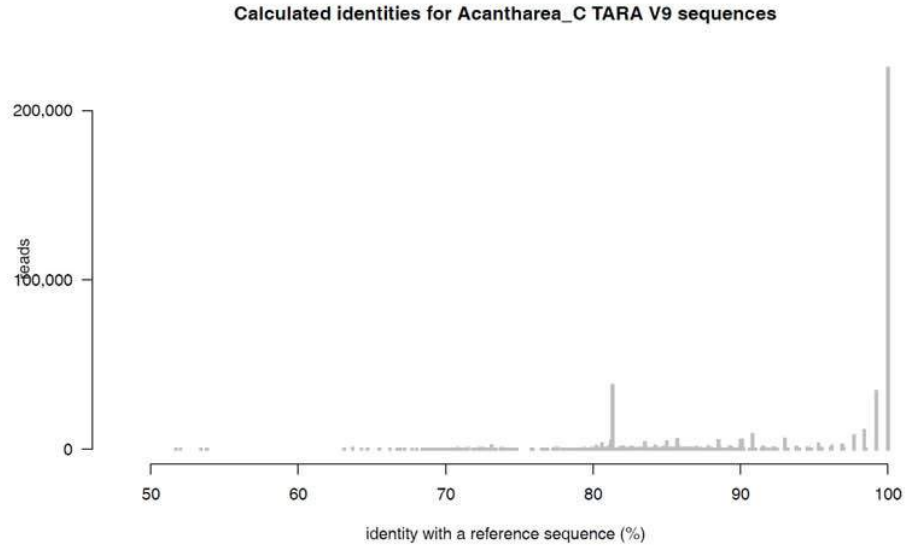

Calculated identities for Acantharea_I TARA V9 sequences
reads
100,000
0
60
70
80
90
100
identity with a reference sequence (%)


Calculated identities for Acantharea_A TARA V9 sequences
reads
20,000
10,000
0
70
80
90
100
identity with a reference sequence (%)


Calculated identities for Acantharea_III TARA V9 sequences
reads
10,000
0
70
75
80
85
90
95
100
identity with a reference sequence (%)


Calculated identities for Acantharea_B1 TARA V9 sequences
50,000
40,000
30,000
20,000
10,000
0
reads
60
70
80
90
100
identity with a reference sequence (%)


Calculated identities for Acantharea_B2 TARA V9 sequences
reads
0
50
60
70
80
90
100
identity with a reference sequence (%)


Calculated identities for Acantharea_C TARA V9 sequences
200,000
100,000
0
reads
50
60
70
80
90
100
identity with a reference sequence (%)

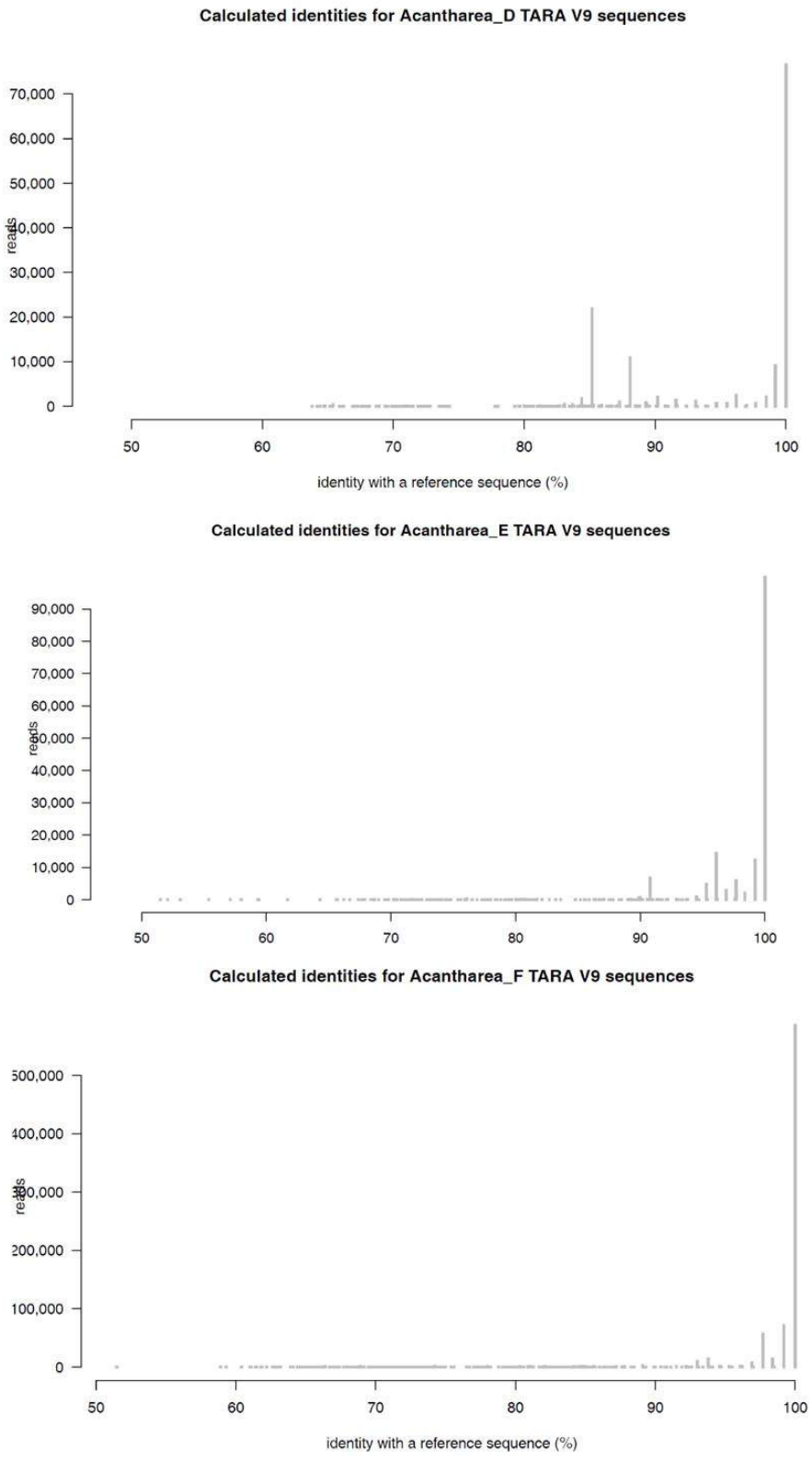


Figure S1: Assignation level of the environmental V9 reads to reference sequences from each acantharian clade.

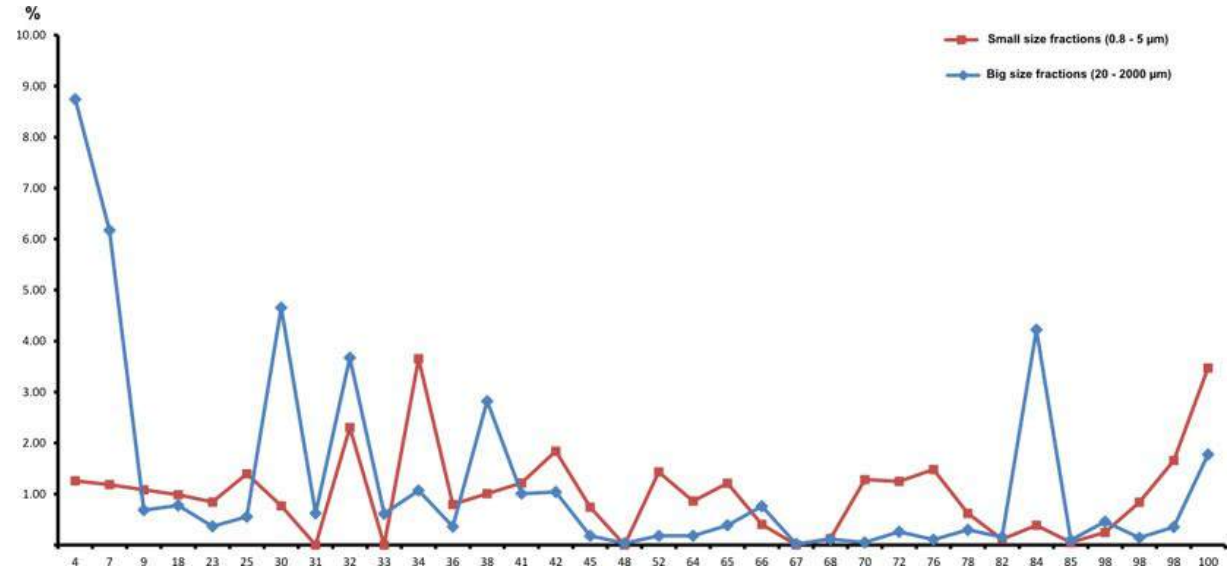


Figure S2: Acantharia contribution to protistan community in the small and the big size fractions

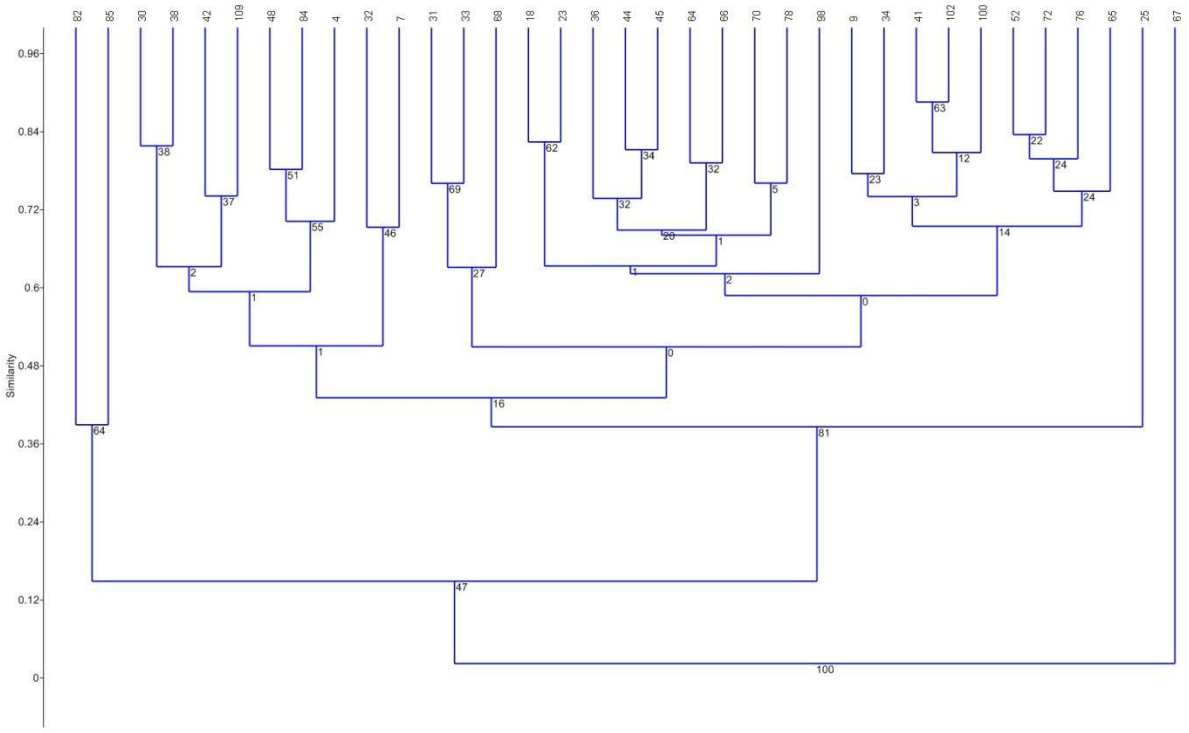


Figure S3: Clustering analysis based on the Bray-Curtis index to compare the acantharia communities between stations. (1000 bootstraps).

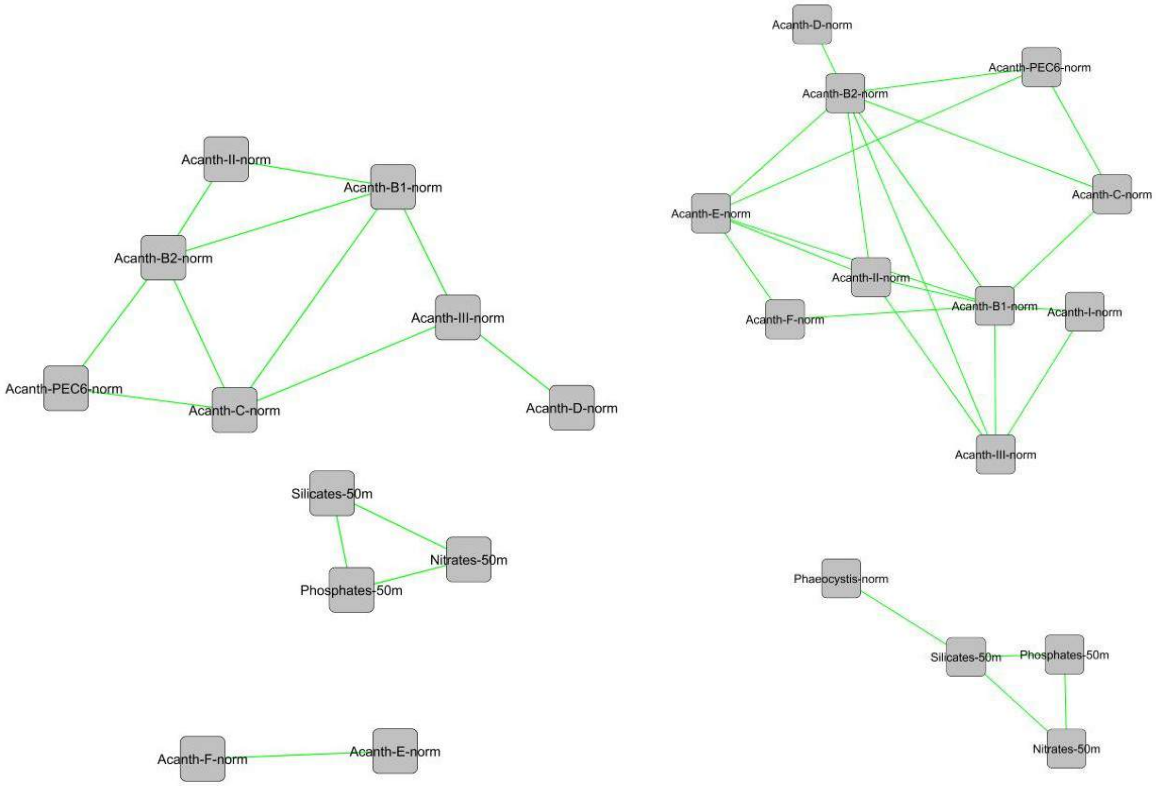


Figure S4: Correlation networks showing the statistically significant co-occurrences between biotic (reads of *Phaeocystis* and acantharian clades) and abiotic (oceanographic parameters) variables. The networks on the left were built without stations 82, 84 and 85.

TABLES

Table S1: Detailed information about each station

Station	Latitude	Longitude	Oceanic region	DATE	Season	Longhurst Province	Longhurst Province Name	Biome	number	Photoperiod length (hours of daylight)
4	35°33.79 N	6°26.835 W	North atlantic	September, 15 2009	summer	NASE	North Atlantic Subtropical Gyral	Westerlies	18	12.25
7	37°1.622 N	1°56.909 E	Mediterranean Sea	September, 23 2009	autumn	MEDI	Mediterranean Sea	Westerlies	16	12.07
9	39°7.101 N	5°53.275 E	Mediterranean Sea	September, 28 2009	autumn	MEDI	Mediterranean Sea	Westerlies	16	11.54
18	35°45.548 N	14°15.391 E	Mediterranean Sea	November, 2 2009	autumn	MEDI	Mediterranean Sea	Westerlies	16	10.41
23	42°10.559 N	17°43.534 E	Mediterranean Sea	November, 18 2009	autumn	MEDI	Mediterranean Sea	Westerlies	16	9.42
25	39°22.503 N	19°24.573 E	Mediterranean Sea	November, 23 2009	autumn	MEDI	Mediterranean Sea	Westerlies	16	9.47
30	33°55.393 N	32°48.17 E	Mediterranean Sea	December, 15 2009	autumn	MEDI	Mediterranean Sea	Westerlies	16	9.55
31	27°9 N	34°49.1 E	Red Sea	January, 9 2010	winter	REDS	Red Sea	Coastal	33	10.33
32	23°22.533 N	37°14.467 E	Red Sea	January, 11 2010	winter	REDS	Red Sea	Coastal	33	10.49
33	22°3.1 N	38°13 E	Red Sea	January, 13 2010	winter	REDS	Red Sea	Coastal	33	10.51
34	18°23.933 N	39°51.767 E	Red Sea	January, 20 2010	winter	REDS	Red Sea	Coastal	33	11.12
36	20°49.231 N	63°30.773 E	Indian Ocean	March, 12 2010	winter	ARAB	NW Arabian Upwelling Province	Coastal	34	11.57
38	19°1.231 N	64°33.225 E	Indian Ocean	March, 15 2010	winter	ARAB	NW Arabian Upwelling Province	Coastal	34	12.01
41	14°34.572 N	70°1 E	Indian Ocean	March, 30 2010	spring	ARAB	NW Arabian Upwelling Province	Coastal	34	12.15
42	6°0.336 N	73°54.063 E	Indian Ocean	April, 4 2010	spring	MONS	Indian Monsoon Gyres Province	Trades	30	12.11
44	2°48.471 N	71°31.496 E	Indian Ocean	April, 11 2010	spring	MONS	Indian Monsoon Gyres Province	Trades	30	12.1
45	1°7.821 N	71°32.872 E	Indian Ocean	April, 12 2010	spring	MONS	Indian Monsoon Gyres Province	Trades	30	12.08
48	9°35.415 S	66°19.351 E	Indian Ocean	April, 19 2010	autumn	MONS	Indian Monsoon Gyres Province	Trades	30	11.51
52	16°2.586 S	54°0.116 E	Indian Ocean	May, 17 2010	autumn	EAFR	East Africa Coastal Province	Coastal	32	11.21
64	29°29.6 S	37°56.561 E	Indian Ocean	July, 7 2010	winter	EAFR	East Africa Coastal Province	Coastal	32	10.2
65	35°45.299 S	26°19.199 E	Indian Ocean	July, 12 2010	winter	EAFR	East Africa Coastal Province	Coastal	32	9.54
66	34°5.444 S	18°0.209 E	South Atlantic	July, 15 2010	winter	EAFR	East Africa Coastal Province	Coastal	32	10.06
67	32°46.793 S	17°25.63 E	South Atlantic	September, 6 2010	winter	SATL	South Atlantic Gyral Province	Trades	10	11.35
68	31°58.416 S	5°21.612 E	South Atlantic	September, 12 2010	winter	SATL	South Atlantic Gyral Province	Trades	10	11.48
70	18°4.229 S	4°54.82 W	South Atlantic	September, 21 2010	winter	SATL	South Atlantic Gyral Province	Trades	10	12.05
72	8°13.415 S	17°5.202 W	South Atlantic	October, 4 2010	spring	SATL	South Atlantic Gyral Province	Trades	10	12.12
76	20°0.78 S	35°43.296 W	South Atlantic	October, 16 2010	spring	SATL	South Atlantic Gyral Province	Trades	10	12.34
78	30°47.629 S	43°41.878 W	South Atlantic	November, 3 2010	spring	SATL	South Atlantic Gyral Province	Trades	10	13.23
82	47°49.694 S	58°58.271 W	South Atlantic	December, 6 2010	spring	FKLD	SW Atlantic Shelves Province	Coastal	21	15.51
84	60°40.744 S	60°27.182 W	Antarctic	January, 3 2011	summer	ANTA	Antarctic Province	Polar	82	18.51
85	62°51.41 S	49°41.817 W	Antarctic	January, 6 2011	summer	APLR	Austral Polar Province	Polar	83	19.39
98	25°6.431 S	111°20.25 W	South Pacific	April, 2 2011	autumn	SPSG	South Pacific Subtropical Gyre	Westerlies	59	11.49
100	13°59.394 S	96°51.343 W	South Pacific	April, 14 2011	autumn	SPSG	South Pacific Subtropical Gyre	Westerlies	59	11.48
102	5°44.135 S	85°46.669 W	South Pacific	April, 21 2011	autumn	PEQD	Pacific Equatorial Divergence	Trades	62	11.57
109	1°58.286 N	84°26.772 W	South Pacific	May, 12 2011	spring	PEQD	Pacific Equatorial Divergence	Trades	62	12.12

Table S2: Oceanographic parameters for the 35 stations sampled in this study

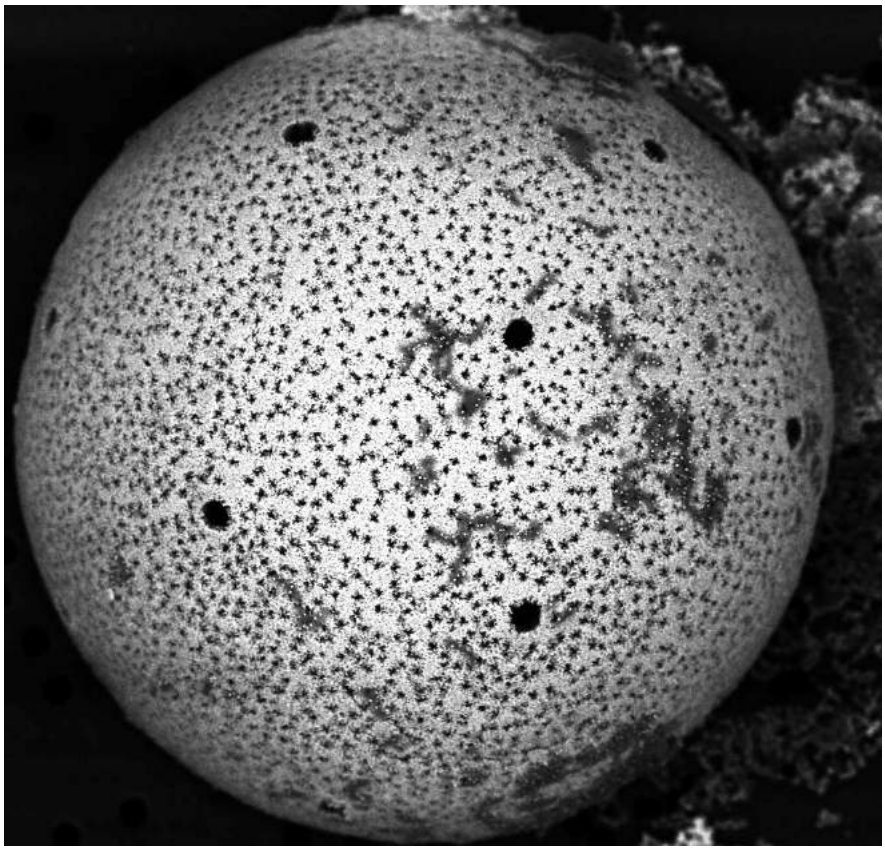
	Temperature [ITS-90, deg C]	Salinity [PSU]	Density [sigma-theta, Kg/m^3]	Oxygen [umol/Kg]	Nitrate_sensor [mol/kg-1]	Chloro_WETLabs [mg Chl/m3]	Chloro_HPLC [mg Chl/m3]	DCM	DCM Chla	Nitrates_50m	Silicates_50m	Phosphates_50m	Photoperiod (hours daylight)
4	NA	NA	NA	NA	NA	NA	NA	NA	NA	2.9862	1.8376	0.1512	12.25
7	21.4959	37.2522	26.0402	NA	NA	0.1998	0.1998	30	0.4409	1.2004	5.4352	0.0852	12.07
9	21.6184	37.6869	26.324	NA	NA	0.1013	0.1013	50	0.465	2.9438	3.3086	0.1176	11.54
18	21.4413	37.8968889	26.58014286	207.6266739	NA	0.011421053	0.011421053	84	0.0534	0.8496	1.7264	0.042	10.41
23	16.67869565	38.3230714	28.14576087	223.7207333	NA	0.058766667	0.058766667	54	0.1109	0.5722	3.1404	0.0446	9.42
25	17.27902632	38.3006296	28.05016667	225.633913	NA	0.045391304	0.045391304	54	0.079	0.2442	1.3222	0.0052	9.47
30	20.43161905	39.4214	28.02454545	207.5006087	0.197814815	0.016875	0.016875	71	0.0538	0.3886	1.4552	0.0952	9.55
31	25.0491	39.9931875	27.10722222	188.8033778	0.124	0.012444444	0.012444444	84	0.0976	0.5838	1.8668	0.1304	10.33
32	25.79635714	39.81	26.72164516	188.2424565	0.133428571	0.00808	0.00808	80	0.0877	0.4834	1.6986	0.0388	10.49
33	27.27992308	39.0093846	25.649	181.911913	0.15116129	0.072628571	0.072628571	37	0.1246	0.0888	0.859	0.0926	10.51
34	27.5805	38.7229333	25.33077778	184.7120652	0.150125	0.052166667	0.052166667	62	0.1144	0.1352	2.4572	0.1998	11.12
36	24.7588	36.5311818	24.57530435	NA	0.180645161	0.158	0.158	23	0.2639	2.3718	2.68	0.5386	11.57
38	25.67581395	36.593	24.37217073	189.9667609	0.212027778	0.123714286	0.123714286	40	0.1293	0.9056	2.362	0.4222	12.01
41	28.69512195	36.3965897	23.2357619	191.5039348	0.163125	0.00575	0.00575	73	0.0985	0.2564	2.3702	0.2248	12.15
42	29.85929032	34.5975217	21.47227586	189.9054348	0.139318182	0.0030625	0.0030625	80	0.1	0.3374	2.1262	0.1638	12.11
44	30.07585366	35.0204348	21.739625	188.8069565	0.127	0.016653846	0.016653846	66	0.118	0.1812	2.4096	0.218	12.1
45	30.05505556	35.2280714	21.84563043	182.9620435	0.201193548	0.050944444	0.050944444	54	0.1746	0.1696	2.9466	0.2376	12.08
48	27.16544444	34.6148824	22.31256098	179.5897391	0.1870625	0.069257143	0.069257143	58	0.1848	0.3518	2.9618	0.1698	11.51
52	27.89920588	34.6307727	22.14924	191.6276136	NA	0.045678571	0.045678571	74	0.116	0.1446	3.2156	0.1608	11.21
64	22.1762	35.335	24.4245	209.606	NA	0.0483125	0.0483125	78	0.0535	0.3336	2.2826	0.1578	10.2
65	21.821375	35.4465	24.6085	206.7096304	NA	0.079904762	0.079904762	27	0.0841	3.7396	4.4694	0.433	9.54
66	15.0278125	35.3235	26.2165	238.6117174	0.205392857	0.150342105	0.150342105	31	0.1759	3.8154	6.7522	0.6574	10.06
67	11.59226667	34.8792162	26.56463636	184.7370652	0.49575	0.416323529	0.416323529	7	1.3738	5.7656	4.6866	0.8262	11.35
68	16.8002381	35.687	26.08975	231.8774444	NA	0.14815	0.14815	59	0.1815	0.1622	4.1964	0.4418	11.48
70	19.78526667	36.3755	25.8695	215.3704889	0.145217391	0.130157895	0.130157895	2	0.1382	0.3932	1.9634	0.3246	12.05
72	25.03109091	36.423	24.4085	198.7495116	0.133666667	0.0095	0.0095	111	0.1167	0.1982	1.5026	0.1446	12.12
76	23.30406452	37.0870625	25.43588235	205.9654222	0.126	0.008	0.008	138	0.0553	0.283	2.2506	0.128	12.34
78	19.831875	36.325	25.81853846	221.5041333	0.135619048	0.011875	0.011875	121	0.0758	0.1196	2.2332	0.0426	13.23
82	7.274352941	34.0468	26.6305	305.3226957	0.729896552	0.414795455	0.414795455	56	0.5051	11.2488	2.5668	0.8968	15.51
84	1.691666667	33.7265455	26.97758333	338.6624667	1.006222222	0.097382353	0.097382353	57	0.215	24.3828	27.6526	1.6908	18.51
85	0.620942857	34.356	27.55430769	343.2515581	1.165714286	0.080025	0.080025	82	0.4686	25.6062	64.7556	1.5274	19.39
98	25.08763636	36.3955	24.37259091	200.7547778	0.17637931	-0.002307692	-0.002307692	182	0.0623	0.0036	0.9422	0.0898	11.49
##	24.63071429	35.7302353	24.12967857	202.6430989	0.3800625	0.132613636	0.132613636	62	0.1691	4.862	1.5828	0.6214	11.48
##	21.20895652	34.9198125	24.29273913	140.0335435	0.744571429	0.241452381	0.241452381	39	0.2789	9.0558	6.7672	1.009	11.57
##	26.20582222	34.253375	22.42746341	196.6383913	0.297457143	0.196405405	0.196405405	29	0.2309	12.3452	9.5936	1.1196	12.12

Table S3: Number of total reads assigned to different acantharian clades, *Phaeocystis* and protists found in the photic zone of each station.

Stations	Acanth_B1	Acanth_B2	Acanth_C	Acanth_D	Acanth_E	Acanth_F	Acanth_I	Acanth_II	Acanth_III	Acanth_IV	Acanth_PEC6	TOTAL Acantharia	%F	%E	% E and F	Protistan reads	*Phaeocystis* reads
4	5066	589	3192	14057	5798	76738	875	387	2728	252	217	109899	69.8	5.3	75.1	3082733	11379
7	880	29	51910	546	3829	34342	2536	744	532	75	1125	96548	35.6	4.0	39.5	3224389	2770
9	3236	113	9968	692	1466	13669	346	704	1476	92	189	31951	42.8	4.6	47.4	3386437	6961
18	1055	1476	6160	10469	10478	14518	713	1238	1040	20	848	48015	30.2	21.8	52.1	4258365	5969
23	416	471	4586	11220	3230	13275	1445	948	1094	7	285	36977	35.9	8.7	44.6	6858854	12126
25	1022	2217	5814	33900	1430	7729	594	1881	974	39	228	55828	13.8	2.6	16.4	5925151	4924
30	2258	1710	3169	2174	2323	18230	473	468	368	148	994	32315	56.4	7.2	63.6	2135326	3412
31	495	39	746	664	386	2787	209	583	640	1822	5	8376	33.3	4.6	37.9	285895	1714
32	3910	3974	33464	4978	4525	15887	562	1724	1969	397	351	71741	22.1	6.3	28.5	2632328	3654
33	1108	78	1329	1360	691	4303	505	1424	1273	4024	8	16103	26.7	4.3	31.0	208924	5124
34	844	3908	9312	4485	1904	16855	3468	1547	1897	179	262	44661	37.7	4.3	42.0	1535403	696
36	241	79	3140	40	11972	14164	94	1088	737	189	14	31758	44.6	37.7	82.3	6215689	3388
38	2462	169	895	93	1688	21640	759	842	163	563	11	29285	73.9	5.8	79.7	2013140	1401
41	4441	687	10657	2541	11448	29233	1156	5753	1132	1214	650	68912	42.4	16.6	59.0	6243196	3232
42	5983	6561	22897	4095	9869	49254	991	10392	2381	2965	1370	116758	42.2	8.5	50.6	8127701	8836
44	100	21	445	167	920	3011	64	717	18	50	7	5520	54.5	16.7	71.2		562
45	123	59	926	173	2801	14902	24	126	124	95	63	19416	76.8	14.4	91.2	3487872	1468
48	24	204	448	194	2896	13425	12	339	38	301	74	17955	74.8	16.1	90.9	40859	417
52	2467	5715	8896	1585	5796	12863	967	5825	1023	735	681	46553	27.6	12.5	40.1	6506038	12286
64	1093	501	3209	1106	900	12367	593	993	648	162	194	21766	56.8	4.1	61.0	5056459	4444
65	6510	1498	5956	977	701	6876	856	1860	1085	75	97	26491	26.0	2.6	28.6	2996455	2835
66	306	237	3409	235	144	7452	215	396	1089	1093	56	14632	50.9	1.0	51.9	2859392	4397
67	3	4	80	5	10	40	15	1	5	1	5	169	23.7	5.9	29.6	867781	13691
68	1411	78	3521	323	362	3610	632	507	849	12	159	11464	31.5	3.2	34.6	2998333	19497
70	350	235	1575	717	242	21326	675	132	748	586	83	26669	80.0	0.9	80.9	2851287	9785
72	3458	4510	15904	5739	8583	15791	1144	4207	2123	1383	1305	64147	24.6	13.4	38.0	5106031	6584
76	4697	2678	22719	2494	6069	16393	1694	2599	2335	253	1320	63251	25.9	9.6	35.5	6214659	3618
78	1015	1493	4075	4149	3859	28883	2187	1021	1009	210	645	48546	59.5	7.9	67.4	7115938	9530
82	9	3	204	5	7	1267	48	54	567	0	39	2203	57.5	0.3	57.8	5316625	317094
84	96	313	6194	1	2	40699	32	475	58	0	14	47884	85.0	0.0	85.0	2302673	64067
85	126	5	369	13	14	1796	507	2205	3329	0	9	8373	21.4	0.2	21.6	7400759	817025
98	267	173	1880	456	2789	8769	156	48	156	73	125	14892	58.9	18.7	77.6	3659069	4406
100	2031	1122	9689	1452	7927	23019	918	3148	935	818	347	51406	44.8	15.4	60.2	4989122	3467
102	2998	1228	13834	4562	7443	40809	370	5232	1293	797	305	78871	51.7	9.4	61.2	4160774	9449
109	14324	7442	28699	16231	18526	96947	45524	12884	5532	1741	37	247887	39.1	7.5	46.6	4910360	9076
TOTAL	74825	49619	299271	131898	141028	702869	71359	72492	41368	20371	12122	1617222	1578.6	302.0	1880.6	134974017	1389284

Chapitre 4

Study of the life cycle of Acantharia: cyst formation



Phylogeny and biogeography of cyst-forming Acantharia shed light on their ecology and biogeochemistry in the oceans

Johan Decelle[1], Patrick Martin[2,4], Katsiaryna Paborstava[2], David W. Pond[3], Geraint Tarling[3], Frédéric Mahé[1], Colomban de Vargas[1], Richard Lampitt[2] and Fabrice Not[1]

1 CNRS, UMR 7144, EPPO, Université Pierre et Marie Curie, Station Biologique de Roscoff - Place Georges Teissier, 29682 Roscoff, France
2 National Oceanography Centre, Southampton, European Way, Southampton SO14 3ZH, United Kingdom
3 British Antarctic Survey, Madingley Road, High Cross, Cambridge Cambridgeshire CB3 0ET United Kingdom
4 Present address: Department of Marine Chemistry & Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

Accepted in ***PLos ONE*** with minor revisions

Abstract

Marine planktonic organisms that undertake active vertical migrations over their life cycle are important contributors to downward particle flux in the oceans. Acantharia, globally distributed heterotrophic protists that are unique in building skeletons of celestite (strontium sulfate), can produce reproductive cysts covered by a heavy mineral shell that sink rapidly from surface to deep waters. We combined phylogenetic and biogeochemical analyses to explore the ecological and biogeochemical significance of this reproductive strategy. Phylogenetic analysis of the 18S and 28S rDNA genes of different cyst morphotypes collected in different oceans indicated that cyst-forming Acantharia belong to three early diverging clades from the orders Chaunacanthida and Holacanthida that are mainly heterotrophic. Environmental high-throughput V9 tag sequences and clone libraries showed that the three clades are widely distributed in the Indian, Atlantic and Pacific Oceans at different latitudes, but seem more abundant in regions of higher primary productivity. Moreover, sequences of cyst-forming Acantharia were distributed evenly in both the photic and mesopelagic zone, a vertical distribution that we attribute to their life cycle where flagellated swarmers are released in deep waters from sinking cysts. Bathypelagic sediment traps in the subantarctic and oligotrophic subtropical Atlantic Ocean showed that downward flux of Acantharia was only large at high-latitudes and during a phytoplankton bloom. High organic carbon export in cold waters would be a putative nutritional source for heterotrophic juveniles ascending in the water column. This study improves our understanding of the life cycle and biogeochemical contribution of Acantharia, and brings new insights into a remarkable reproductive strategy in marine protists.

Introduction

The dark ocean, the largest habitat in the biosphere [1], is of major importance in global ocean biogeochemistry. Spanning the mesopelagic (200-1000 m) and bathypelagic (1000-4000 m) zones, the dark ocean is a key reservoir of organic carbon that is produced in the photic layer and sinks downward in form of particles. Large particles, such as fecal pellets and phytodetrital aggregates, but also groups of unicellular planktonic organisms, such as Radiolaria, are major contributors to the downward particle organic carbon (POC) flux [2]. Among the Radiolaria, Acantharia have been shown to contribute to mesopelagic and bathypelagic POC flux via cyst formation [3,4], thus playing an important role in the biological pump of carbon.

Acantharia are cosmopolitan and abundant protists, and are the only organisms known to precipitate the mineral celestite (strontium sulfate, $SrSO_4$). They also have the capacity to completely change their morphology over the course of their life cycle in order to form cysts. Unlike the cysts of other planktonic single-celled organisms, such as dinoflagellates, diatoms or ciliates, which are resting stages produced to overcome unfavorable environmental conditions [5,6,7], acantharian cysts are directly linked to reproduction [8 (p. 729),9,10,11]. Thousands of flagellated cells (~2-3 µm), called swarmers (probably gametes), are formed within each cyst and released either through pores or upon rupture of the cyst wall. However, some acantharian species do not form cysts and in this case, swarmers are produced directly through the vegetative stage (adult form) without encystment.

During cyst formation, the spicules of the skeleton and the main cellular components (myonemes, axopods, ectoplasm etc.) are resorbed. The cell loses its buoyancy and starts depositing a robust seed-like celestite shell of 5-7 µm in thickness, and generally bigger than the vegetative cell, up to 1 mm. [9,11]. Cysts exhibit a large variety of shape and forms, presumably produced by distinct acantharian families and species: oval, round, elongated, with or without mineral plates and pores etc. Some of them have been described in the past as new acantharian families and species [12] or classified according to their form [13] (e.g. forma Allas, Folium, Bimamma, Ampulla, Olive etc). However, since the morphological characters used for species identification completely disappear during encystment, and cyst formation is very rarely observed upon collecting adult stages, it is not clear which taxonomic groups are able to encyst. DNA identification could therefore be a useful tool to unambiguously assign certain types of cyst to adult morphospecies within molecular-based clades, and allow taxonomic distinction between cyst-forming and non-cyst-forming Acantharia.

Although Acantharia are mostly found in the surface ocean, numerous environmental DNA sequences (18S rDNA gene) have been found deeper in the mesopelagic and bathypelagic zones [14,15,16,17,18]. The deep sequences probably originate at least partly from sinking cysts and release of swarmers. Ballasted with the very dense mineral celestite (density = 3.96 $g.cm^{-3}$), about twice more than calcium carbonate of other protistan tests, acantharian cysts rapidly sink from the surface to depth, up to 500 $m.day^{-1}$ for the largest specimens [4], and presumably release swarmers far below the surface. Acantharian cysts have been collected in sediment traps and net samples around the world, both at high latitudes (Antarctic, East Greenland sea, Iceland

basin, [3,4,19]), and in tropical and subtropical waters (the Mediterranean Sea, the Atlantic, Indian and Pacific Oceans, [11,8,20,13]).

Acantharia contribute to the downward POC flux, which is one important factor regulating atmospheric CO_2 concentration [21]. Celestite is highly soluble in seawater and the cyst shell appears to dissolve quickly once cysts start sinking [3,22]. Most previous studies of acantharian fluxes, focusing mainly on low latitudes, have hence concluded that cysts contribute significantly to shallow POC fluxes, but not appreciably below the upper 200–300 m [3,19,23]. However, a particularly large type of acantharian cyst was recently found in sediment traps from 2000 m in the Iceland Basin, contributing up to 48% of POC flux during a 2-week sampling interval [4]. It was hypothesized that deep sedimentation of cysts might be limited to high latitudes and exhibit a seasonal pattern, where high primary productivity during seasonal phytoplankton blooms triggers encystment such that juvenile Acantharia at depth can exploit the subsequent pulse of sinking phytodetritus as a food source upon bloom collapse [4]. However, a more thorough study of the contribution of Acantharia to deep particle fluxes, and its latitudinal variability, has not yet been undertaken. Moreover, the biogeography and vertical distribution of the vegetative stage of cyst-forming Acantharia needs closer examination to understand how their ecology and life cycle influence the downward POC and strontium fluxes.

Beyond implications for the carbon biogeochemistry, Acantharia affect the oceanic cycling of strontium and barium, which are highly concentrated in cyst shells [3,24]. By precipitating celestite, Acantharia create vertical and horizontal gradients in strontium concentration [3,25,26]. Depletion of strontium in seawater due to Acantharia is sometimes claimed to be problematic for the use of strontium/calcium ratios in corals and planktonic foraminifera to reconstruct past sea-surface temperatures [26]. While its role remains enigmatic, strontium is also required for calcification in many marine organisms [27,28]. This does call for more examination of acantharian ecology and biogeochemical flux inherent to the life cycle.

Here, we first used molecular tools to phylogenetically identify the cyst-forming Acantharia, and then investigated their vertical distribution in different oceanic regions worldwide through V9 tag sequences. Finally, we assessed the biogeochemical contribution and seasonality of cyst sedimentation by measuring strontium fluxes at four sites in the Atlantic Ocean.

Materials and Methods

Collection of cysts

Cysts were collected in distinct oceanic regions. Some cysts were isolated from plankton net tows between 0 and 50 m in the Mediterranean Sea (Villefranche-sur-mer; Table 1) in 2010 and 2011. They were individually photographed and placed in a guanidine-containing extraction buffer (GITC). Additional cysts were collected during 2006–2008 in the Iceland basin, either at 160 and 620 m in 3-day deployments of neutrally buoyant PELAGRA sediment traps [2,29], or at 2000 m using year-long deployments of bottom-tethered sediment traps [4]. A few adult cells (vegetative stage) were also found in the PELAGRA traps. All trap samples were fixed in

formaldehyde preservative (see below). Cysts and adults were individually picked from the three depths, washed several times in PBS buffer to remove the formaldehyde, isolated in GITC buffer and preserved at -20°C before processing. All sampling locations are shown in Figure 1.

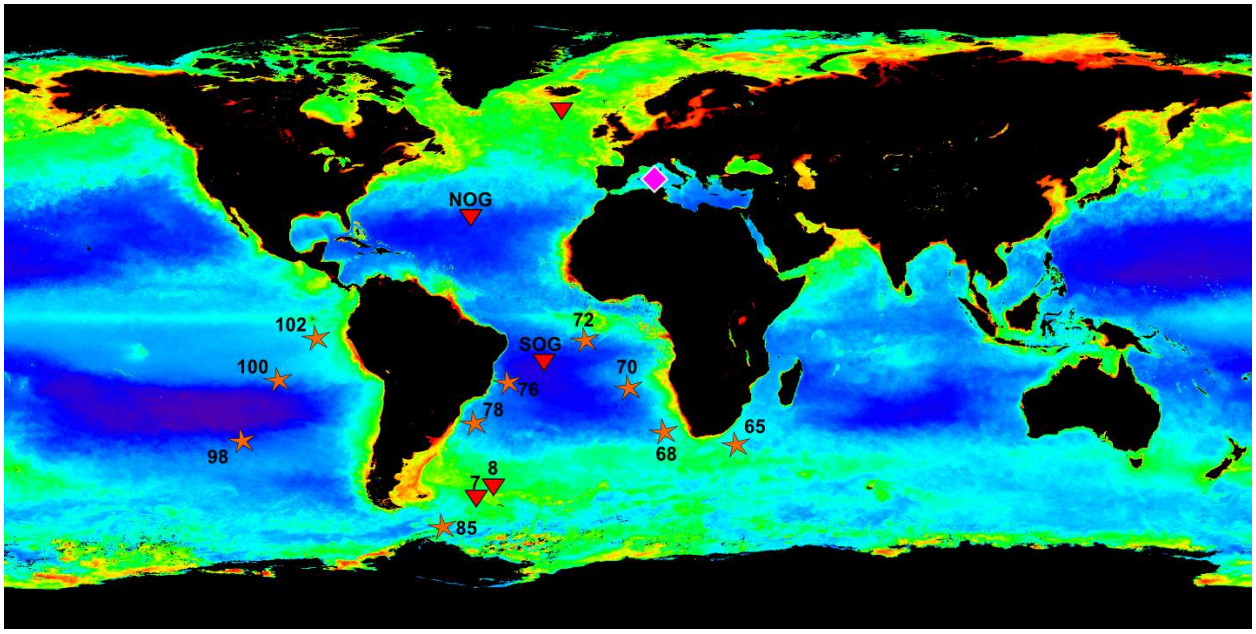


Figure 1: Sampling sites of this study. Global chlorophyll distribution map (higher concentrations indicated in green and red) showing the sampling locations of meso-and bathypelagic traps (▼), and surface plankton nets (◆). Genetic sequences (V9 region of the 18Sr DNA gene) have been also obtained from the surface and mesopelagic area (★). (Map obtained from OceanColor website: http://oceancolor.gsfc.nasa.gov/)

DNA extraction, amplification and phylogenetic analyses

DNA extraction was carried out on each isolated cyst as previously described [30]. Single-cell Polymerase Chain Reaction (PCR) was used to amplify 18S and 28S (D1 and D2 regions) ribosomal DNA markers with Radiolaria-specific primers (see 30 for primer details). Following amplification, the PCR products were purified by EXOSAP-IT (GE Healthcare Bio-Sciences Corp.), and bidirectionally sequenced using the ABI-PRISM Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems). Accession numbers of the 18S and 28S rDNA sequences of cysts are given in Table 2.

The matrices of 18S rDNA and partial 28S rDNA sequences were individually aligned with Muscle, implemented in Seaview [31]. As the topology of 18S and 28S rDNA genes are congruent [30], the two aligned matrices were concatenated with Sequence Matrix [32], forming a dataset of 104 taxa and 2386 bp-long. The GTR+G+I model of nucleotide substitutions was selected by MEGA5 [33] according to the BIC and AIC criteria. The phylogenetic inference by

Maximum Likelihood (ML) was then performed with PhyML v3.0 [34], and resampled 500 times by non-parametric bootstrapping. Bayesian inference was conducted on the concatenated dataset using Beast v.1.6.1 and companion software [35] under the GTR model taking into account 4-class gamma and invariant sites. Two Markov Chain Monte Carlo (MCMC) chains were run for 30 million generations, sampling every 1000 generations. The two runs were combined with LogCombiner v1.6.1, and convergence of log-likelihoods and parameter values were assessed in Tracer v1.4.1. 10% of the total trees were discarded as burn-in, and the remaining trees were used by TreeAnnotator 1.5.4 to build the consensus tree and to calculate the posterior probabilities (PP) of each node.

Biogeography and diversity of Acantharia in the surface and mesopelagic zones

Clone libraries from previous environmental surveys: All the 18S rDNA environmental sequences from clone libraries (Sanger technology) related to Acantharia (n = 260) were retrieved from Genbank version 190 in March 2012, noting any associated information for each sequence, such as location and depth of sampling. We used pplacer to compare the phylogenetic relationships between the environmental sequences and cyst sequences [36] with the 18S rDNA reference alignment. The advantage of using this program is that all the 18S rDNA environmental sequences can be comprehensively analyzed whatever their length and region. The 260 environmental sequences (query sequences) were aligned with a Hidden Markov Model built from the reference alignment using tools from the HMMER v3.0 suite (http://hmmer.org/), with default parameters. Using this alignment and the inferred reference tree, pplacer determined the most probable location for each query sequence and represented them as additional branches in the tree.

V9 sequence tags and taxonomic assignation: Seawater samples were collected with Niskin bottles in the mesopelagic zone at different locations in the South Indian, Atlantic and Pacific oceans during the TaraOceans expedition ([37], details in Table 1). Around 90 liters of seawater were immediately filtered onto successive a 0.8 µm mesh polycarbonate membrane, which was preserved in liquid nitrogen. DNA extraction was performed using the using the Nucleospin® DNA II kit. The V9 hypervariable region of the 18S ribosomal DNA gene was PCR amplified with general-eukaryote primers (1380F/1510R from [38]) and sequenced using the Illumina technology. The V9 reads were assigned with a reference database of V9 sequences of Acantharia [30] and eukaryotes (Guillou et al submitted 2012). Each environmental sequence was compared to all reference sequences using an exact global pairwise alignment algorithm [39], and received the taxonomic assignation of its nearest-neighbor in the reference database (or of the last common ancestor in case of a tie). Pairwise alignments were computed with ggsearch, a tool from the version 36 of the FASTA program package [40], using default parameters, and results were stored in a sqlite database. Only reads assigned with more than 85 % of identity with a reference sequence were considered in this study.

Strontium measurements in sediment trap samples

In addition to the Iceland Basin sediment traps from which we isolated Acantharia (described above), we measured dissolved Sr in the preservative solution of bathypelagic time-series sediment traps deployed at four locations in the subtropical and subantarctic (Scotia Sea) Atlantic Ocean. Since celestite is highly soluble and dissolves upon collection in sediment traps, the preservative solution of trap samples containing Acantharia becomes highly elevated in Sr [4]. Since no other organisms are known to precipitate Sr, very high Sr concentrations in sediment trap preservatives should unambiguously identify a contribution by Acantharia to the particle flux.

Deployment locations, depths, and collection periods for all traps, including the Iceland Basin ones, are listed in Table 1. Sr fluxes collected by the Iceland Basin traps were measured previously and are already published [4,29]. The subtropical and Scotia Sea traps were McLane Parflux Mark78H traps with 21 collection bottles, and were deployed at 2000 m (Scotia Sea) and 3000 m (subtropical). Collection bottles on the subtropical traps were filled with formaldehyde brine prior to deployment (2-5% formaldehyde with 0.5% w/v NaCl in seawater, buffered with 25 $mg.L^{-1}$ sodium tetraborate), those on the Scotia Sea traps with mercuric chloride-poisoned seawater. Upon recovery of the traps, samples were stored at 4ºC.

Aliquots of preservative solution from the collection bottles were diluted to between 1:500 and 1:4000 by volume with 3% sub-boiled HNO_3, spiked with 5 ng g^{-1} In and Re as internal standards, and Sr and Ca concentrations measured on a Thermo X series inductively coupled plasma mass spectrometer. The instrument was calibrated with mixed Ca + Sr standards containing 5 ng g^{-1} In and Re.

Sr fluxes into collection bottles with elevated Sr concentration were calculated after subtracting the background Sr concentration (7.24–8.42 µg mL^{-1}, depending on trap) by multiplying by the bottle volume (300 mL or 500 mL, depending on location) and dividing by the trap collection area (0.66 m^{-2}) and by the collection period (14–31 days). The background concentration of Sr in seawater is around 8 µg mL^{-1}, and in the sediment trap preservative samples was mostly between 7.1 and 8.4 µg mL^{-1}. We hence calculated the background Sr concentration in the samples as the mean concentration of all bottles containing less than 9 µg mL^{-1} Sr, and considered samples to be significantly elevated in Sr if their concentration was higher than mean + 2 * standard deviation of the bottles with <9 µg mL^{-1}. While the cut-off of 9 µg mL^{-1} is somewhat arbitrary, this is in practice not relevant, since even a modest Sr flux would elevate the preservative concentration by several µg mL^{-1} above the background.

Uncertainties for the Sr measurements were propagated by assuming 1% analytical uncertainty in each Sr measurement [4] and from the calculated standard deviation of the Sr concentration in bottles with <9 µg Sr mL^{-1}, i.e. the standard deviation of the background concentration.

Results

Molecular identification of acantharian cysts

We collected and PCR-sequenced the 18S and 28S rDNA genes of 3 morphotypes of acantharian cysts (pear-shaped, elongated and round), and reconstructed their phylogenetic relationships with other acantharian sequences (Figure 2). Photos of the cyst types and the most genetically related adult specimens (vegetative stage) are shown in Figure 2. The walls of all cysts showed numerous pore openings, but no small mineral plates were observed. All cysts sampled in this study belonged to the early diverging clades A, B and C, representing the taxonomic orders Chaunacanthida and Holacanthida. No cyst sequences were affiliated with clades E and F, which consist of taxa that harbor symbiotic microalgae. The pear-shaped cyst (cyst 50), found at 620 meters depth, branched within clade A (previously named clade II in [30]), and was closely related to the genus *Acanthoplegma* sp. (Ei 59) in the order Holacanthida. All the elongated cysts (> 10 specimens were sequenced), collected at different depths up to 2000 meters, are genetically identical and hold a phylogenetic position close to the specimen Ei 68 in subclade B1 (information about cysts are detailed in Table 2). The round cysts (cyst 25, 28, 45, 48) sampled at the same depths grouped within clade C (subclade C3), which represents the order Chaunacanthida. For these elongated and round cysts, we observed different developmental stages during the sinking process: spicules were still visible at 160 m but appeared partly resorbed at greater depths, and were totally absent in cysts from 2000 m. Two other round cysts with large orifices at one pole (Vil 162 and Pec 9) were found at the surface in the Mediterranean Sea and belonged to subclade C4. The individual Vil 162 was particularly closely related to the adult species *Gigartacon muelleri* (Vil 105, 117, 41, 52, 53, and 61): the 18S rDNA and 28S rDNA sequences were 100% and 99% identical, respectively.

In addition, several encystment events of the species *Acanthochiasma* sp. (Holacanthida) were observed and isolated through the microscope (supplementary videos). While crawling rapidly on the Petri dish, the cells expelled all their spicules but one, and excreted cytoplasmic remains and microalgal cells (presumably symbionts). Cyst formation was not completed in most cases because the cells died beforehand, but the final shape of the cyst that we could observe was round. These cysts had a similar genetic identity, belonging to clade B2.

Vegetative cells were also found in trap samples at 160 m (adult 20) and 620 m (adult 43 and 52), and their sequences grouped within clade C3 and F, respectively.

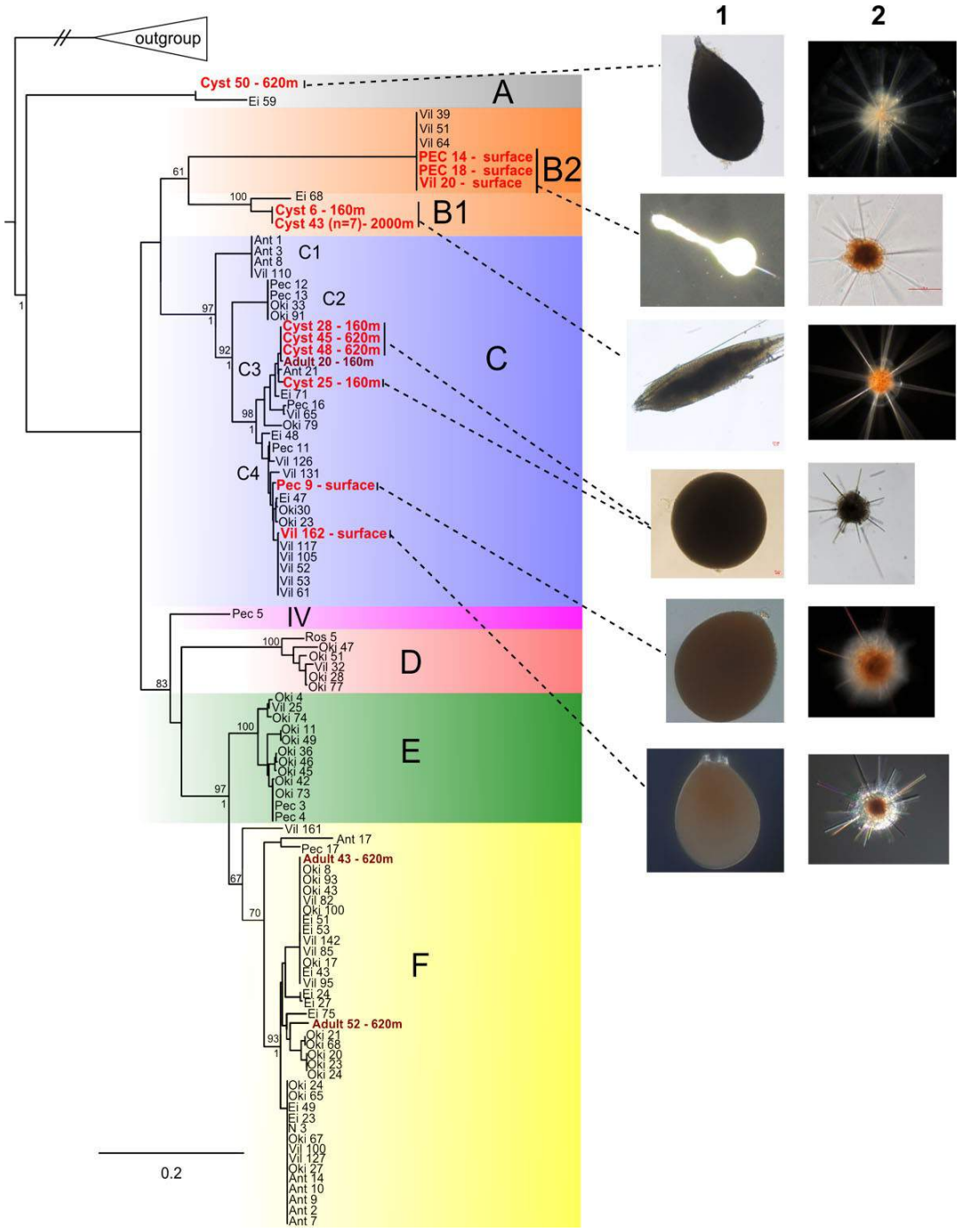


Figure 2: Phylogenetic placement of the cyst-forming Acantharia. Molecular phylogeny obtained with Maximum Likelihood analysis based on a concatenated matrix of 18S and 28S rDNA sequences of Acantharia (104 taxa and 2386 bp-long, GTR+I+G model), including cysts (bold red font) and vegetative (bold brown font) specimens collected in plankton nets and sediment traps at different depths. Each picture of a cyst (column 1) is shown together with the vegetative stage (column 2) that shares the highest genetic similarity. PhyML bootstrap percentages based on 500 pseudo-replicates and Bayesian posterior probabilities (PP) are indicated at each node (above and below, respectively) when support values were higher than 60% and 0.8, respectively.

Biogeography and vertical distribution of cyst-forming Acantharia

Clone libraries: We retrieved all the 18S rDNA sequences from GenBank related to Acantharia that were found in environmental clone libraries, and placed them with pplacer onto a reference phylogeny of Acantharia using the pplacer program (Figure S1). These environmental sequences were sampled between the surface and 3000 m in the Mediterranean Sea, and the Atlantic and Pacific Oceans, and belonged to different size fractions (total to < 3µm). We found 88 sequences between 0-100 m representing a wide genetic diversity of Acantharia (Figure 3A). Clades E and F accounted for about 50% of all 18S rDNA sequences, while the cyst forming ones accounted for 10%, 6% and 16% for A, B and C, respectively. Between 100-1000 m, which is essentially below the photic zone in most oceanic regions, clades B and C clearly dominated (more than 75% of the 114 acantharian sequences obtained at this depth), whereas the symbiotic clades E and F represented only 3% of sequences. Clade A contributed 3% between 100 and 1000 m, and clade I appeared in this zone with 10%. Clades I and III were named in a previous study [30] and are only represented by 18S rDNA environmental sequences (Figure S1). Between 1000 and 3000 m, the bathypelagic zone, clades B and C still represented a large proportion of acantharian sequences (60% out of 57 sequences), and clades I and A accounted for 21% and 9%, respectively. While no sequences of clades E were found in bathypelagic waters, clade F was represented by only one sequence in this zone. Thus, the cyst-forming clades A, B and C seemed to be present throughout the water column from the surface to 3000 m depth, but they clearly dominated below the photic zone compared to other clades. Clade D was also present from the surface to the bathypelagic zone.

Many environmental sequences from clone libraries were very close genetically to the cyst sequences retrieved in this study. For the round and elongated cysts from clade C and B1 respectively, more than 50 associated sequences were found between 200 and 1000 meters, some from the anoxic Cariaco Basin, Venezuela [41]. The environmental sequences that were related to the cysts of clade B2 were sampled in the hydrothermal vent sediment of the Guaymas Basin [42], and in a low-oxygenated layer at 1000 m depth in the Marmara Sea [18]. Remarkably, the large majority of environmental sequences related to clades A, B and C, retrieved in mesopelagic and bathypelagic waters were found in the picoeukaryote size fraction (< 3 µm). Clone libraries associated with Sanger sequencing generally represent the most abundant taxa in the environment because of the low sequencing depth (< 1000 clones). Generating much more reads, the new high-throughput sequencing methods can allow us to have now a more comprehensive view of the protistan community.

High-throughput environmental V9 tag sequences: Environmental tags (V9 region of the 18S rDNA gene) were obtained at 10 locations in the South Indian, Atlantic and Pacific Oceans, both the surface and in the mesopelagic zones. The sampling depth in the mesopelagic zone varied between stations from 177 to 800 m (Table 1). In total, 1,288,018 and 780,268 V9 sequences of Acantharia were retrieved from the surface and mesopelagic zone respectively, and assigned to different acantharian clades. As observed in the clone libraries analysis before, the acantharian

community in the surface was generally greatly diversified with a high contribution of symbiotic clades E and F at most stations (Figure 3B), even though at stations 65, 68, 72 and 76, cyst-forming clades A, B and C accounted for 44% to 60 % of acantharian sequences. In the Southern Ocean (station 85), sampled in December, the surface community differed greatly from other stations with clades III and A representing 40% and 26%, respectively.

In the mesopelagic zone, the acantharian community was relatively similar throughout the Indian, Atlantic and Pacific Oceans, but systematically very different from that found at the surface. In these deep waters, acantharian sequences contributed a significant proportion of total protist sequences obtained in this study: between 4.4 and 11 % in the Pacific and between 1.2 and 5.3 % in the Atlantic. Clade I was clearly predominant in the mesopelagic, representing 55 to 96% of the acantharian sequences, whereas it only accounted for ≤6% at the surface. In contrast, clades A, B and C were equally present at the surface and in the mesopelagic zone at most stations.

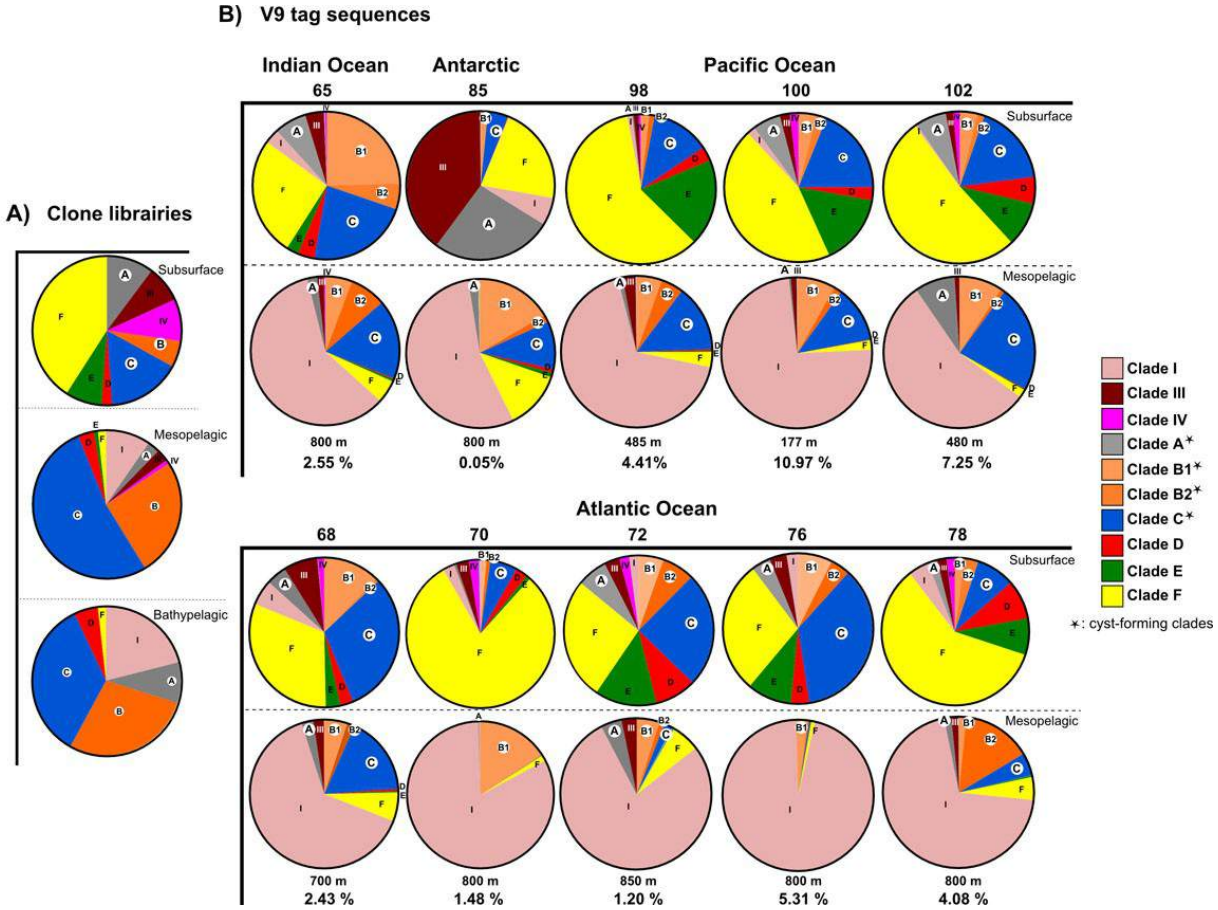


Figure 3: Vertical distribution of Acantharia based on 18S rDNA environmental sequences. A) clone libraries obtained from previous studies in different oceanic regions; B) V9 tag sequences from the present study sampled in the Indian (station 65), Antarctic (station 85), Pacific (stations 98 to 102) and Atlantic Oceans (stations 68 to 78). In the pie-charts, cyst-forming clades A, B and C are highlighted with a white circle, and clades representing <1% of the total acantharian sequences are not shown. For the mesopelagic zone, the sampling depth and the contribution of acantharian sequences to total protistan sequences are indicated below the pie-charts.

In order to gain a more accurate picture of the vertical distribution at the cyst-level, we focused on the environmental V9 sequences that were closely related to the cysts collected in this study (>97%). The end of the 18S rDNA marker (including the V9 region) was not PCR-amplified for the round (48, 45, 28 and 25) and pear-shaped cysts (cyst 50) because of the low PCR success for formaldehyde-preserved cysts. Still, we obtained the V9 region for the cysts Vil 162, Pec 9, Cyst 6 and Vil 20 from subclades C4, B1 and B2, respectively. The vertical distribution of V9 sequences that shared more than 97% with the cyst sequences is shown in Table S1. While no V9 sequences related to Vil 162 were found anywhere, V9 sequences assigned to elongated cysts (cyst 6) were more numerous in the mesopelagic zone and in some cases were not even detected in the photic zone (station 68). V9 sequences of Pec 9 and Vil 20 were either more abundant in the mesopelagic zone or in the photic layer depending on the station.

Strontium and carbon fluxes of Acantharia across the Atlantic Ocean

We assessed the temporal pattern of acantharian vertical flux at different latitudes in the Atlantic Ocean with sediment traps in the Scotia Sea and the subtropical North and South Atlantic (NOG and SOG), deployed for one or two years. Sr fluxes were highly episodic, with no more than a third of all samples on a trap containing elevated Sr and never in more than two consecutive samples (Figs 4, S2). Fluxes were very low in the subtropical traps, with only two samples in two separate years at each site containing significant fluxes: these were very low at the southern subtropical location (0.03 and 0.04 mg Sr m^{-2} d^{-1}), and higher at the northern subtropical site (0.13 and 0.23 mg Sr m^{-2} d^{-1}). Interestingly, these fluxes were all caught in boreal autumn: in November 2007 and September 2008 at the northern site, and in November 2008 and 2009 at the southern site, implying that there might be a seasonal cycle to Acantharian sedimentation in both locations.

Fluxes in the Scotia Sea were similarly low at the southern site, with 0.02 and 0.04 mg Sr m^{-2} d^{-1} in March and November 2008. At the northern site, however, Sr fluxes of 0.03–0.64 mg Sr m^{-2} d^{-1} were caught in February, March, October, and November 2008.

We did not directly measure the acantharian POC fluxes. However, the C:Sr ratio reported by Martin et al. [4] from the Iceland Basin at 2000 m (0.120 ± 0.022 mg mg^{-1}) would imply a contribution by Acantharia, during those collection periods in which Sr concentrations were elevated in the traps, of 0.004–0.003 mg C m^{-2} d^{-1} and 0.015–0.027 mg C m^{-2} d^{-1} at the southern and northern subtropical sites, respectively, and 0.005–0.008 mg C m^{-2} d^{-1} and 0.007–0.136 mg C m^{-2} d^{-1} at the southern and northern Scotia Sea site, respectively. Fluxes of total POC have not yet been measured from the subtropical traps, but are known for the Scotia Sea traps (G. Tarling unpubl.). We hence estimate a contribution by Acantharia of 0.6-3.1% to the total monthly POC flux.

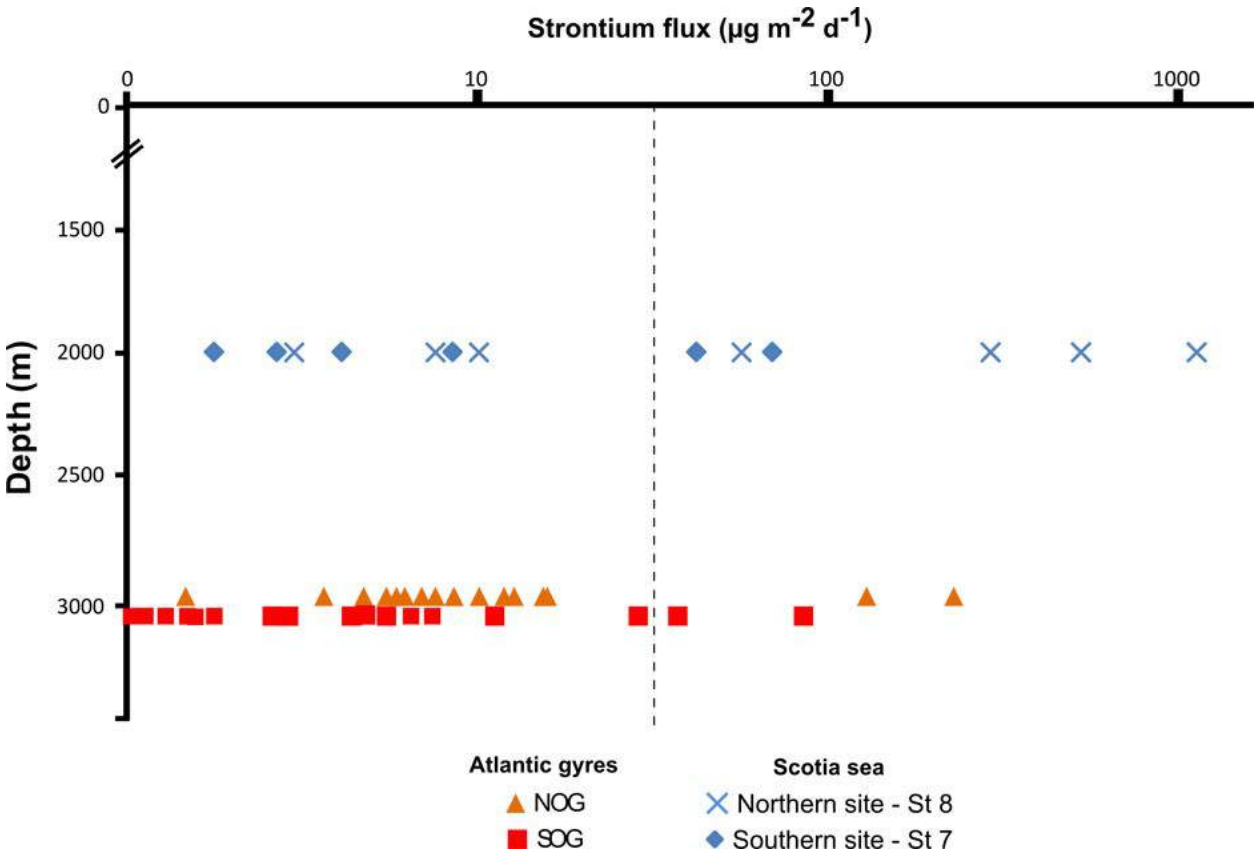


Figure 4: Geographic variation of the strontium flux. Strontium fluxes were measured in different sediment traps deployed across the Atlantic Ocean, in the Northern (NOG) and Southern (SOG) subtropical gyres at 3000 m (2007-2010) and in the Scotia Sea at 2000 m (St 7 and St 8; in 2008). Each value corresponds to a collection period of 14-31 days and is significant beyond the dashed line (the strontium concentration for each sample is shown in Fig. S2).

Discussion

Morphology of cysts and their vegetative stages

Encystment in Acantharia entails a remarkable morphological transformation whereby the diagnostic morphological features of the vegetative stage disappear. In this study, DNA identification allowed us to link cyst morphologies to vegetative cells of known taxonomy, and to highlight the cyst-forming acantharian clades. Cyst formation appears to be an old life-history trait in Acantharia since all cysts we sampled belonged to the early diverging clades A, B and C, representing the taxonomic orders Holacanthida and Chaunacanthida. Acantharia from these clades are characterized by identical simple (no lateral extensions, called apophyses) and long spicules with a loose central junction [30]. Some morphotypes of clade C (Chaunacanthida) can fold their spicules forming a new umbrella-like morphology (litholoph form), and the ones of clade B (Holacanthida) can totally shed their spicules. The cyst-forming Acantharia have thus a

flexible mineral skeleton during their vegetative stage, and dissociation of the central junction seems to be a required stage for encystment. In contrast, non-cyst-forming individuals from other clades (IV, D, E and F), mainly the orders Symphiacanthida and Arthracanthida, have a fixed and unfolded skeleton because of a very tight central junction and robust spicules with apophyses. It has been observed that some Symphiacanthida (e.g. *Amphibelone* sp.) are also able to form cysts, whereby the cyst shell is deposited upon the skeleton of the vegetative cell [9,13], but no molecular identification has confirmed the phylogenetic affiliation of these particular cysts. More cyst morphotypes have to be isolated and genetically identified to see whether other clades like I and III are also able to encyst and whether different encystment processes exist.

Diversity of Acantharia in the mesopelagic and bathypelagic zones

Planktonic protists hold key roles in the trophic chain and metabolism of the mesopelagic and bathypelagic zones but, essentially due to sampling difficulties, little is known about their diversity in these deep waters [1]. Cloning and sequencing analyses suggest that the diversity of protists, like prokaryotes, appears to decline in deep layers, where Radiolaria (including Acantharia), Alveolata and Euglenozoa have been reported to be the most abundant taxa [15,17,43]. Our study confirms that Acantharia, and more specifically clades I, A, B and C, can represent an important component of deep-sea protists worldwide, up to 11% of the environmental tag sequences. Acantharia from clade I, the morphology of which is totally unknown, are predominantly found in the dark ocean (up to 3000 m) and constitute up to 75% of the acantharian community in the mesopelagic zone. This clade appeared to be specific, possibly endemic, to deep waters but we cannot say whether it can undertake encystment for reproduction. For clades A, B and C, the DNA identification of cysts leads us to suggest that their presence in deep waters is probably linked to their encystment capacity.

Ecology and Life cycle of cyst-forming Acantharia

The vertical distribution pattern of the cyst-forming Acantharia revealed in this study can improve our understanding of the life cycle. The wide vertical distribution of clades A, B and C from the surface to the mesopelagic zone, detected by environmental clone libraries and V9 tag sequences (Figure 3), suggests that the vegetative stages occur in the upper layers and reproduction takes place in the dark ocean following encystment. Although Acantharia-derived detritus might be detected with our methods, since DNA is a stable molecule, we argue that the many sequences related to these clades found in deep waters in the picoeukaryote size fraction (< 3µm), much smaller than the size range of cysts (200-1000 µm), probably corresponds to swarmers shed by cysts. Despite the unknown ploidy of the swarmers, we presume that they fuse and form juveniles, which then ascend into the upper layers. Deep-sea bacterial and protistan communities are markedly different from the ones dwelling in the photic layer [15,17] but this study shows that the life cycle of particular microorganisms can lead to their occurrence throughout the water column. DNA identification of two stages of life cycle, the vegetative and cyst stage here, can help interpreting the numerous forthcoming molecular surveys investigating

the eukaryotic community in the photic zone and dark ocean [14,17,18,44,45], but also in sediment trap samples [46], ultimately linking water layers, biology, and geochemistry, through the ecology of organisms.

Remarkably, our phylogenetic analysis indicates that most cyst-forming Acantharia do not harbor symbiotic microalgae. The species *Acanthochiasma* sp. of clade B2 is an exception, since microalgae have been observed inside (pers. observation). Symbiotic microalgae have otherwise only been found in the more recently diverging clades E and F (Arthracanthida and Symphiacanthida), in which sexual reproduction takes place directly from the vegetative stage without encystment [8]. This obligate symbiosis would constrain the hosts to remain and complete their life cycle in the photic zone for rapid acquisition and maintenance of their photosynthetic symbionts. This is consistent with very few sequences from clades E and F being found in the mesopelagic and below in most oceanic regions (Figure 3). In contrast, non-symbiotic Acantharia could theoretically occupy any depth in the water column provided that food availability was sufficient. Besides, in contrast to the orders Arthracanthida and Symphiacanthida, the Holacanthida and Chaunacanthida have a less developed buoyancy apparatus with few myonemes (non-actin filaments) per spicules [8,47], suggesting that fine buoyancy control might be less critical for cyst-forming Acantharia. Of course, we cannot rule out the possibility that acantharian species within a single clade have different niches in the water column.

Distinct shallow and deep modes of reproduction have been also reported in planktonic Foraminifera, a sister group of the Radiolaria [48,49]. Likewise, foraminiferan species that have part of their life cycle in deep layers, below the photic layer, are not associated with symbiotic microalgae, indicating that photosymbiosis strongly influences the life cycle and consequently the vertical flux of hosts. A hypothetical scenario for the life cycle of cyst-forming and symbiotic Acantharia is proposed in Figure 5. Our study suggests that cyst-forming Acantharia do not contribute to primary production in the photic layer, a biogeochemical role specifically found in symbiotic forms [50], but do participate in the vertical particle flux in the oceans.

Geographic variation of the acantharian contribution to downward particle flux

Mineralized planktonic organisms, such as Foraminifera, coccolithophorids, diatoms and Radiolaria, are important contributors to downward particle flux in the ocean and thus play key roles in global element cycles [2,51,52]. Encystment in Acantharia is another example where changes in the life cycle of upper ocean organisms have a direct effect on downward particle flux. Acantharian fluxes in the bathypelagic were very low in the subtropical Atlantic, but higher in the Scotia Sea. Together with the data presented by Martin et al. [4,29], our results support the conclusion that Acantharia only sink past the upper mesopelagic in appreciable numbers at high latitude. Our estimates of acantharian POC fluxes at the northern Scotia Sea site are some of the highest reported from the bathypelagic, even though the percentage contribution to total POC flux was minor. In the Iceland Basin, acantharian cyst flux could contribute up to 48% and 59% of total bathypelagic POC and particulate nitrogen flux during the two weeks before spring

bloom phytodetritus was caught in the sediment traps [4]. In the East Greenland Sea, acantharian cysts comprised up to 90% of the POC sinking at 100 m during a short-term sampling period [19]. Thus, Acantharia can contribute to long-term carbon sequestration in the deep-sea. In subtropical regions, mainly oligotrophic, the vertical flux of acantharian cyst is lower and explains why modest values were detected at 3000 m depth in this study. In the Mediterranean Sea and Indian Ocean, cysts were generally reported only between 50 and 250 m [9,11].

Since cysts are more heavily mineralized and thus less susceptible to dissolution than vegetative cells, we suspect that cysts would have contributed the majority of the fluxes in our traps. We do not believe that artefacts may have significantly biased our regional sediment trap comparison. We can rule out a significant contribution of Sr by $CaCO_3$ dissolution, as Ca was never elevated by more than 300 µg mL^{-1} above samples that contained only background Sr concentrations. Since $CaCO_3$ typically contains no more than ~1 mg Sr g^{-1}, $CaCO_3$ is unlikely to have contributed more than 0.3 µg Sr mL^{-1}. While the Sr fluxes we report from the northern Scotia Sea are still lower than those reported from the Iceland Basin [4], they are nevertheless comparable to fluxes reported elsewhere above 500 m [3,24,53].

The difference in vertical flux of Acantharia between low and high latitudes is potentially due to larger cysts with higher sinking rates being found in high latitudes. Cysts of clade B1 collected in the Iceland Basin were up to 1 mm long and had sank of 490 ± 150 m d^{-1} [4], while cysts reported by Bernstein et al. [3] in the North Pacific Ocean were only a few hundred µm in diameter. Yet, cysts morphologically similar to the Iceland Basin cysts (clade B1) have been observed not only in the Southern [54] but also in the Indian Ocean [11] and the South China Sea [13], indicating that large cysts may not necessarily be restricted to cold waters in high-latitudes. Moreover, Antia et al. [19] reported small cysts with low sinking rates from the Greenland Sea, one of the highest latitude sites sampled for Acantharia.

A different possibility is that cyst-forming Acantharia in general might be more abundant at higher latitudes. We found that cyst-forming clades are widely distributed in surface waters of the Indian, Atlantic and Pacific Oceans but tend to be more abundant at the surface in nutrient-rich coastal areas (Figure 3): clades B and C in stations 65, 68 off South Africa, and clade A in Southern Ocean (station 85). It is interesting to note that station 85 was sampled in austral summer (December), a period matching with a significant strontium flux in the Scotia Sea (Figure S2). In addition, Popofsky [10] observed that the species *Acanthochiasma* sp. (a species found in clade B) becomes abundant during the summer in the Antarctic. In contrast, clades E and F that do not encyst are predominant in the oligotrophic open ocean. Higher latitudes might thus be home to larger numbers of cyst-forming species at a given time than subtropical regions, giving rise to higher downward fluxes of acantharian cysts. However, it is difficult to truly compare with certainty the acantharian communities of different oceanic regions because we could not address temporal variation at each sampling location.

Seasonal variation of acantharian cyst sedimentation

Long-term sediment trap deployments across a latitudinal gradient allowed us to follow the seasonal dynamics of the acantharian sinking flux. Strontium fluxes in both subtropical and Scotia Sea traps showed a seasonal pattern, as also observed in the Southern Ocean and Iceland Basin, where cysts were only occurring in summer (like in Scotia Sea) and spring (April-May), respectively [4,54]. For traps deployed in the Scotia Sea, the months December and January were not sampled, though it is possible that strontium flux was significant at this time, given the temporal pattern (Fig S3). Cyst formation and sedimentation are hence presumably driven by environmental cues. The strong difference in Sr flux between the two Scotia Sea sites suggests that acantharian fluxes are indeed linked to phytoplankton blooms, as hypothesized by Martin et al. [4]: the northern Scotia Sea site experienced intense phytoplankton blooms starting in October/November, and the northern sediment trap showed peaks in particulate carbon flux in October and November 2008, rising from around 50 mg C m^{-2} d^{-1} in September to >400 mg C m^{-2} d^{-1} in October and November [55]. The southern trap shows no such sharp peak, and instead had the highest carbon fluxes in March and May. Thus, formation and deep sinking of cysts leading to reproduction in Acantharia does appear to be linked to primary productivity. This mirrors the periodicity of encystment in marine ciliates in summer [56], and that in a freshwater ciliate in May and September [57], which are both related to chlorophyll-*a* concentration. We therefore suggest that cyst-forming Acantharian show a clear seasonality in high-latitude waters that tracks phytoplankton stocks, and then complete their life cycle by sinking deep into the interior of the ocean.

Hypotheses on the ecological role of encystment in Acantharia

Encystment in Acantharia is not a reversible resting stage like in many planktonic protists, but rather a key step for reproduction. Cyst formation is clearly an active process, and evidently an efficient mechanism for rapid sinking. While we do not know anything specific about the bioenergetics of celestite precipitation, it is possible that the extensive biomineralisation and remodelling of the cyst shell is not a trivial energy expense. To our knowledge, such a reproductive strategy has not been reported for other marine protists. The robust and closed seed-like cysts shells not only act as ballast but might also protect cysts from grazing during descent. Although it is difficult to assert the true ecological role of this life-history trait, sinking in the water column for reproduction must have an adaptive value in heterotrophic protists. In stratified waters, rapid sinking might allow individuals to quickly reach depths at which predation risk is reduced. In phytoplankton, vertical migration for nutrient uptake is a common strategy to increase nutrient uptake when surface waters are nutrient-depleted (e.g. the diatom *Rhizosolenia* sp., [58,59]). However, for heterotrophs like cyst-forming Acantharia, sinking into the dark ocean where food concentrations are lower than in surface waters may entail trade-offs. While we know little about their diet, Acantharia are known to ingest a wide range of prey, such as bacteria, microalgae and even zooplankton [60]. It is interesting to note in this context that Schewiakoff ([8] p. 686) reported that 73% of Acantharia in which he could see food vacuoles

were caught below 100 m depth. We hence suggest that cyst-forming Acantharia at high latitudes and during phytoplankton blooms, when downward particle flux shows sudden spikes, sink particularly deep because juveniles could still exploit the fresh phytodetritus that sinks right to the seafloor at this time. Conversely, in low-latitude settings in which production and downward particle flux are far lower and more constant, sinking very deep may be less advantageous as far less fresh particulate matter sinks to great depths here. The regional and temporal flux data we have presented here are consistent with this view.

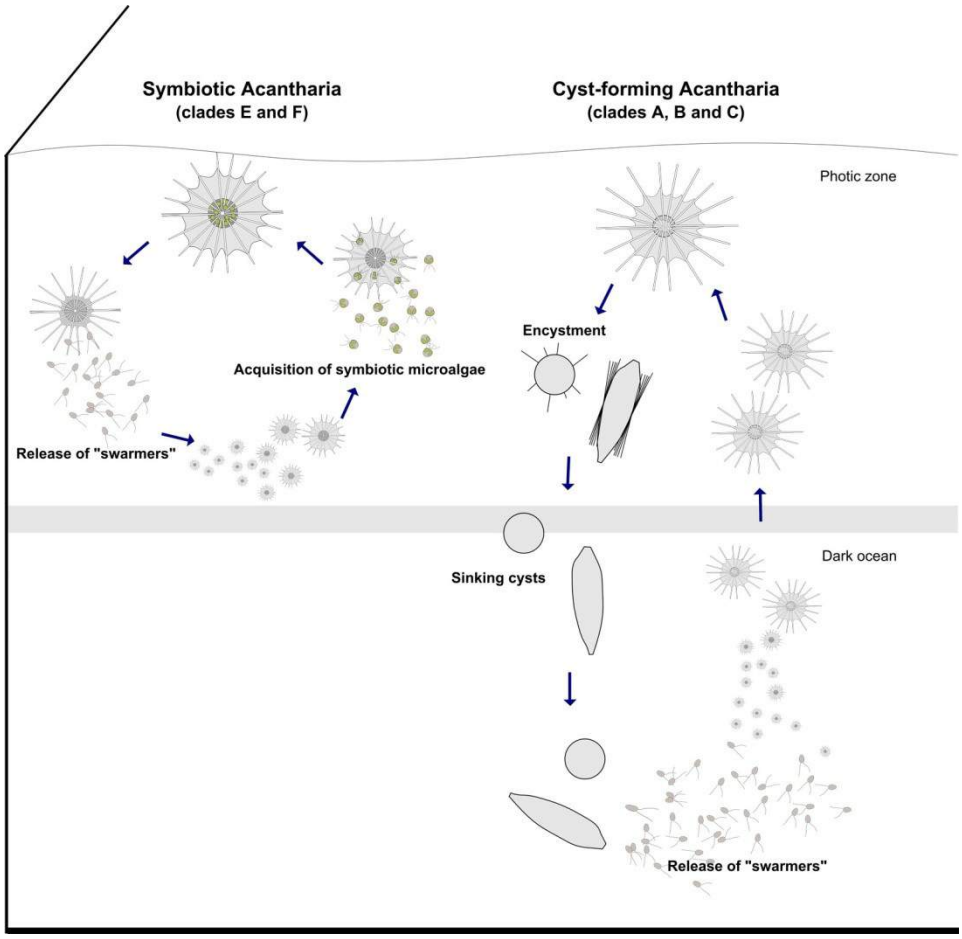


Figure 5: Hypothetical scenario of the life cycle in symbiotic and cyst-forming Acantharia with shallow and deep reproduction, respectively.

Acknowledgments

This work was supported by the PhD fellowship DIPHOPE and SYMFORAD project from the Region Bretagne. We thank S. Romac for technical assistance and J. Dolan from the LOV in Villefranche-sur-Mer for providing sampling facilities. From the Station Biologique de Roscoff, we thank Morgan Perennou and Gwen Tanguy from the GENOMER platform. We are grateful to the TaraOceans expedition (doi:10.1371/journal.pbio.1001177), which supported the collection of some samples. We also thank Matthew Cooper for help with Sr measurements and Corinne Pebody for help with sediment trap deployments.

References

1. Aristegui J, Gasol JM, Duarte CM, Herndl GJ (2009) Microbial oceanography of the dark ocean's pelagic realm. Limnol Oceanogr 54(5): 1501-1529.

2. Lampitt RS, Salter I, John D (2009) Radiolaria: major exporters of organic carbon to the deep ocean. Global Biogeochem Cycles 23 GB1010, doi:10.1029/2008GB003221.

3. Bernstein RE, Feely RA, Byrne RH, Lamb MF, Michaels AF (1987) Acantharian fluxes and strontium to chlorinity ratios in the north Pacific Ocean. Science 237: 1490-1494.

4. Martin P, Allen JT, Cooper MJ, *et al.* (2010) Sedimentation of acantharian cysts in the Iceland Basin: Strontium as a ballast for deep ocean particle flux, and implications for acantharian reproductive strategies. Limnol Oceanogr 55: 604-614.

5. Smetacek VS (1985) Role of sinking in diatom life-history cycles: ecological, evolutionary and geological significance. Mar Biol 84: 239-251.

6. Kremp A, Rengefors F, Montresor M (2009) Species specific encystment patterns in three Baltic cold-water dinoflagellates: the role of multiple cues in resting cyst formation. Limnol Oceanogr 54(4): 1125-1138.

7. Von Dassow P, Montresor M (2011) Unveiling the mysteries of phytoplankton life cycles: patterns and opportunities behind complexity. J. Plankton Res. 33(1): 3-12.

8. Schewiakoff WT (1926) The Acantharia. Fauna e Flora del Golfo di Napoli 37: 1-755.

9. Hollande A, Cachon J, Cachon-Enjumet M (1965) Les modalités de l'enkystement présporogénétique chez les acanthaires. Protistologica 1: 91-112 [In French].

10. Popofsky A (1908) Die Radiolarien der Antarktis (mit ausnahme der tripyleen) Deutsche Südpolar-Expedition 1901-1903, Zoologie II 10(3): 183-305 [In German].

11. Bottazzi EM (1973) Ulteriori ritrovamenti di cisti di Acantari (protozoa). Mt Lomb Sci Lett 107B: 3-26.

12. Haeckel E (1888) Die Radiolarien (Rhizopoda Radiolaria). Eine Monographie, Dritter Teil, Die Acantharien oder Actinopyleen-Radiolarien. Verlag von Georg Reimer, Berlin, pp 33.

13. Zhiyuan T, Xinghui S (189) Studies on the acantharian cysts of the south china sea? Studia marina sinica 30: 127-142.

14. Gilg IC, Amaral-Zettler LA, Countway PD, Moorthi S, Schnetzer A, Caron D (2010) Phylogenetic affiliations of mesopelagic acantharia and acantharian-like environmental 18S rRNA genes off the Southern California coast. Protist 161: 197-211.

15. Not F, Gausling R, Azam F, Heidelberg JF, Worden AZ (2007) Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. Environ Microbiol 9(5): 1233-1252.

16. Schnetzer A, Moorthi SD, Countway PD, Gast RJ, Gilg IC, Caron D (2011) Depth matters: Microbial eukaryote diversity and community structure in the eastern North Pacific revealed through environmental gene libraries. Deep-Sea Res I 58: 16-26.

17. Countway PD, Caron DA, Gast RJ, Savai P (2007) Comparison of protistan diversity in deep (2500m) vs euphotic zone assemblages in the Sargasso Sea and Gulf Stream (N.Atlantic). Environ Microbiol 9: 1219-1232.

18. Quaiser A, Zivanovic Y, Moreira D, López-García P (2010) Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. ISME 5: 285-304.

19. Antia AN, Bauerfeind E, Bodungen Bv, Zeller U (1993) Abundance, encystment and sedimentation of acantharia during autumn 1990 in the East Greenland Sea. J Plankton Res 15: 99-114.

20. Hollande A, Cachon-Enjumet M (1957) Enkystement et reproduction isosporogénétique chez les acanthaires. CR Acad Sci 244: 508-510 [In French].

21. Kwon EY, Primeau F, Sarmiento JL (2009) The impact of remineralization depth on the air-sea carbon balance. Nat Geosci 2: 630-625.

22. Beers JR, Stewart GL (1970) The preservation of acantharians in fixed plankton samples. Limnol Oceanogr 15: 825-827.

23. Michaels AF, Caron DA, Swanberg NR, Howse FA, Michaels CM (1995) Planktonic sarcodines (Acantharia, Radiolaria, Foraminifera) in surface waters near Bermuda – abundance, biomass and vertical flux. J Plankton Res 17: 131-163.

24. Bernstein RE, Byrne RH, Betzer PR, Greco AM (1992) Morphologies and transformations of celestite in seawater: the role of acantharians in strontium and barium geochemistry. Geochim Cosmochim Acta 56: 3273-3279.

25. De Villiers S (1999) Seawater strontium and Sr/Ca variability in the Atlantic and Pacific oceans. Earth Planet Sci Lett 171: 623-634.

26. De Dekker P (2004) On the celestite-secreting Acantharia and their effect on seawater strontium to calcium ratios. Hydrobiologia 517: 1-13.

27. Bidwell JP, Paige JA, Kuzirian AM (1986) Effects of strontium on the embryonic development of *Aplysia californica*. Biological Bulletin (Woods Hole, MA, United States) 170(1): 75-901.

28. Hanlon RT, Bidwell JP, Tait R (1989) Strontium is required for statolith development and thus normal swimming behavior of hatchling cephalopods. J Exp Biol 141: 187-95.

29. Martin P, Lampitt RS, Perry MJ, Sanders R, Lee C, D'Asaro E (2011) Export and mesopelagic particle flux during a North Atlantic spring diatom bloom. Deep-Sea Res I 58: 338-349.

30. Decelle J, Suzuki N, Mahé F, de Vargas C, Not F (2012) Molecular phylogeny and morphological evolution of the Acantharia (Radiolaria) Protist 163: 435-450.

31. Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27(2): 221-224.

32. Vaidya G, Lohman DJ, Meier R SequenceMatrix: concatenation software for the fast assembly of multigene datasets with character set and codon information. Cladistics, accepted.

33. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 28: 2731-2739.

34. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 694-704.

35. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.

36. Matsen FA, Kodner RB, Armbrust EV (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinformatics 11: 538.

37. Karsenti E, Acinas SG, Bork P, Bowler C, De Vargas C, et al. (2011) A Holistic Approach to Marine Eco-Systems Biology. PLoS Biol 9(10): e1001177. doi:10.1371/journal.pbio.1001177.

38. Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS ONE 4(7): e6372. doi:10.1371/journal.pone.0006372.

39. Needleman SB, Wunsch, CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of Molecular Biology 48 (3): 443-453.

40. Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85: 2444-2448.

41. Edgcomb VP, Orsi W et al (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. ISME 1-13.

42. Edgcomb VP, Kysela DT, Teske A, Gomez AD, Sogin ML (2002) Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. Proc Natl Acad Sci USA 99: 7658-7662.

43. Shah Salani F, Arndt H, Hausmann K, Nitsche F, Scheckenbach F (2012) Analysis of the community structure of abyssal kinetoplastids revealed similar communities at larger spatial scales. ISME 6: 713-723.

44. López-García P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature 409: 603-607.

45. López-García P, Philippe H, Gail F, Moreira D (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the mid-Atlantic ridge. Proc Natl Acad Sci USA 100: 697-702.

46. Amacher J, Neuer S, Anderson I, Massana R (2009) Molecular approach to determine contributions of the protist community to particle flux. Deep-Sea Res I 56: 2206-2215.

47. Febvre J, Febvre-Chevalier C (2001) Acantharia. In: eLS. John Wiley & Sons Ltd, Chichester. http://www.els.net [doi: 10.1038/npg.els.0002102.

48. Schiebel R, Hemleben C (2005) Modern planktic foraminifera. Paläontologische Zeitschrift 79: 135-148.

49. Shaked Y, de Vargas C (2006) Pelagic photosymbiosis: rDNA assessment of diversity and evolution of dinoflagellate symbionts and planktonic foraminiferal hosts. Mar Ecol Prog Ser 325: 59-71.

50. Michaels AF (1988) Vertical distribution and abundance of Acantharia and their symbionts. Marine Biology 97: 559-569.

51. Salter I, Kemp AES, Moore CM, Lampitt RS, Wolff GA, Holtvoeth J (2012) Diatom resting spore ecology drives enhanced carbon export from a naturally iron-fertilized bloom in the Southern Ocean. Global Biogeochem Cycles 26: GB1014, doi:10.1029/2010GB003977.

52. Schiebel R (2002) Planktic foraminiferal sedimentation and the marine calcite budget. Global Biogeochem Cycles 16: 1065.

53. Lamborg CH et al (2008) The flux of bio- and lithogenic material associated with sinking particles in the mesopelagic ‘‘twilight zone’’ of the northwest and North Central Pacific Ocean. Deep-Sea Res II 55: 1540-1563.

54. Spindler M, Beyer K (1990) Distribution, abundance and diversity of Antarctic acantharian cysts. Mar Micropal 15: 209-218.

55. Whitehouse MJ, Atkinson A, Korb RE, Venables HJ, Pond DW, Gordon M (2012) Substantial primary production in the land-remote region of the central and northern Scotia Sea. Deep-Sea Res II 59-60: 47-56.

56. Reid PC (1987) Mass encystment of a planktonic oligotrich ciliate. Mar Biol 95: 221-230.

57. Müller H, Wünsch C (1999) Seasonal dynamics of cyst formation of pelagic strombidiid ciliates in a deep prealpine lake. Aquat Microb Ecol 17: 37-47.

58. Richardson TL, Cullen JJ, Kelley DE, Lewis MR (1998) Potential contributions of vertically migrating *Rhizosolenia* to nutrient cycling and new production in the open ocean. J Plankton Res 20(2): 219-241.

59. Villareal TA, Altabet MA, Culver-Rymsza K (1993) Nitrogen transport by vertically migrating diatom mats in the North Pacific Ocean. Nature 363: 709-712.

60. Swanberg NR, Caron DA (1991) Patterns of sarcodine feeding in epipelagic oceanic plankton. J Plankton Res 13(2): 287-312.

Table 1: List of sampling sites and equipment used.

Oceanic region	Sampling sites	Latitude	Longitude	Date	Sampling	Depth (m)
Mediterranean Sea	Villefranche-sur-mer	43°40.552 N	7°18.447 E	Nov-10	Net samples	0
North Atlantic Ocean	Iceland basin	60°N	20°W	Nov 06 - July 07	Trap samples	2000
North Atlantic Ocean	Iceland basin	61°N	26°W	May-08	Trap samples	160-620
North Atlantic Ocean	NOG	23° 46.119 N	41° 5.419 W	Nov 07 - Nov 09	Trap samples	3000
South Atlantic Ocean	SOG	18° 31.48 S	25° 6.0 W	May 08 - June 10	Trap samples	3000
South Atlantic Ocean	St 7	55° 15.320 S	41° 21'.273 W	Mar 08 - Dec 08	Trap samples	2000
South Atlantic Ocean	St 8	52° 48.288 S	40° 6.522 W	Feb 08 - Dec 08	Trap samples	2000
Indian Ocean	Station 65	35°45.299 S	26°19.199 E	Jul-10	Niskin bottles	0 and 800
South Atlantic Ocean	Station 68	31°58.416 S	5°21.612 E	Sep-10	Niskin bottles	0 and 700
South Atlantic Ocean	Station 70	18°4.229 S	4°54.82 W	Sep-10	Niskin bottles	0 and 800
South Atlantic Ocean	Station 72	8°13.415 S	17°5.202 W	Oct-10	Niskin bottles	0 and 850
South Atlantic Ocean	Station 76	20°0.78 S	35°43.296 W	Oct-10	Niskin bottles	0 and 800
South Atlantic Ocean	Station 78	30°47.629 S	43°41.878 W	Nov-10	Niskin bottles	0 and 800
Antarctic	Station 85	62°51.41 S	49°41.817 W	Jan-11	Niskin bottles	0 and 800
South Pacific Ocean	Station 98	25°6.431 S	111°20.25 W	Apr-11	Niskin bottles	0 and 485
South Pacific Ocean	Station 100	13°59.394 S	96°51.343 W	Apr-11	Niskin bottles	0 and 177
South Pacific Ocean	Station 102	5°44.135 S	85°46.669 W	Apr-11	Niskin bottles	0 and 480

Table 2: Information about the cysts collected in this study and the GenBank accession numbers of their associated 18S and 28s rDNA.

Code name	Form	Molecular clade	Taxonomic order	Sampling site	Depth (m)	Accession number	
						18S	28S
Cyst 6	elongated	B1	Holacanthida	Iceland Basin	160	XXXX	XXX
Cyst 25	round	C	Chauncanthida	Iceland Basin	160	XXXX	XXX
Cyst 28	round	C	Chauncanthida	Iceland Basin	160	XXXX	XXX
Cyst 43	elongated	B1	Holacanthida	Iceland Basin	2000	XXXX	XXX
Cyst 45	round	C	Chauncanthida	Iceland Basin	620	XXXX	XXX
Cyst 48	round	C	Chauncanthida	Iceland Basin	620	XXXX	XXX
Cyst 50	pear-shaped	A	Holacanthida	Iceland Basin	620		XXX
Pec 9	oval	C	Chauncanthida	Mediterranean Sea	surface	XXXX	XXX
Pec 14	round	B2	Holacanthida	Mediterranean Sea	surface	XXXX	XXX
Pec 18	round	B2	Holacanthida	Mediterranean Sea	surface	XXXX	XXX
Vil 162	oval	C	Chauncanthida	Mediterranean Sea	surface	XXXX	XXX

Supplementary material

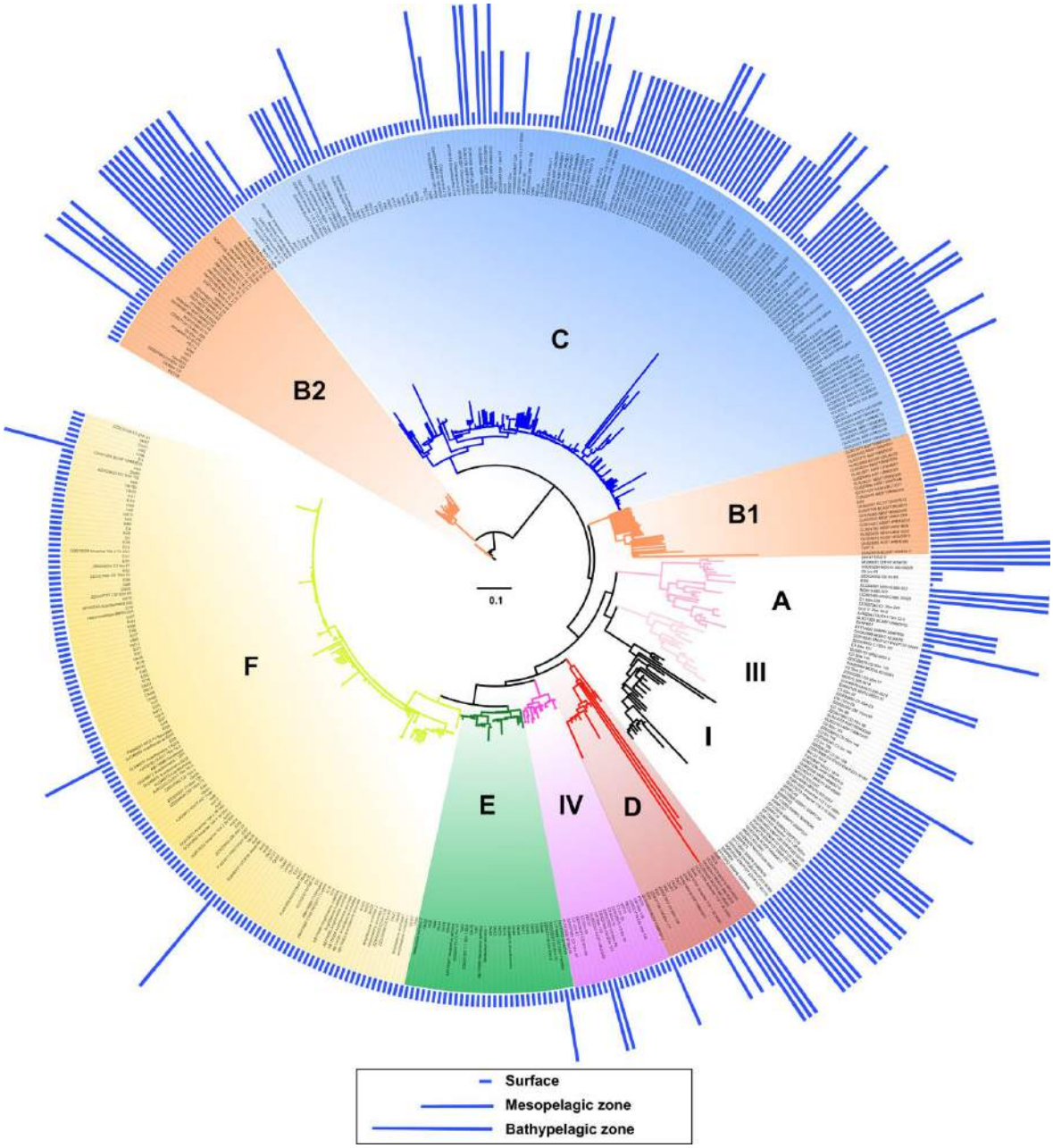


Figure S1: Diversity of acantharian sequences found in deep waters. Pplacer tree showing the phylogenetic placement of 260 environmental 18s rDNA sequences sampled in previous studies into reference clades of Acantharia (I, III, IV, A-F). The depth at which the acantharian sequences were collected is indicated with the blue bars.

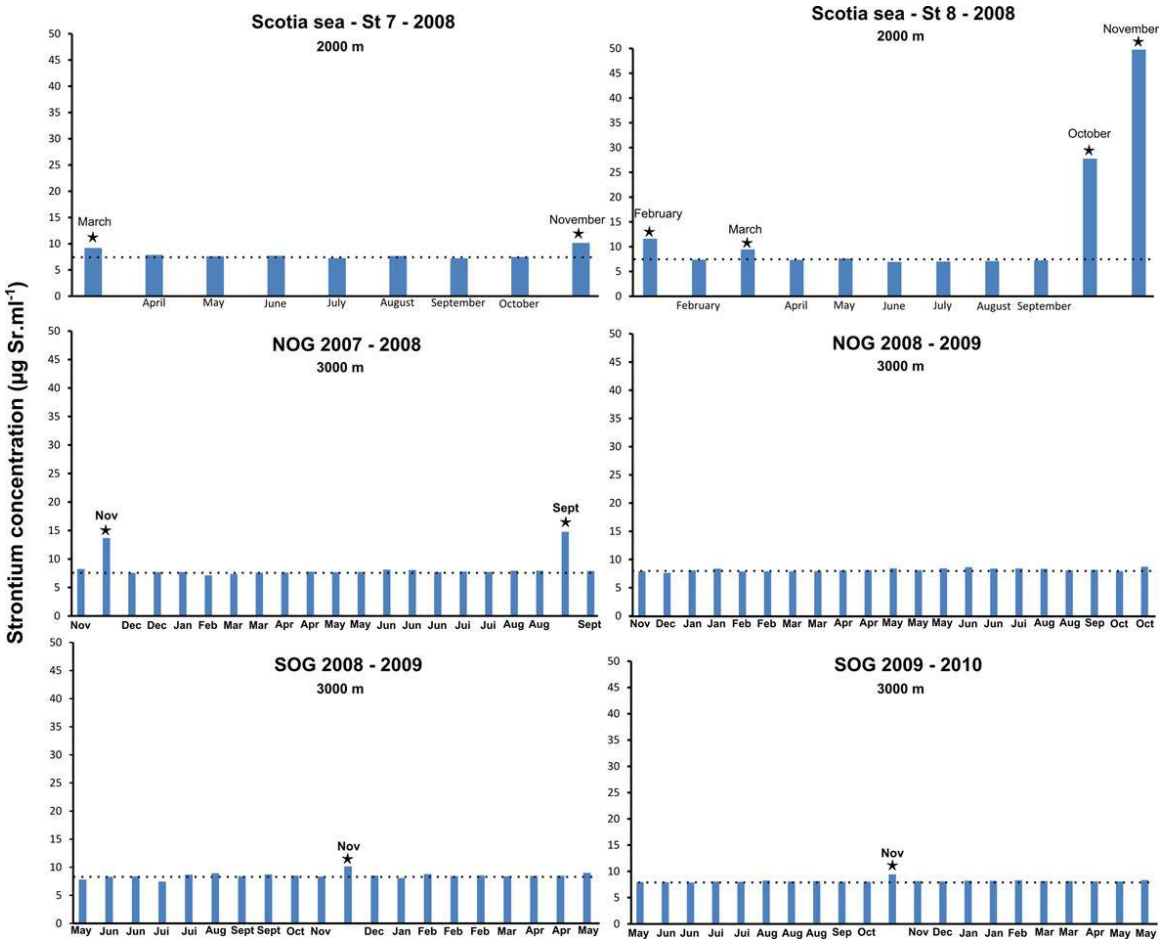


Figure S2: **Temporal pattern of strontium flux across the Atlantic Ocean.** Concentration of dissolved strontium measured in each sample cup of the four bathypelagic sediment traps, which were deployed at 2000 m in the Scotia Sea (St 7 and St 8; 2008) and at 3000 m in the Northern (NOG) and Southern (SOG) subtropical gyres (2007-2010). The horizontal dashed line represents the background strontium concentration above which the Sr flux from Acantharia was deemed significant, as highlighted by a black star.

Table S1: Comparison of the number of V9 tag sequences assigned to different cysts (> 97%) isolated in this study between the photic and the mesopelagic zones.

Depth	Clade	Cyst	65	68	70	72	76	78	85	98	100	102
Photic zone	C4	PEC9	535	2352	134	1335	3146	259	49	213	1101	2241
Mesopelagic	C4	PEC9	1001	587	10	64	47	294	20	81	1597	1374
Photic zone	B1	Cyst 6	35	0	10	10	6	73	172	8	17	52
Mesopelagic	B1	Cyst 6	309	738	7148	122	176	462	163	575	20	290
Photic zone	B2	Vil 20	281	5	6	3440	1322	614	1	17	252	350
Mesopelagic	B2	Vil 20	1203	496	31	68	37	236	7	2804	91	165

Discussion et perspectives

La photosymbiose a été probablement une des plus grandes innovations dans l'évolution des eucaryotes donnant naissance à tous les organismes photosynthétiques sur terre. Ce type d'interaction est également une composante essentielle pour le fonctionnement des écosystèmes pélagiques actuels et constitue un moteur pour l'évolution du plancton. Les différents chapitres de ce travail de thèse ont permis de mettre en évidence l'influence de la photosymbiose sur l'évolution, l'écologie, le cycle de vie, et le rôle biogéochimique des acanthaires, un groupe de protistes planctoniques abondant et cosmopolite dans les océans. Ces travaux ont également contribué à l'élaboration d'hypothèses plus générales quant au possible fonctionnement des photosymbioses en milieu pélagique.

1. Vers une nouvelle taxonomie intégrative des acanthaires

Afin de positionner la symbiose dans un contexte taxonomique et évolutif, l'histoire évolutive des acanthaires, inconnue jusqu'alors, a été mise en évidence par des analyses phylogénétiques à partir de séquences ribosomales d'individus isolés et identifiés sur des critères morphologiques (chapitre I). De nombreux conflits entre cette phylogénie et la taxonomie sont apparus, suggérant la nécessité d'une révision taxonomique des acanthaires. Par exemple, l'ordre Holacanthida n'apparait pas comme monophylétique et est présent dans différents clades phylogénétiquement éloignés (clades A, B et D). Peu d'acanthaires appartenant à l'ordre des Symphiacanthida ont été échantillonnés dans cette étude mais leur position dans le clade F parmi les Arthracanthida remet en doute son existence. A un niveau taxonomique inférieur, certains genres et familles sont polyphylétiques. Clairement, cette phylogénie n'est pas une classification en elle-même mais fournit un support indéniable pour redéfinir la pertinence de certains critères diagnostiques synapomorphiques, comme la jonction centrale. Ce type d'analyse est d'autant plus profitable dans le cas des acanthaires et autres protistes incultivables, où l'ontogénèse continue de certains critères est très importante et encore incomprise (Fig. 1). Le squelette des acanthaires ne cesse de croître et de se modifier au cours du cycle de vie d'un individu et il est particulièrement difficile de statuer sur un phénotype final pour chaque espèce. La morphologie du squelette semble néanmoins porter un signal phylogénétique fort comme chez de nombreux protistes biominéralisants: frustules de diatomées (Medlin et Kaczmarska, 2004), la lorica des tintinnides (Dolan, 2012; Bachy et al., 2012). Couplées aux analyses phylogénétiques, de nouvelles techniques d'observation (ex : microscopie électronique à balayage et à épifluorescence) doivent être développées pour définir plus de caractères morphologiques sur le squelette et certains composants cellulaires de chaque cellule (myonèmes, axopodes etc.). La multiplication des

caractères morphologiques observés sur une cellule améliorera l'identification et permettra également d'envisager une révision complète de la taxonomie des acanthaires.

Sur les 150 espèces d'acanthaires et 18 familles du schéma taxonomique actuel, environ 75 espèces et 15 familles ont été échantillonnées dans cette thèse, ce qui donne relativement une bonne couverture et vision de la diversité génétique et morphologique des acanthaires. Certaines des cellules collectées lors de cette thèse n'ont pu être identifiées à partir des taxonomies existantes d'Haeckel (1888) et de Schewiakoff (1926) et constituent très probablement de nouvelles espèces voire de nouveaux ordres à définir (ex: clades D1 et IV). Il est important de souligner que la taxonomie des acanthaires, qui n'a pas évolué depuis plus d'un siècle, contient très probablement un nombre surestimé d'espèces. Certaines sont par exemple décrites qu'à partir d'un seul et même spicule et la morphologie globale de la cellule est totalement inconnue (ex: *Acanthochiasma quadrangulum*). Ces espèces définies sur une description extrêmement superficielle ne devraient plus être considérées dans la taxonomie des acanthaires. De plus, une attention particulière doit également se porter sur le nombre important d'espèces synonymes : plusieurs noms d'espèces ont été donnés par Müller, Haeckel, Schewiakoff sur un même morphotype. Une clarification et une révision de la taxonomie sur une base morpho-moléculaire, en accord avec le code de nomenclature zoologique (ICZN), est actuellement en cours d'élaboration en collaboration avec le Dr Noritoshi Suzuki, seul expert au monde de la taxonomie des Acanthaires !

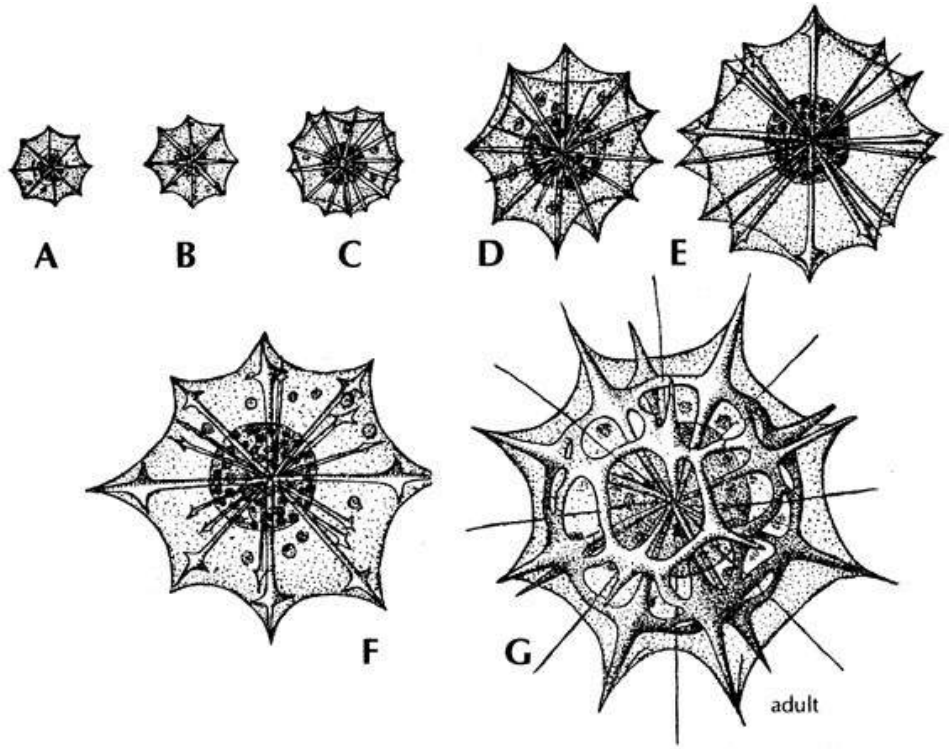


Figure 1. Exemple d'ontogénèse du squelette des acanthaires (cas de *Pleuraspis costata*) où les caractères morphologiques utilisés pour l'identification et la classification de l'espèce peuvent changer drastiquement au cours du cycle de vie (d'après Reshetnjak, 1981 et Schewiakoff, 1926).

Depuis l'écriture du chapitre I, l'ADNr 5.8S et les espaces intergéniques (ITS1 et ITS2) de l'opéron ribosomal, généralement plus résolutifs que l'ADNr 18S et l'ADNr 28S chez plusieurs lignées eucaryotes (Coleman, 2007), ont été séquencés sur plusieurs individus de chaque clade. Une phylogénie concaténée a pu être construite en maximum de vraissemblance à partir de tout l'opéron ribosomal soit 3654 positions nucléotidiques (Fig. 2). L'apport de l'ITS dans le jeu de données a confirmé et amélioré les relations phylogénétiques de la plupart des clades d'acanthaires. La position du clade C et B2 reste toutefois encore incertaine (BV < 0.6). A l'avenir, un gène codant pour une protéine comme l'actine pourrait donner une vision alternative sur la phylogénie des acanthaires et conforter leur histoire évolutive (ex : foraminifères; Flakowski et al., 2005).

Parallèlement, la poursuite de l'échantillonnage et l'isolement de nouvelles cellules d'acanthaires a permis de caractériser le clade A (clade II dans le chapitre I) et le clade IV. Actuellement, seuls les clades I et III sont composés uniquement de séquences environnementales dont le phénotype reste inconnu. Une base de données de référence morpho-génétique, comprenant environ 150 séquences du gène de l'ADNr 18S et 200 séquences du gène de l'ADNr 28S avec la morphologie associée, a été construite durant cette thèse. Cette base fut un outil crucial pour assigner toutes les séquences environnementales générées, clones et marqueurs V9 de l'ADNr 18S (chapitre I, III, IV). La couverture de l'échantillonnage taxonomique n'est évidemment pas exhaustive mais le pourcentage d'assignation des séquences environnementales indique qu'il n'existerait pas d'autres nouveaux clades, phylogénétiquement très éloignés, et que la plupart de la diversité génétique a été mise en évidence. Par exemple, pour les clades C, D, E et F, les séquences environnementales (V9) sont pour la grande majorité identiques ou très proches des séquences de références (100 ou 98%; figures dans *supplementary information*, chapitre III). Cependant, un effort particulier d'échantillonnage de cellules devra cibler les clades basaux, tels que les clades I, III, A et B.

Enfin, l'acquisition de différents gènes ribosomaux a permis de comparer leur taux d'évolution et donc leur pouvoir résolutif. Par rapport à la région V4 (370 pb) de l'ADNr 18S, la région D2 (228 pb) de l'ADNr 28S évolue environ 2,5 fois plus vite, et constitue un « *barcode* » moléculaire intéressant pour les acanthaires comme c'est le cas chez d'autres protistes (ciliés, Santoferrara et al., 2012 et diatomées, Hamsher et al., 2011). Cette étude a participé au projet international de barcodes des protistes mené par le groupe ProWG dans le consortium de Barcode of Life (Pawlowski et al., 2012 ; annexe 4), et la base de référence des acanthaires a été intégrée dans une base de données plus large représentant tous les protistes: (*The Protist Ribosomal Reference Database,* Guillou et al., 2012 annexe 4).

Des planches taxonomiques de chacun des clades d'acanthaires sont présentées en Annexe 2

2. Hypothèses sur l'histoire évolutive des acanthaires: de la morphologie à l'écologie

Les conflits taxonomiques et l'absence totale de traces fossiles au cours des temps géologiques n'ont pas permis jusqu'ici de mettre en évidence l'évolution des acanthaires. Dans cette étude, la combinaison de données issues de la phylogénie morpho-génétique (chapitre I), de l'écologie (chapitre III) et du cycle de vie (IV) permet de proposer une hypothèse sur l'histoire évolutive et écologique des acanthaires. L'approche à l'échelle d'une cellule unique a permis de faire un lien direct entre la morphologie et le mode de vie comme la symbiose.

2.1 Du simple cristal au squelette

L'utilisation du sulfate de strontium chez les acanthaires reste une énigme. Comment et pourquoi seuls les acanthaires font des squelettes avec ce minéral ? Ce travail de cette thèse n'a pas eu pour but de répondre à ces questions mais permet d'émettre une hypothèse sur l'origine évolutive de la biominéralisation du sulfate de strontium chez les acanthaires.

Des monocristaux de sulfate de strontium ont été observés dans les «*swarmers*», petites cellules flagellées qui correspondraient aux gamètes, chez d'autres radiolaires qui utilisent la silice pour fabriquer leur squelette (Hollande et Martoja, 1974; Anderson et al., 1990; Perry et Hugues, 1990). Ces cristaux de 50 à 100 nm disparaissent ensuite au cours du cycle de vie des cellules. Leur rôle serait possiblement lié à la détection de la gravité ou à la flottabilité de la cellule (Anderson et al., 1990). Bien que la présence de ces cristaux dans les «*swarmers*» d'acanthaires reste à démontrer, leur observation chez des groupes frères soulève des questions sur l'histoire évolutive de la biominéralisation du strontium. On peut émettre l'hypothèse que ces monocristaux sont un caractère ancestral de tous les radiolaires et qu'ils ont été gardés au cours des stades ultérieures du cycle de vie chez les acanthaires pour constituer un squelette minéral (plésiomorphie). Par simple accroissement dans la cellule, dix de ces cristaux sont devenus des spicules simples. Les clades A et B2 de l'ordre des Holacanthida, premiers clades divergeant dans la phylogénie, présentent cette morphologie avec des spicules longs et simples qui traversent la cellule de façon irrégulière (Fig. 2). Ces morpho-espèces sont très certainement fragiles, ce qui expliquerait le peu d'individus récupérés dans les traits de filets lors de cette thèse (deux individus adultes dans le clade A). D'ailleurs, l'espèce *Acanthochiasma* sp. du clade B2 n'a pu être échantillonnée qu'avec des traits de filets effectués à la nage, moins destructifs que les traits de filets tirés à partir d'un bateau. Selon la phylogénie réalisée avec une partie du gène ribosomal 18S (Fig. 5, chapitre I), le clade environnemental I aurait divergé avant le clade A, ce qui suppose, selon l'hypothétique chemin évolutif, que le squelette ait une morphologie au moins équivalente voire plus simple ou bien même absent (Fig. 2). L'étude dans le chapitre IV a montré que ce clade I est endémique de la zone mesopélagique et représente une grande majorité des acanthaires à cette profondeur. Un échantillonnage spécifique dans ces eaux profondes pourrait ainsi dévoiler dans le futur le phénotype de ce clade et venir alimenter l'hypothèse sur l'évolution du squelette des acanthaires.

Puis, au cours de l'évolution, chacun des dix spicules s'est divisé en deux. Les morphotypes ont désormais vingt spicules qui se joignent de manière plus ou moins lâche au milieu de la cellule dans une matrice contractile (Fig. 3.1 et 3.2). C'est le cas du squelette des organismes du clade C, monophylétique, représentant l'ordre des Chaunacanthida. La jonction lâche de leurs spicules fait que certains spécimens peuvent adopter une forme de parapluie lors d'un stress physique. Chez un même morphotype, les spicules peuvent être de différentes tailles (*Heteracon biformis*) et avoir de toutes petites excroissances ou «dents». La récente analyse phylogénétique avec l'ensemble de l'opéron ribosomal montre que le clade B1 branche avec le clade C avec un support acceptable (BV = 0.77), et non pas avec le clade B2 comme le montraient les précédentes analyses du chapitre I. Malheureusement, la jonction centrale des individus du clade B2 n'a pu être identifiée car ce clade est composé de séquences provenant de kystes et de l'individu Ei 68 (espèce non déterminée) dont l'endoplasme est très dense et opaque (cf annexe 2).

Ne vivant pas avec des microalgues symbiotiques, les clades A, B1 et C sont largement hétérotrophes. Leur reproduction par enkystement, probablement dû à leur squelette flexible, entraine une partie du cycle de vie en profondeur jusque dans le méso- et bathypélagique (chapitre IV). Ces acanthaires contribuent ainsi au transfert et à la séquestration du carbone et du strontium au-delà de la zone photique, plus particulièrement dans les hautes latitudes. La niche écologique de ces clades serait principalement les couches profondes même si certains comme le clade C peuvent être présents également en surface. Ceci corrobore les observations de Schewiakoff (1926) qui observaient les Holacanthida (clades A) et Chaunacanthida (clade C) plus en profondeurs avec des réserves lipidiques et de nombreuses vacuoles digestives.

Dans le cadre d'une collaboration, une étude de paléogénomique a trouvé de l'ADN ancien dans les sédiments océaniques profonds (jusqu'à 15 000 ans) affilié aux acanthaires, et plus spécifiquement à ceux qui réalisent l'enkystement. On suppose que leur position dans la colonne d'eau et la sédimentation importante des kystes augmentent les chances d'atteindre rapidement le sédiment des fonds marins sans reminéralisation complète de leurs composants cellulaires et de leur ADN. (Lejzerowicz et al., *in prep*; annexe 4). Ceci confirme le rôle biogéochimique important, soulevé dans le chapitre IV, des clades A, B et C sur le flux et la séquestration de matière en profondeur.

Dans le clade D (ordre Holacanthida), les spicules se joignent au milieu de la cellule avec un imbriquement et une fusion spécifiques. C'est l'apparition d'une première jonction centrale, qui rend sans aucun doute le squelette plus robuste (Fig. 3.3). Les organismes du clade IV (non identifiés taxonomiquement) présentent également cette morphologie mais sa position phylogénétique reste à confirmer (BV = 0.65). Contrairement aux précédents clades, les clades D et IV occuperaient plutôt la zone photique des océans (chapitre IV). Aucune microalgue endosymbiotique n'a été observée en microscopie à fluorescence ou détectée par PCR dans les différentes espèces de ces clades. Est-ce que l'accession aux eaux de surface, où la pression de prédation est plus forte, est liée à l'apparition d'un squelette plus robuste?

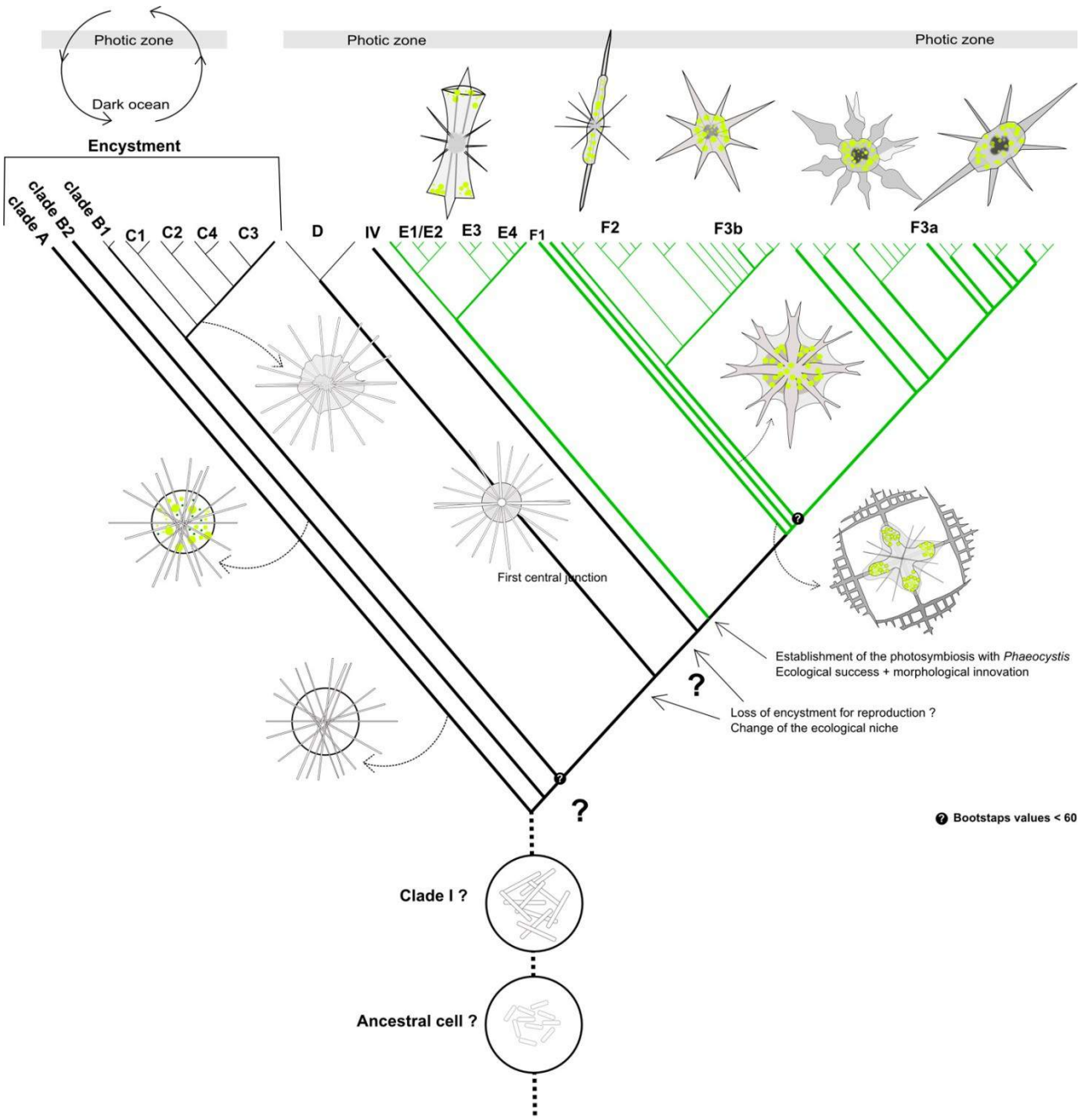


Figure 2. Schéma hypothétique de l'évolution de la morphologie et de l'écologie des acanthaires basé sur une phylogénie moléculaire en maximum de vraissemblance à partir de la concaténation des gènes de l'opéron ribosomal (ADNr 18S, 28S et ITS).

Viennent ensuite les clades E et F (ordres Arthracanthida et Symphiacanthida), synonymes de profonds changements morphologiques et écologiques. Le squelette minéral subit une modification phénotypique très importante donnant naissance à des formes complexes et très variées (Fig. 4). Une cellule peut avoir des spicules de tailles et d'épaisseurs très différentes (*Amphilonche elongata*), avec des structures supplémentaires comme des extensions latérales sur les spicules (apophyses chez *Xiphacantha* sp., *Lithoptera* sp.). La jonction centrale est très robuste car il y a une fusion complète des spicules, ce qui peut aussi entraîner l'apparition d'un corps central occupant une grande place dans l'endoplasme (*Acanthostaurus conacanthus*, Fig. 3.4 et 3.5). Dans le clade E, la présence de coquille est une synapomorphie claire. La phylogénie moléculaire a permis de distinguer deux groupes qui fabriquent différemment cette coquille : les acanthaires des sous-clades E1 et E2 ont une coquille faite d'agencement d'apophyses qui forment un treillage particulier (ex: *Phractopelta* sp.) tandis que ceux des sous-clades E3 et E4 fabriquent une coquille épaisse et robuste par agencement de plaques (*Hexaconus* sp.; *Diploconus* sp.). La complexification du squelette chez les clades E et F rend d'une manière générale le squelette plus dense. Ceci est d'autant plus significatif que le sulfate de strontium est le biominéral le plus dense dans le milieu marin (presque deux fois plus que le carbonate de calcium, Martin et al., 2010). L'augmentation du nombre de myonèmes (Schewiakoff, 1926), système régulant la flottabilité, permettrait aux acanthaires symbiotiques de se maintenir en haut de la colonne d'eau malgré leur densité, et de bénéficier ainsi d'une énergie lumineuse suffisante pour accommoder leurs microalgues intracellulaires.

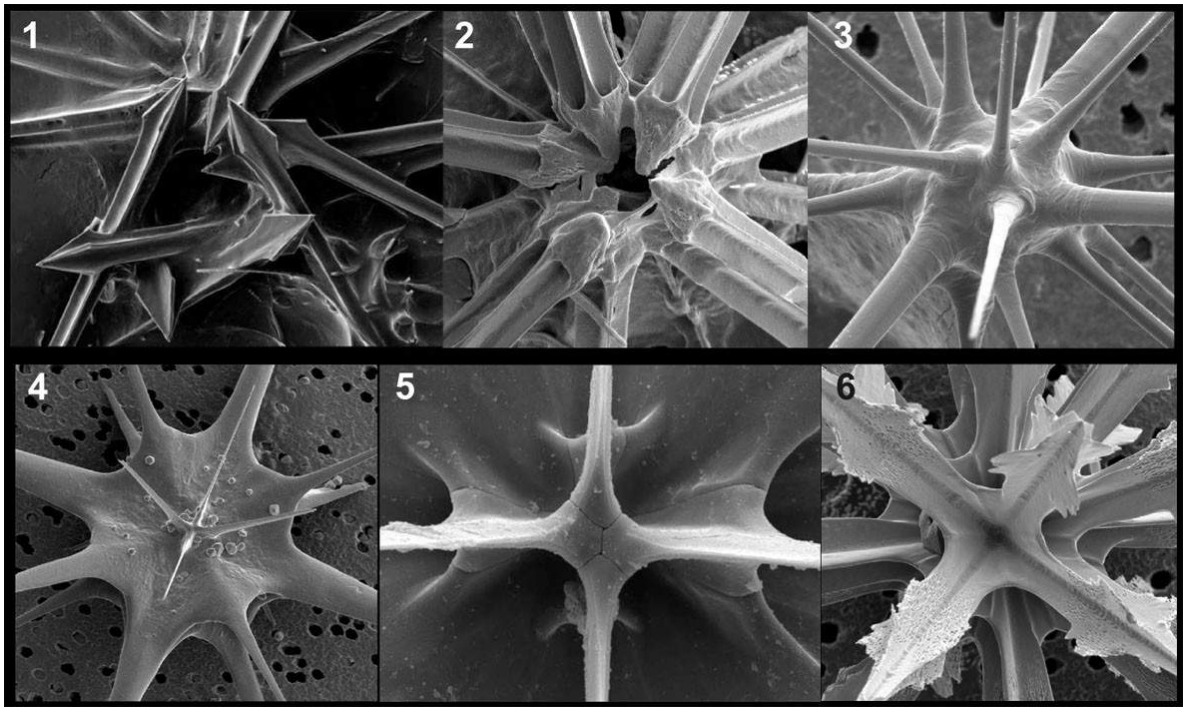


Figure 3. Evolution de la jonction centrale des spicules au milieu de la cellule d'acanthaire : absence de jonction chez le clade C (1 et 2), première jonction par fusion des spicules chez le clade D (3), et jonction développée et robuste chez le clade F (4-6) **(images de J. Decelle et O. Thomson).**

2.2 Innovation et complexité morphologique liées à la photosymbiose

La symétrie, la finesse et la beauté des squelettes exhibés par les unicellulaires marins ont fasciné les scientifiques du 19ème siècle comme Johannes Müller et Ernst Haeckel. Leur esprit de curiosité était tourné vers les mécanismes expliquant l'émergence et l'évolution de ces formes au cours du temps. Les acanthaires étaient parmi les premières observations de photosymbiose à l'époque, mais le lien entre la morphologie et la symbiose n'était pas encore démontré. La symbiose est aujourd'hui reconnue comme une source d'innovation et de complexité phénotypique au cours de l'évolution (Moran, 2007) et les acanthaires en sont une belle illustration (Fig. 4). La complexité morphologique est une tendance générale dans l'évolution des eucaryotes (Caroll, 2001), y compris chez les protistes (Schaap et al., 2006), mais reste un grand mystère aux yeux des biologistes. Bien qu'il n'existe pas de définition consensus, la complexité peut être définie par l'apparition de différents types cellulaires, une augmentation du nombre de compartiments physiques et de structures spécialisées, qui créent plus d'interactions dans un système donné (McShea, 1996; Caroll, 2001). Quels sont les processus physiologiques et/ou évolutifs qui expliquent la complexité et l'exceptionnelle diversité de formes chez les acanthaires?

La complexification et la grande densité des squelettes suggèrent un lien, plus ou moins direct, entre la photosymbiose avec *Phaeocystis* et la biominéralisation du sulfate de strontium, dont les mécanismes restent inconnus. A l'instar des coraux ou des foraminifères (Lea et al., 1995; Venn et al., 2011), le microenvironnement autour de la cellule d'acanthaire, influencé par l'activité photosynthétique, pourrait stimuler l'incorporation de strontium, qui est d'ailleurs beaucoup moins abondant que le calcium dans l'eau de mer. Il est également plausible que les produits métaboliques de *Phaeocystis* comme les sulfures méthylés (DMS, DMSP, DMSO), transférés à l'hôte (chapitre II), interviennent dans le processus de biominéralisation du sulfate de strontium. D'une manière générale, la complexification du squelette s'expliquerait d'un point de vue énergétique où la photosymbiose apporterait les besoins nutritionnels nécessaires pour la biominéralisation et ses mécanismes cellulaires associés. La photosymbiose a également entraîné une complexification morphologique chez les foraminifères benthiques comme l'apparition d'un cloisonnement et de structures particulières sur le test calcifié (Lee, 2010). La kleptoplastidie chez les foraminifères du genre *Elphidium* en est un exemple supplémentaire (Pillet et al., 2011). Mais ceci n'explique pas la raison pour laquelle les acanthaires symbiotiques, véritables architectes du minéral, ont construit des formes aussi variées (Fig. 4).

2.3 Réflexion sur la diversité des formes chez les acanthaires symbiotiques

La compréhension des processus qui influencent la diversification morphologique est fondamentale pour comprendre l'évolution de la biodiversité. Les formes des organismes sont généralement sous le contrôle de forces extérieures biotiques et abiotiques (Thompson, 1917). Dans le monde turbulent du plancton, les interactions entre les microorganismes sont perpétuelles, plus particulièrement dans la zone photique où le nombre et le type d'interactions

augmenteraient. La prédation est considérée comme la pression sélective agissant de façon prépondérante sur la morphologie des protistes marins, probablement plus que la lutte contre la gravité et la compétition pour les ressources (« *Watery arms race* » Smetacek, 2001). L'apparition de la biominéralisation de nanostructures comme les frustules de diatomées, les coccolithes des haptophytes, les tests des foraminifères, et les squelettes des radiolaires seraient autant de «boucliers» pour faire face à des prédateurs 10 à 1000 fois plus grands (Tillmann, 2004; Porter, 2011). Chez les acanthaires, toutes ces formes variées pourraient être une illustration de cette course à l'armement à la surface des océans où l'activité de prédation est forte. Outre la prédation, il est également possible que la photosymbiose impose elle-même une pression sélective sur la forme de l'hôte comme c'est le cas chez les coraux (Anthony et al., 2005; Kaniewska et al., 2008). La croissance du squelette de certains coraux permet d'améliorer de façon très importante la capture de la lumière, hypothèse émise également chez certaines espèces de coccolithophores (Young, 1994), offrant ainsi des conditions optimales pour leurs microalgues symbiotiques. De plus, en réfractant la lumière, le squelette minéral peut être un moyen de protection contre les forts rayonnements dans la zone photique (Armstrong et Brasier 2005; Reef et al., 2009). La forme du squelette des acanthaires symbiotique serait donc le résultat de toutes ces pressions multiples. Mais pourquoi existe-il une disparité phénotypique si grande chez ces acanthaires, qui *a priori* ont une physiologie similaire et font face aux mêmes pressions abiotiques et biotiques ?

Partageant la même niche écologique, il est possible que les espèces d'acanthaires symbiotiques rentrent en compétition. Selon l'hypothèse évolutive du «*character displacement*», les caractères morphologiques tendent à se différencier de manière plus importante lorsque les espèces partagent une même niche (Brown et Wilson, 1956; Dayan et Simberloff, 2005). Cette différenciation permettrait de réduire la compétition interspécifique, favorisant ainsi la coexistence des espèces. Par exemple, les espèces d'oiseaux sympatriques ont une plus grande diversité de becs que des espèces vivant dans différentes niches écologiques (Schluter, 1988).

Enfin, les différentes formes des acanthaires pourraient ne refléter qu'une importante variance morphologique aléatoire (Maynard-Smith 1970; McShea, 1996), possible grâce aux propriétés intrinsèques des cristaux de sulfate de strontium qui peuvent adopter des formes très diverses (Li et al., 2008). Ceci n'est pas sans rappeler l'incroyable diversité de formes des cristaux de glace (Libbrecht, 2005).

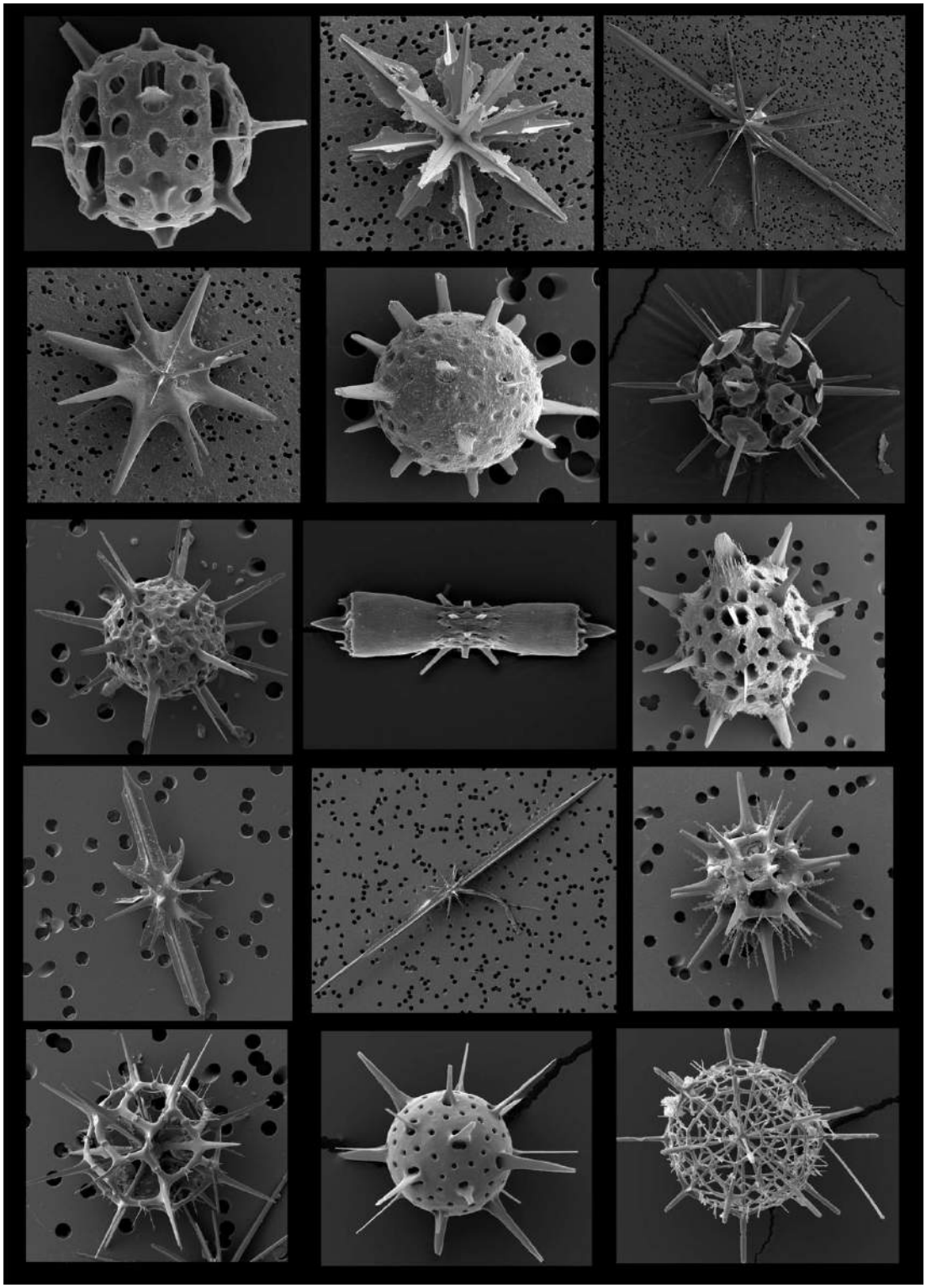


Figure 4. Diversité des formes chez les acanthaires en symbiose avec *Phaeocystis* (clades E et F). (images de J. Decelle et O. Thomson).

2.4 La photosymbiose chez les acanthaires: une possible radiation évolutive

La radiation évolutive (*adaptive radiation*) a été très étudiée pour expliquer l'évolution des formes (Simpson, 1953; Mayr, 1963). Les pinsons de Darwin et les poissons cichlidés des lacs africains sont des exemples iconiques de ce processus évolutif (Salzburger, 2009). Une radiation évolutive peut apparaître lorsque des espèces divergent et deviennent plus aptes à exploiter d'autres niches écologiques par rapport aux espèces ancestrales (Losos, 2010). L'exploitation de nouvelles niches écologiques, ou «opportunité écologique», est généralement catalysée par des innovations clés (Galis, 2001), définie comme l'émergence d'une structure ou d'une propriété permettant la performance d'une nouvelle fonction (Mayr, 1963). Cette innovation peut radicalement changer le mode d'interaction d'un organisme avec son environnement, et ouvrir la voie vers de nouvelles zones adaptatives. L'établissement de la symbiose avec *Phaeocystis* constitue clairement une radiation évolutive chez les acanthaires. La photosymbiose a été cette innovation clé qui a permis aux organismes d'exploiter la zone photique, habitat où la prédation et le stress trophique sont importants.

L'ouverture écologique déclenche typiquement une extraordinaire diversification ou « *burst* » de la biodiversité (Simpson, 1953). Les analyses phylogénétiques bayesiennes et en maximum de vraisemblance semblent montrer que les clades E et F représentent une plus grande diversité taxonomique et génétique que les clades non symbiotiques (même si cela dépend grandement de la couverture de l'échantillonnage) et que cette diversification est apparue récemment sur un laps de temps relativement court. De plus, comme observé chez les clades E et F, les taxa qui ont subi une radiation évolutive montrent généralement un degré exceptionnel de disparité phénotypique (Losos et Mahler, 2010). Les acanthaires ne sont pas le seul exemple de radiation évolutive liée à la photosymbiose. L'acquisition de microalgues symbiotiques a été aussi le moteur de la diversification des coraux sclératinaires au début du Mésozoïque (Stanley, 2006) ainsi que des foraminifères benthiques et pélagiques (Norris, 1996; Richardson, 2001).

En revanche, la photosymbiose chez *Acanthochiasma* sp du clade B2 n'a pas entrainé les changements morphologiques et écologiques observées chez les clades E et F. Au contraire, le squelette de cette espèce est très simple. Apparue avant celle des clades E et F avec *Phaeocystis*, cette symbiose n'est pas du tout spécifique d'un type de symbiote (plusieurs genres, de plusieurs lignées) et s'apparenterait plutôt à une symbiose généraliste et opportuniste de microalgues (chapitre II-B). Cette évolution de la symbiose avec des microalgues chez les acanthaires pourrait faire l'objet d'une analogie avec l'agriculture chez les insectes ou chez l'Homme qui, au départ fut plutôt généraliste puis s'est ensuite spécialisée sur certaines plantes ou champignons qui ont fournit un meilleur rendement à leur hôte (Salamini et al., 2002 ; Mueller et al., 1998).

2.5 Succès écologique de la photosymbiose avec *Phaeocystis*

Outre les conséquences majeures sur la morphologie, la photosymbiose a bouleversé l'écologie et le rôle biogéochimique des acanthaires appartenant aux clades E et F. Le caractère obligatoire de cette symbiose implique une acquisition rapide des microalgues lors du développement de la cellule. La perte de l'enkystement associée à une reproduction directement de la forme végétative, hypothèse émise dans le chapitre IV, permettraient un cycle de vie contenu dans la zone photique où les chances d'interactions avec le symbiote sont plus grandes. La contribution biogéochimique des clades E et F aux flux de matières organiques serait alors moindre mais ils participent en revanche à la production primaire *via* leurs microalgues symbiotiques (Michaels, 1988). Leur présence prépondérante dans la zone photique expliquerait également la diminution de la concentration en strontium en surface, par rapport aux zones plus profondes où le minéral est reminéralisé par la dissolution des squelettes (De Villiers, 1999; De Dekker, 2004).

La photosymbiose avec *Phaeocystis* représente indéniablement un succès écologique. Grâce à la photosynthèse, l'holobionte bénéficie d'une source nutritionnelle carbonée, et possiblement azotée, permettant de s'affranchir de certains facteurs limitants comme les nutriments, et ainsi d'exploiter la surface oligotrophique des océans. Par rapport aux autres clades, les acanthaires symbiotiques des clades E et F sont les plus abondants dans la zone photique des océans mondiaux (chapitres III et IV). Certes, le « *metabarcoding* » environnemental utilisé au cours de ces chapitres ne donne qu'une abondance relative de séquences, mais les données moléculaires confirment les observations de la littérature (Stoecker et al., 1996) et faites sous le microscope au cours de cette thèse lors des prélèvements dans différentes régions.

Egalement observé chez les foraminifères planctoniques (Norris, 1996), les acanthaires symbiotiques auraient une distribution plus cosmopolite que les non-symbiotiques (chapitre III). Cette capacité importante de dispersion dans les océans serait directement liée à la nature ubiquitaire du symbiote *Phaeocystis*. De plus, la mixotrophie des acanthaires symbiotiques leur confèreraient probablement une zone optimale physiologique de croissance plus large permettant de faire face à de longues périodes de stress trophique. Par exemple, les foraminifères symbiotiques (*Globigerinoides ruber*, *Globigerinoides sacculifer*, et *Orbulina universa*), qui ont une physiologie similaire, tendent à maintenir leur taux de croissance sur une large gamme de températures contrairement aux espèces non-symbiotiques (Lombard et al., 2009). Il est intéressant de constater que cette photosymbiose ne se situe pas uniquement dans les zones tropicales pauvres en nutriments comme attendu, mais se retrouve également dans les régions côtières, plus productives, comme la Manche avec *Phaeocystis globosa* et l'Antarctique avec *Phaeocystis antarctica*. La photosymbiose chez les acanthaires pourrait présenter des avantages de nature différente selon la zone géographique et le régime trophique de la masse d'eau. Les plastes ont d'autres fonctions métaboliques que la photosynthèse. Ils peuvent fournir de l'oxygène à l'hôte, jouer un rôle pour dans le métabolisme de l'azote, produire des acides aminés ou des lipides essentiels (Grzymski et al., 2002). Le fait de trouver des protistes contenant un plaste dans des zones sans lumière démontre bien que cet organite n'est pas que

photosynthétique (Bernhard et Bowser, 1999). Les bénéfices apportés par la microalgue symbiotique peuvent être ainsi très diverses et variables selon l'environnement. Il est possible que l'association symbiotique acanthaire-*Phaeocystis* reçoive des pressions sélectives environnementales différentes (entre la Mer Méditerranée et l'Antarctique par exemple), qui selon la théorie de la coévolution en mosaïque géographique (Thompson, 1999), pourrait coévoluer indépendamment.

Dans l'Océan Austral, des expériences d'enrichissement en fer déclenchent une forte croissance de la population de *P. antarctica* (Hoffmann et al., 2006). Du côté du microzooplancton, seuls les acanthaires répondent positivement à cet enrichissement, contrairement aux autres radiolaires, foraminifères et ciliés dont l'abondance ne varie pas (Henjes et al., 2007). Les acanthaires représentent alors plus de 90% de la biomasse du microzooplancton. L'hypothèse émise par les auteurs est une réponse classique de type proie-prédateur, les acanthaires se nourrissant spécifiquement de l'efflorescence de *Phaeocystis* déclenchée par l'apport de fer. Ce travail de thèse a permis de dévoiler une interaction plutôt du type symbiotique entre *Phaeocytis* et les acanthaires et apporte un éclairage nouveau pour l'interprétation de ces observations à l'échelle des populations. Dans le Golf de Naples, depuis le début du suivi d'observations en 2004, une efflorescence d'acanthaires symbiotiques du clade F (*Acanthometra pellucida* en 2010) a lieu chaque année à la fin du printemps (mai-juin). Ce phénomène a été aussi rapporté dans d'autres régions comme la Mer des Sargasses (Massera Bottazzi et Andreoli, 1981). Est-ce que cette augmentation d'acanthaires symbiotiques précède ou non un « *bloom* » de *Phaeocystis*? Est-ce que l'écologie de *Phaeocystis* est à son tour influencée par la présence des acanthaires ? A l'avenir, il serait intéressant d'examiner les variations temporelles des populations libres de *Phaeocystis* et de leurs hôtes à un point spatial donné pour mieux comprendre les influences écologiques réciproques entre partenaires symbiotiques.

3. Hypothèses sur le fonctionnement et l'évolution de la photosymbiose pélagique

Figure 5. Comparaison des propriétés physiques entre le milieu pélagique (gauche, photo de J. Decelle) et le milieu récifal (droite, photo de National Geographic).

3.1 Fonctionnement des photosymbioses marines

La photosymbiose est un processus répandu en milieu marin chez les eucaryotes pour acquérir la photosynthèse mais ce type d'interaction semble prendre des configurations différentes selon les conditions environnementales. Ces travaux ont mise en évidence deux types de photosymbioses impliquant des microalgues comme *Heterocapsa*, *Phaeocystis*, *Chrysochromulina*, qui sont connues pour être abondantes et cosmopolites dans les océans. Certaines forment même des efflorescences très importantes sur les côtes, entrainant des effets néfastes sur la chaîne alimentaire mais aussi pour la pêcherie et les activités touristiques (Edvardsen et Imai, 2006). Bien que leur succès écologique ait fait l'objet de nombreuses études, leur mode de vie symbiotique n'était pas connu auparavant. Une symbiose mutualiste mettant en jeu un symbiote très abondant et écologiquement prépondérant en phase libre n'était pas la vision communément admise avant cette thèse. Les symbiotes décrits à partir d'associations terrestres et marines sont généralement très rares dans l'environnement en phase libre (Nyholm et al., 2004), et certains auteurs les considèrent même comme membres de la « biosphère rare » (Webster et al., 2010; Pedrós Alió, 2012). C'est d'ailleurs la raison pour laquelle leur écologie en dehors de l'hôte est quasiment inconnue (Douglas, 1996), une composante pourtant primordiale pour comprendre précisément le mode d'interaction en question. Quelles sont les raisons qui expliquent ces différents modes de vie libres des symbiotes entre le milieu pélagique et les autres écosystèmes? Nous orienterons ici notre comparaison sur la photosymbiose en milieu pélagique et récifal.

L'écosystème récifal et la surface des eaux océaniques sont des environnements caractérisés par un fort ensoleillement et un appauvrissement en nutriments. Cependant, ils présentent des propriétés physiques radicalement différentes, qui pourraient avoir une influence directe sur le mode d'interaction des organismes (Fig. 5). Tout d'abord, le milieu océanique est un habitat

largement tridimensionnel qui représente un volume d'eau gigantesque. La zone photique peut atteindre jusqu'à 200 mètres de profondeur dans les eaux les plus oligotrophes et s'étendre horizontalement sur des centaines de milliers de kilomètres. Les microorganismes y sont faiblement concentrés et leurs interactions physiques sont soumises aux mouvements browniens chaotiques. La turbulence du milieu peut également limiter les communications chimiques durables et à distance entre les organismes. Dans le sol, les symbioses entre plantes-champignons et légumes-*Rhizobia* s'établissent à distance dans un premier temps grâce à la communication chimique. Le rapprochement physique peut ensuite être facilité par un tropisme où un partenaire oriente sa croissance vers l'autre et/ou un tactisme où le symbiote mobile se déplace vers l'hôte (ex : Rhizobiacées, Selosse, 2010). Dans le milieu pélagique, l'établissement d'associations symbiotiques où une microalgue particulière doit être acquise dans l'environnement à un moment précis est *a priori* plus compliqué et imposerait de fortes pressions sélectives. Nous avons émis l'hypothèse dans le chapitre II qu'une association avec un partenaire abondant comme *Phaeocystis* augmenterait les chances d'interactions physiques entre cellules et donc d'établir plus facilement une relation symbiotique. La description récente d'autres photosymbioses pélagiques avec les cyanobactéries, *Synechoccocus* et *Prochlorococcus*, cosmopolites et parmi les organismes les plus abondants dans les océans (Foster 2006a et b; Zwirglmaier et al., 2008; Yuasa et al., 2012), viennent conforter notre hypothèse sur ce mode d'associations symbiotiques dans le milieu pélagique.

De plus, la durée du cycle de vie de l'hôte aurait des implications majeures sur le mode de fonctionnement de la relation. Dans le milieu pélagique, les hôtes unicellulaires comme les radiolaires et les foraminifères ont des temps de générations très courts, de 3 à 4 semaines environ (Anderson, 1983; Bijma et al., 1990), ce qui implique une acquisition *de novo* récurrente des symbiotes. L'interaction avec un partenaire abondant est d'autant plus avantageuse dans ce cas. Les travaux de cette thèse n'ont pas permis de montrer si l'acquisition des *Phaeocystis* se fait une seule fois au début du cycle de vie ou plusieurs fois durant la croissance de l'hôte. Cette dernière possibilité implique que le symbiote soit «disponible» non seulement à un moment précis mais durant toute la durée du cycle de l'hôte.

Dans le milieu récifal, les microalgues en phase libre sont au contraire relativement rares mais les caractéristiques de cet écosystème et celles des organismes photosymbiotiques qui y vivent semble constituer des obstacles moins critiques pour l'acquisition d'un symbiote libre. Se trouvant au niveau de zones côtières peu profondes, entre 5 et 20 mètres, le volume de l'écosystème récifal est beaucoup moins conséquent que l'immensité océanique. Comparée aux protistes, les hôtes pluricellulaires des récifs ont une durée de vie très longue, de plusieurs centaines d'années pour certains groupes (Barnes et Lough, 1996; Fairbanks et al, 1997). L'acquisition des microalgues semble par conséquent moins dynamique et moins critique pour la survie de l'hôte par rapport à l'océan ouvert.

De plus, les invertébrés hôtes (coraux, mollusques, éponges etc.) sont quasiment tous benthiques et ont un contact étroit avec le sédiment dès les premiers stades de vie. C'est

d'ailleurs dans le sédiment que la population libre de symbiotes serait la plus grande (1000–2000 cells/mL) par rapport à la colonne d'eau (< 20 cells/mL) (Littmann et al., 2008). Les interactions symbiotiques sont également facilitées par le fait que les larves de ces invertébrés peuvent se mouvoir et leurs capacités de chimiotactisme et même acoustiques leur permettent de sélectionner très tôt au cours de leur développement une zone de recrutement appropriée (Gleason et al., 2009; Birrell, 2008; Vermeij et al., 2010). Basée sur des expériences *in vitro*, les microalgues symbiotiques *Symbiodinium* seraient également capables de se mouvoir et de se diriger vers un hôte par chimiotactisme (Pasternak et al., 2006; Yacobovitch, 2004). Cette capacité de chimiotactisme, qui facilite l'interaction avec un hôte, est d'ailleurs grandement diminuée lorsque la turbulence du milieu augmente (Yacobovitch, 2004). Contrairement au milieu pélagique, la turbulence et les courants sont généralement très faibles dans le milieu récifal (Lesser et al., 1994; Goldshmid et al., 2004), ce qui peut favoriser une communication chimique pour établir des relations de type hôte-symbiote. Enfin, on pourrait qualifier la transmission de pseudo-verticale car la densité d'hôtes partageant et relarguant continuellement les mêmes symbiotes dans un microenvironnement du récif est telle que la transmission se ferait quasiment d'hôte à hôte avec un passage très court en phase libre.

Les caractéristiques physiques du milieu pélagique, qui imposent des pressions fortes sur les interactions écologiques, expliqueraient ainsi le mode original de la photosymbiose mise en évidence au cours de cette thèse. Des études supplémentaires doivent néanmoins confirmer les observations et les hypothèses de ce travail. En particulier, l'écologie de la phase libre d'autres microalgues symbiotiques (ex : *Pelagodinium béii*, symbiote d'*Acanthochiasma* sp. - chapitre II - et des foraminifères planctoniques) doit être étudiée par diverses techniques (*metabarcoding* environmental avec des marqueurs génétiques à haute résolution, l'hybridation *in situ* de sondes fluorescentes (*FISH*) etc…).

Je propose ici un schéma hypothétique décrivant les configurations possibles des photosymbioses pélagiques et récifales (Fig. 6). Dans le récif, la population des microalgues symbiotiques serait majoritairement dans les tissus d'un hôte (plusieurs millions de cellules par cm^2, Stimson et al., 2001) tandis que dans le milieu pélagique la population est en phase libre. Ceci aurait une influence sur leur évolution car les pressions sélectives agiraient différemment dans les deux milieux (Fig. 6). Il est intéressant de noter que les *Symbiodinium* évoluent six fois plus vite environ que les *Pelagodinium* (Shaked et de Vargas, 2006). De plus, par rapport aux autres dinoflagellés qui forment des « *blooms* » comme *Heterocapsa* sp., les *Symbiodinium* possèdent un très petit génome, jusqu'à 50 fois moins important (LaJeunesse et al., 2005). Ces deux études indiqueraient l'empreinte génétique du mode de vie symbiotique sur *Symbiodinium* et confirmerait le schéma hypothétique (Fig. 6). Avec une population plus petite en phase libre et une reproduction principalement clonale *in hospite* (plusieurs générations par semaine), le brassage génétique serait limité chez les *Symbiodinium* laissant place à une possible dérive génétique, phénomène observé chez les bactéries symbiotiques (Moran, 1996). La coévolution

hôte-symbiote serait donc plus marquée dans le milieu récifal, menant à des dépendances de la microalgue vis-à-vis de son hôte (ex : voies métaboliques). Au contraire, les microalgues symbiotiques de l'océan ouvert garderaient tout leur répertoire génétique. Nous verrons plus précisément dans le point 3.3 de cette discussion les conséquences possibles de cette configuration sur l'évolution et le maintien des photosymbioses dans le plancton.

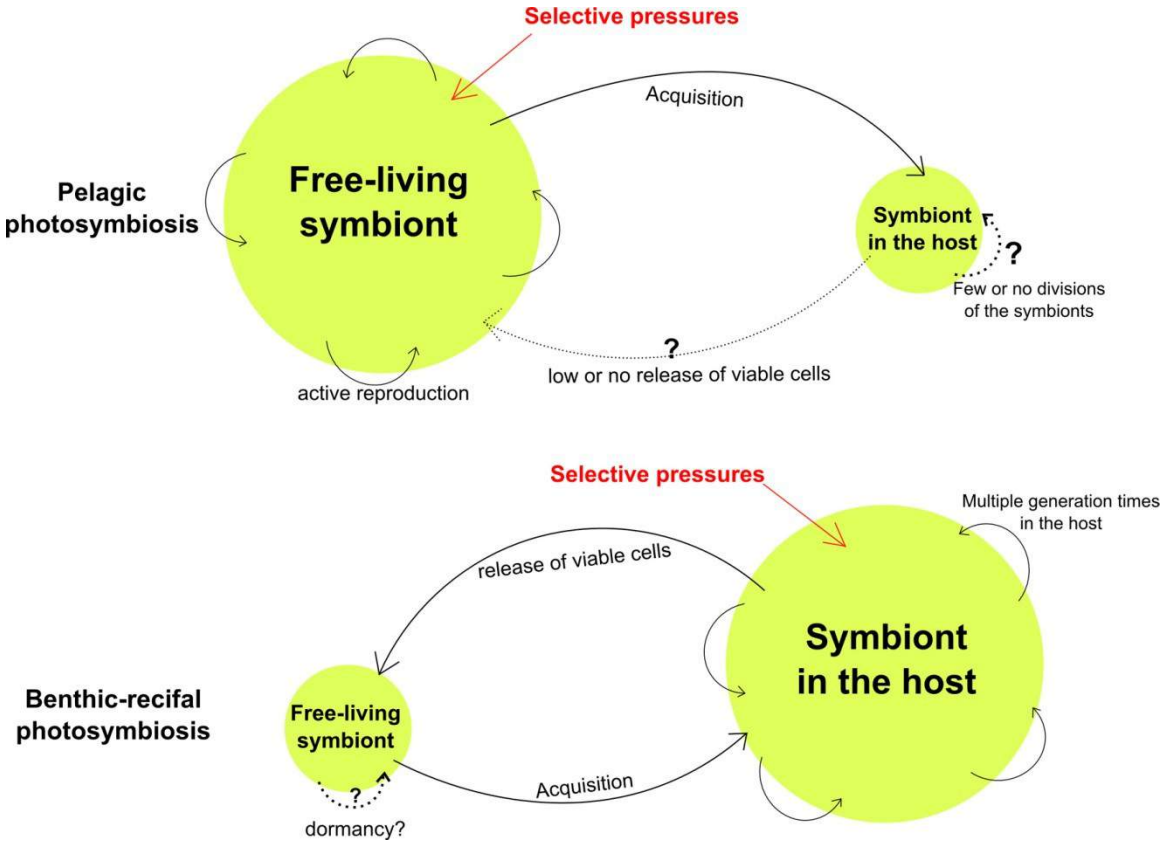


Figure 6. Modes hypothétiques de la photosymbiose en milieu pélagique et récifal. La taille du cercle vert est corrélée avec la taille de la population.

3.2 Prédispositions des microalgues à la vie symbiotique

Une des grandes interrogations concernant la photosymbiose marine est la raison pour laquelle certaines algues se retrouvent en symbiose avec une grande diversité d'hôtes, phylogénétiquement éloignés. Il semblerait que ces microalgues aient des prédispositions à vivre en mode symbiotique mais les raisons sous-jacentes n'ont pas été clairement élaborées. Ce travail de thèse tente d'apporter certaines réponses sur ce point.

Tout d'abord, à un niveau taxonomique très large, les microalgues décrites en photosymbiose dans le milieu marin à ce jour sont pour la grande majorité issues de la lignée « rouge » (Fig. 1 de l'introduction; Stoecker et al., 2009). Nos travaux, qui révèlent des haptophytes (*Phaeocystis*, *Chrysochromulina*) et dinoflagellés (*Heterocapsa*, *Pelagodinium*) en

symbiose avec les acanthaires, viennent corroborer cette tendance générale. Depuis le Mésozoïque (-251 millions d'années), la diversité du phytoplancton dans les océans est dominée par les microalgues « rouges » (Falkowski, 2004). Par rapport aux microalgues vertes, leur nombre important de gènes plastidiaux les rendrait plus versatiles métaboliquement et ainsi plus aptes à vivre dans des environnements plus contrastés comme par exemple au sein d'un hôte (*The portable plastid hypothesis*, Grzebyk et al., 2003).

La lignée «rouge» représente une grande diversité de microalgues. Certaines d'entre elles peuvent être en symbiose avec un spectre d'hôtes important. Par exemple, le dinoflagellé *Scrippsiella nutricula* est en symbiose avec des radiolaires (collodaires et spumellaires, Gast et al., 2000) mais aussi avec une méduse (*Vellela vellela*). Le chapitre II a mis en évidence que certaines microalgues symbiotiques d'*Acanthochiasma* sp. sont connues par ailleurs pour être en symbiose avec d'autres hôtes planctoniques (*Chrysochromulina* et *Pelagodinium béii*, Gast et al., 2000; Shaked et de Vargas, 2006). Ce « partage des symbiotes » est aussi une configuration retrouvée dans d'autres systèmes photosymbiotiques comme les invertébrés récifaux et les lichens (Doering et Piercey-Normore, 2009; Finney et al., 2010). Certaines algues qui interagissent avec un large spectre d'hôtes semblent donc avoir des prédispositions à vivre en symbiose. Quels sont les critères qui font d'une microalgue un symbiote potentiel ?

Nous avons suggéré précédemment que les photosymbioses pélagiques impliquent des microalgues qui sont abondantes pour augmenter les chances d'interactions à chaque génération mais aussi au cours de la croissance de l'hôte. A ma connaissance, aucune symbiose pélagique n'a été décrite avec *Symbiodinium* jusqu'à présent. Pourquoi alors une microalgue symbiotique si « serviable » dans le milieu récifal ne pourrait pas l'être dans les régions océaniques ? *Symbiodinium* est probablement adapté au milieu côtier récifal et serait beaucoup moins abondant au large par rapport à d'autres microalgues comme *Phaeocystis* et possiblement *Pelagodinium béii* ou *Scrippsiella vellela*. Il est intéressant de noter que la méduse *Cassiopea xamachana*, qui vit fixée sur le sédiment contrairement à son homologue pélagique *Vellela vellela*, entretien une photosymbiose non pas avec *Scrippsiella vellela* mais avec le genre *Symbiodinium* (Rowan et Powers, 1992). Les acoeles évoluant sur les côtes récifales et littorales vivent respectivement avec les genres *Symbiodinium* et *Tetraselmis* (prasinophyte), alors que les espèces soeurs dans le milieu pélagique sont pour la plupart en symbiose avec des diatomées (Mc Adam et Balzer, 2004). L'abondance et la distribution géographique des symbiotes libres constitueraient une première prédisposition d'ordre écologique.

La microalgue symbiotique doit avoir également des prédispositions physiologiques qui procurent à un hôte une fitness importante. Outre les compétences requises liées à la température ou à la qualité de la lumière (Putman et al., 2012), nos études se sont penchées sur une adaptation physiologique particulière. Nous avons trouvé de grandes concentrations de DMSP/DMS/DMSO dans les acanthaires en symbiose avec *Phaeocystis* (chapitre II-A). Cette microalgue est extrêmement bien connue pour produire ces composés sulfurés car cela a des répercutions majeures dans le cycle global du soufre et plus généralement sur le climat (Simó, 2001). D'autres études ont également détectés de fortes concentrations en DMSP dans différents hôtes comme les

coraux, les anémones et les acoeles (Yost et Mitchelmore, 2009; Van Bergeijk, 2001) mais les raisons n'ont pas été abordées dans ces études. Compte tenu du pouvoir antioxydant de ces composés soufrés (Sunda et al., 2002), nous avons émis l'hypothèse dans le chapitre II que ces molécules sont accumulées dans l'holobionte pour se protéger des radicaux libres produits par la photosynthèse et le fort ensoleillement de la zone photique. Chez les coraux, des acides aminés appelés « mycosporine-like amino acids » ou MAAs sont des molécules anti-UV présentes également en grandes concentrations dans les tissus. A l'instar du DMSP, les MAAs sont produites par les microalgues (*Symbiodinium*) et transférées à l'hôte. *Phaeocystis* en culture est aussi connue pour être un producteur important de MAAs (Jeffrey et al., 1999; Moisan et Mitchell, 2001), ainsi que la cyanobactérie *Prochloron didemni*, endosymbiote de certaines ascidies des récifs coraliens (Donia et al., 2011). Toutes ces observations montrent l'importance de la photoprotection dans les relations photosymbiotiques. Les microalgues doivent ainsi être capables de produire des métabolites essentiels pour le fonctionnement et le maintien de la photosymbiose. Des analyses supplémentaires devront être entreprises pour étudier plus en détail les mécanismes de photoprotection et de stress oxydatif chez les radiolaires symbiotiques et appuyer notre hypothèse sur cette compétence symbiotique nécessaire de la microalgue.

3.3 Evolution et stabilité de la photosymbiose pélagique

Malgré leur rôle important et leur caractère commun dans les écosystèmes, les symbioses dites mutualistes restent un paradoxe insolvable aux yeux des théoriciens de l'évolution contrairement aux autres interactions hôte-parasite et proie-prédateur qui peuvent s'expliquer par des modèles théoriques (Herre et al., 1999; Bronstein, 2003; Sachs et al., 2004). En effet, la coopération entre organismes phylogénétiquement éloignés pose un problème central dans la théorie de l'évolution: la sélection naturelle favoriserait les tricheurs qui peuvent acquérir « égoïstement » tous les bénéfices de la relation sans fournir aucun retour. Il a été montré que la transmission verticale, transfert direct des symbiotes à la progéniture, peut stabiliser une relation mutualiste (Herre et al., 1999, Wilkinson et Sherratt, 2001; Sachs et al., 2004). Mais malgré cette source de conflit et d'instabilité, les relations mutualistes à transmission horizontale ont persisté au cours de l'évolution et prospèrent dans de nombreux écosystèmes actuels (microalgues-coraux, rhizobia-légumes, arbres-champignons etc.). Pour quelles raisons ces symbioses n'ont pas évolué vers la transmission verticale ? Plus particulièrement, pourquoi et comment la photosymbiose pélagique à transmission horizontale s'est maintenue et n'a pas évolué vers l'acquisition définitive de la photosynthèse?

La transmission verticale est considérée comme un moyen d'assurer l'acquisition d'un symbiote sans coût important de rapprochement, plus particulièrement lorsque l'accès aux symbiotes libres est limité (Nobre et Aanen, 2010). Cependant, nous avons vu que certains hôtes planctoniques comme les acanthaires interagissent avec des symbiotes relativement abondants, ce qui allège cette pression d'interaction obligatoire.

La transmission verticale permet également l'uniformité des génotypes et la fidélité du symbiote, conditions importantes pour la stabilité de la coopération (Franck, 1996; Sachs et al., 2004). Une population symbiotique hétérogène peut être une source de parasitisme si les symbiotes entrent en compétition dans l'hôte pour les ressources et l'espace. Cela diminue la fitness de l'holobionte et peut déclencher la rupture de la relation symbiotique. (Franck, 1996). Cependant, l'existence de nombreuses interactions mutualistes à transmission horizontale qui évoluent depuis plusieurs centaines de millions d'années contredisent la théorie. Concernant les photosymbioses, il est très commun de trouvé une population hétérogène de *Symbiodinium* chez les coraux durs et les bénitiers géants (Baker et Romanski, 2007; Carlos et al., 2000). Environ 11% des cellules de foraminifères planctoniques abritent plusieurs génotypes de *Pelagodinium béii* (Shaked et de Vargas, 2006). Les photosymbioses sont généralement spécifiques d'un genre de microalgue, mais flexibles sous ce niveau taxonomique. L'« uniformité des génotypes » se ferait alors au niveau du genre. Les approches moléculaires utilisées lors de cette thèse ne permettent pas de connaître la diversité intraspécifique de *Phaeocystis* dans un hôte. On suppose que la flexibilité observée à l'échelle des océans peut aussi se retrouver dans une même localité où plusieurs espèces de *Phaeocystis* cohabitent. Un même hôte pourrait vivre simultanément avec différentes espèces de *Phaeocystis* car le caractère rendant compatible l'association est spécifique du genre. La symbiose dévoilée chez *Acanthochiasma* sp. est un cas très singulier car elle implique des symbiotes de genres différents, voire de lignées distinctes (haptophytes et dinoflagellés) au sein de la même cellule. Ce type de photosymbiose n'a jamais été caractérisé à ma connaissance et appelle à d'avantages d'investigations pour comprendre son fonctionnement et les mécanismes de spécificité.

L'importance théorique de la transmission verticale n'explique donc pas la stabilité évolutive des photosymbioses. Ce mode de transmission où les symbiotes doivent être accomodés durant tout le cycle de vie d'un hôte peut parfois être coûteux (Yamamura, 1997). Le symbiote est bénéfique mais peut être aussi parasite à certains stades de vie de l'hôte (Koide, 1985). Des barrières physiques peuvent également expliquer l'impossibilité d'une transmission directe à la descendance (Douglas, 1994). Le cas des acanthaires ou des foraminifères planctoniques a été évoqué durant cette thèse (chapitre II- A). La production de «*swarmers*» d'une taille inférieure à la microalgue empêcherait la transmission verticale. Cependant, le peu de connaissances sur la reproduction de ces organismes n'écarte pas la possibilité d'une reproduction asexuée qui permettrait de transmettre verticalement les symbiotes, mais cela reste à démontrer.

La transmission horizontale, considérée comme peu stable et ancestrale (Sachs et al., 2011), présente aussi des avantages dans certains environnements par rapport à la transmission verticale. En maintenant un brassage génétique avec la population libre, les microalgues gardent tout leur répertoire génétique et continuent de répondre aux pressions sélectives de l'environnement. Les microalgues restent bénéfiques pour l'hôte même si les conditions environnementales changent. A l'inverse, une transmission verticale déclenche typiquement une réduction drastique du génome du symbiote (Müller's ratchet), ce qui peut diminuer sa flexibilité

vis-à-vis de l'environnement et perdre son pouvoir bénéfique pour l'hôte (Sachs et al., 2011). D'ailleurs, chez les bactéries, la détérioration importante de génomes pourrait expliquer le renouvellement de symbiotes avec un génome complet (Douglas, 2010). En transmission horizontale, l'hôte a l'avantage de sélectionner un symbiote adapté à son environnement qui va lui apporter la meilleur fitness (Wilkinson et Dickinson, 1995). Interagir avec des symbiotes génétiquement différents, ayant des capacités physiologiques distinctes, est fréquent chez un hôte comme les coraux, exposé à des conditions changeantes de températures ou de lumière (Baker, 2003; LaJeunesse, 2001). Concernant les acanthaires, il serait intéressant de voir s'ils changent leurs symbiotes au cours d'une année, selon les saisons, ou s'ils gardent le même génotype *Phaeocystis*.

L'évolution vers une transmission verticale, étape qui permettrait l'acquisition complète d'un organite comme le plaste, ne dépendrait pas seulement de transferts de matériel génétique ou de protéines de transport (Bhattacharya et al., 2007), mais dépendrait également des paramètres environnementaux de l'habitat et des composantes morphologiques des partenaires. La transformation des symbiotes en organites reste un événement rare dans l'évolution, suggérant la difficulté de ce processus chez les eucaryotes (Cavalier-Smith et Lee, 1985).

Quelles sont donc les raisons qui expliquent le maintien de la transmission horizontale de la photosymbiose pélagique ? Basée sur un modèle mathématique, une étude a démontré que le mutualisme peut être stable et évoluer sans transmission verticale si la population libre du symbiote est très grande (Genkai-Kato et Yamamura, 1999). Le modèle a pris en compte les coûts de la transmission verticale et la capacité du symbiote à se reproduire en dehors de l'hôte. Le résultat de ce modèle rejoint la configuration de la photosymbiose pélagique mettant en jeu de grandes populations libres de symbiotes, et serait une explication concernant l'histoire évolutive de cette interaction. Au-delà du modèle mathématique théorique, une interprétation peut être donnée. La sélection naturelle agirait seulement sur la population en phase libre plutôt que sur la petite population symbiotique dans l'hôte (Fig. 6). Dans ce cas, les pressions sélectives sur la population libre n'auraient donc peu voire aucune influence sur l'évolution de l'interaction. D'ailleurs, chez les acanthaires, l'entrée en symbiose de *Phaeocystis* semble être un cul-de-sac pour la microalgue car aucune cellule viable ne reviendrait alimenter le réservoir libre de symbiotes.

Pour comprendre la dynamique évolutive des interactions écologiques, l'hypothèse de la Reine Rouge est l'explication principale, où chaque espèce essaye d'évoluer plus vite pour gagner un avantage sur l'autre espèce (Van Valen, 1973). Des taux d'évolution rapides sont donc favorisés pour les relations antagonistes de type proie-prédateur et parasite-hôte. Mais concernant les relations mutualistes, une étude a montré que ce serait au contraire des taux d'évolution lents qui permettraient le maintien de l'interaction: c'est l'hypothèse de « l'effet du Roi Rouge » (Bergstrom et Lachmann, 2003). Pour la photosymbiose des acanthaires, les microalgues symbiotiques auraient des taux d'évolution relativement lents due à la taille très importante de

leur population effective (Piganeau et al., 2011; Lynch 2006, 2008). L'effet du Roi Rouge serait donc un processus évolutif additionnel pour expliquer le maintien de la relation photosymbiotique chez les acanthaires et est une hypothèse à tester pour explorer ce type d'interaction.

3.4 Mutualisme ou parasitisme inverse chez les acanthaires et autres photosymbioses pélagiques ?

The more we learn about the diversity of life and the structure of genomes, the more it appears that much of the evolution of biodiversity is about the manipulation of other species—to gain resources and, in turn, to avoid being manipulated (John Thomspon, 1999)

Les microalgues symbiotiques dans le milieu pélagique semblent largement indépendantes vis-à-vis de leurs hôtes. Peut-on considérer alors la photosymbiose pélagique comme une relation mutualiste? C'est une question qui remonte à la découverte de ces interactions (cf point 1 de l'introduction), et est plutôt du domaine philosophique tant les interprétations deviennent invérifiables et surement anthropocentristes. Déjà Maurice Caullery, étudiant de la Sorbonne, déclarait au début du 20ème siècle «*L'association des algues et des animaux n'est pas, à tout le monde, une parfaite symbiose et que, dans certaines circonstances, l'animal vit en parasite sur les algues.*» (Caullery, 1922, p.290). Selon Sapp (1994), c'est une exploitation mutuelle: morphologiquement, l'algue est le parasite et physiologiquement, l'animal serait le parasite. Le mot mutualisme, comme le mot symbiose, sont des termes ambigus. Il n'est pas clairement défini si le mutualisme implique que la symbiose est obligatoire pour les deux partenaires. Dans le cas des photosymbioses, l'interaction est indéniablement obligatoire pour l'hôte mais l'est probablement peu ou pas pour la microalgue. La question du mutualisme s'est déjà posée pour les coraux (Douglas et Smith, 1989; Wooldrige, 2010) et pour les lichens (Lücking et al., 2009), mais le cas de *Phaeocystis* est un exemple particulièrement frappant. On parle de mutualisme lorsque les deux partenaires augmentent leur fitness en interagissant. Pour *Phaeocystis*, les avantages éventuels de vivre en symbiose dans les acanthaires pourraient être la protection vis-à-vis des prédateurs et des attaques virales, et un moyen également de dispersion. Mais étant donné les immenses populations libres dans les océans, il est peu probable que cette symbiose soit déterminante pour l'évolution et l'écologie de *Phaeocystis* (Fig. 6).

L'influence de l'hôte sur le symbiote semble corrélée avec le degré d'intégration spatiale du symbiote, de l'ectosymbiose à l'endosymbiose (Smith et Douglas, 1987). Chez les acanthaires, les symbiotes sont dans l'endoplasme ce qui montre un niveau d'intégration important. En effet, les acanthaires semblent prendre contrôle sur la croissance de leurs symbiotes et les transforment pour en tirer un maximum de bénéfices (chapitre II, Fig. 2). Des études supplémentaires devront examiner si les cellules de *Phaeocytis* se divisent dans l'hôte (1

ou plusieurs générations) pour mettre plus en exergue le contrôle de l'hôte. La relation entre les acanthaires et *Phaeocystis* serait plutôt du parasitisme inverse: c'est le plus grand, l'hôte, qui parasite le petit, le symbiote. Des protistes hétérotrophes opportunistes comme les acanthaires « profiteraient » du succès écologique d'espèces phytoplanctoniques en s'associant avec une toute petite partie de leur population. Le terme mutualisme n'est donc pas approprié dans ce cas et nous proposons plutôt le terme de cytokleptie (littéralement vol de cellule : *cyto*- cellule et *klepto*-vol) pour qualifier la symbiose acanthaire-*Phaeocystis*.

« *La science va sans cesse se raturant elle-même. Ratures fécondes*.» Victor Hugo

Références bibliographiques

Références citées dans l'introduction et la discussion

Ahmadjian V (1988) The lichen alga *Trebouxia*: does it exist free living ? *Plant Systematics and Evolution* 158: 243-247.

Alexander E, Stock A, Breiner HW, Behnke A, Bunge J, Yakimov MM, Stoeck T (2009) Microbial eukaryotes in the hypersaline anoxic L'Atalante deep-sea basin. *Environ Microbiol* 11: 360-381.

Amaral-Zettler L, Sogin ML, Caron DA (1997) Phylogenetic relationships between the Acantharea and the Polycystinea: a molecular perspective on Haeckel's Radiolaria. *Proc Natl Acad Sci USA* 94: 11411-11416.

Anderson OR (1993) The trophic role of planktonic foraminifera and radiolaria. *Mar. Microbial Fd. Webs* 7: 31-51.

Anderson OR, Matsuoka A (1992) Endocytoplasmic microalgae and bacteroids within the central capsule of the radiolarian *Dictyocoryne truncatum*. *Symbiosis* 12: 237-247.

Anderson OR, Perry CC, Hughes NP (1990) Transmission and scanning electron microscopic evidence for cytoplasmic deposition of strontium sulphate crystals in colonial radiolaria. *Philos Trans R Soc Lond B* 329: 81-86.

Anderson OR (1983) Ed., *Radiolaria* (Springer-Verlag, NewYork, 355 pp).

Anderson OR (1976) A cytoplasmic fine structure study of two spumellarian radiolaria and their symbionts. *Mar Micropaleontol* 1: 81-99.

Anthony KRN, Hoogenboom MO, Connolly SR (2005) Adaptive variation in coral geometry and the optimization of internal colony light climates. *Functional Ecology* 19: 17–26.

Antia AN, Bauerfeind E, Von Bodungen B, Zeller U (1993) Abundance, encystment and sedimentation of Acantharia during autumn 1990 in the East Greenland Sea. *J Plankton Res* 15: 99–114.

Bachy C, Gómez F, López-García P, Dolan JR, Moreira D (2012) Molecular Phylogeny of tintinnid ciliates (tintinnida, ciliophora). *Protist* 163: 873-887.

Baker AC, Romanski AM (2007) Multiple symbiotic partnerships are common in scleractinian corals, but not in octocorals: Comment on Goulet (2006). *Mar Ecol Prog Ser* 335: 237-242.

Baker AC (2003) Flexibility and specificity in coral-algal symbiosis: diversity, ecology, and biogeography of *Symbiodinium*. *Annual Review of Ecology, Evolution, and Systematics* 34: 661-689.

Baker AC, Rowan R, Knowlton N (1997) Symbiosis ecology of two Caribbean acroporid corals. Procedings of the Eighth International Coral Reef Symposium, vol. 2, pp. 1295–1300.

Barnes DJ, Lough JM (1996) Coral skeletons: Storage and recovery of environmental information. *Global Change Biology* 2: 569-582.

Bhattacharya D, Archibald JM, Weber AP, Reyes-Prieto A (2007) How do endosymbionts become organelles? Understanding early events in plastid evolution. *Bioessays* 29: 1239-1246.

Bé AWH, Anderson OR, Faber WW, Caron DA (1983) Sequence of morphological and cytoplasmic changes during gametogenesis in the planktonic foraminifer *Globigerinoides sacculifer* (Brady). *Micropaleontology* 29: 310-325.

Bé AWH, Spero HJ, Anderson OR (1982) Effects of Symbiont Elimination and Reinfection on the Life Processes of the Planktonic Foraminifer *Globigerinoides sacculifer*. *Marine Biology* 70: 73-86.

Bé AWH, Caron DA, Anderson OR (1981) Effects of feeding frequency on life processes of the planktonic foraminifer Globigerinoides sacculifer in laboratoty culture. *J Mar Biol Ass UK* 61: 257-277.

Beers JR, Stewart GL (1970) The preservation of acantharians in fixed plankton samples. *Limnol Oceanogr* 15:825-827.

Bernhard JM, Bowser SS (1999) Benthic foraminifera of dysoxic sediments: chloroplast sequestration and functional morphology. *Earth Sci Rev* 46: 149-165.

Bernstein R, Kling SA, Boltovskoy D (1999) Acantharia. In Boltovoskoy D (ed) South Atlantic Zooplankton, vol.1. Backhuys Publishers, Leiden, The Netherlands, pp 75-147.

Bernstein RE, Betzer PR, Feely RA, Byrne RH, Lamb MF, Michaels AF (1987) Acantharian fluxes and strontium to chlorinity ratios in the North Pacific Ocean. *Science* 237: 1490-1494.

Bijma J, Erez J, Hemleben C (1990) Lunar and semi-lunar reproductive cycles in some spinose planktonic foraminifers. *J Foraminif Res* 20: 117-127.

Birrell CL, McCook LJ, Willis BL, Diaz-Pulido GA (2008) Effects of benthic algae on the replenishment of corals and the implications for the resilience of coral reefs. *Oceanogr Mar Biol Ann Rev* 46: 25-63.

Bonfante P (2003) Plants, Mycorrhizal Fungi and Endobacteria: a Dialog Among Cells and Genomes. *Biol Bull* 204: 215-220.

Bronstein JL (2003) The scope for exploitation within mutualistic interactions. In Genetic and cultural evolution of cooperation (ed. P. Hammerstein), pp. 185–202. Cambridge, MA: MIT Press.

Brandt K (1881) Über das Zusammenleben von Algen und Tieren, *Biologisches Zentrallblatt* 1: 524-527.

Braun A, Chen J, Waloszek D, Maas A (2007) in First early Cambrian Radiolaria. (Geol Soc, London, Spec Publ) 286: 143-149.

Brown WL, Wilson EO (1956) Character displacement. *Syst Zool* 5: 49-64.

Carlos AA, Baillie BK, Maruyama T (2000) Diversity of dinoflagellate symbionts (zooxanthellae) in a host individual. *Mar Ecol-Prog Ser* 195: 93-100.

Carroll SB (2001) Chance and necessity: the evolution of morphological complexity and diversity. *Nature* 409: 1102-1109.

Caron DA (2000) Symbiosis and mixotrophy among pelagic microorganisms. In: Kirchman, D.L. (Ed.), *Microbial ecology of the oceans*, John Wiley & Sons, Inc., New York, pp. 495-523.

Caron DA, Bé AWH, Anderson OR (1981) Effects of variations in light intensity on life processes of the planktonic foraminifer Globigernoides sacculifer in laboraroty culture. *J Mar Biol Ass UK* 62: 275-287.

Caullery M (1922) Le parasitisme et la symbiose. Paris, Doin.

Cavalier-Smith T, von der Heyden S (2007) Molecular phylogeny, scale evolution and taxonomy of centrohelid heliozoan. *Molecular Phylogenetics and Evolution* 44: 1186-1203.

Cavalier-Smith T, Lee JJ (1985) Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. J Protozool 32: 376-379.

Cienkowski L (1871) Über Schwarmer-Bildung bei Radiolarien. *Archiv für Mikroskopische Anatomie* 7: 372-381.

Coffroth MA, Lewis CF, Santos SR.Weaver JL (2006) Environmental populations of symbiotic dinoflagellates in the genus Symbiodinium can initiate symbioses with reef cnidarians. *Current Biology* 16: 985-987.

Coleman AW (2007) Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Res* 35: 3322-3329.

Cuvelier ML, et al. (2010) Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc Natl Acad Sci USA* 107: 14679-14684.

Darwin CR (1859) On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. London: John Murray edition.

Dorrell RG, Howe CJ (2012) What makes a chloroplast? Reconstructing the establishment of photosynthetic symbioses. *J Cell Sci* 125: 1865-1875.

Douglas AE (2001) Algal symbioses. Encyclopedia of Life Sciences. DOI: 10.1038/npg.els.0000327

De Bary A (1879) De la symbiose. *Revue internationale des Sciences* III, 301-309.

De Dekker P (2004) On the celestite-secreting Acantharia and their effect on seawater strontium to calcium ratios. *Hydrobiologia* 517: 1-13.

De Villiers S (1999) Seawater strontium and Sr/Ca variability in the Atlantic and Pacific oceans. *Earth Planet Sci Lett* 171: 623-634.

Doering M, Piercey-Normore MD (2009) Genetically divergent algae shape an epiphytic lichen community on jack pine in Manitoba. *Lichenologist* 41: 69-80.

Dolan JR (2012) Tintinntinid Ciliates: an Introduction and Overview. in *Biology and Ecology of Tintinnid Ciliates: Models for Marine Plankton*, Dolan, J.R., Agatha, S., Coats, D.W., Montagnes, D.J.S., Stocker, D.K. (eds). Wiley-Blackwell, Oxford, in press.

Donia MS, Fricke WF, Partensky F, Cox J, Elshahawi SI, White JR, Phillippy AM, Schatz MC, Piel J, Haygood MG, Ravel J, Schmidt EW (2011) Complex microbiome underlying secondary and primary metabolism in the tunicate-*Prochloron* symbiosis. *Proc Natl Acad Sci USA*. 108: E1423-1432.

Douglas AE (1998) Host benefit and the evolution of specialization in symbiosis. *Heredity* 81: 599-603.

Douglas AE (1994) Symbiotic interactions. Oxford University Press, Oxford, UK.

Douglas AE and Smith DC (1999) Are endosymbioses mutualistic? *TREE* 4: 350-352.

Edgcomb VP, Kysela DT, Teske A, Gomez AD, Sogin ML (2002) Benthic eukaryoticdiversity in the Guaymas Basin hydrothermal vent environment. *Proc Natl Acad Sci USA* 99: 7658-7662.

Edvardsen B, Imai I (2006) The ecology of harmful flagellates within Prymnesiophyceae and Raphidophyceae. In: Granéli E, Turner JT (eds) Ecology of harmful algae, ecological studies, Vol. 189. Springer, Berlin, pp 67–79

Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. *Nature* 440: 623-630.

Faber WW Jr, Anderson OR, Lindsey JL, Caron DA (1988) Algal-foraminiferal symbiosis in the planktonic foraminifer Globigerinella aequilateralis: I. Occurrence and stability of two mutually exclusive chrysophyte endosymbionts and their ultrastructure. *Journal of Foraminiferal Research* 18: 334-343.

Fairbanks RG, Evans MN, Rubenstone JL, Mortlock RA, Broad K, Moore MD, Charles CD. (1997) Evaluating climate indices and their geochemical proxies measured in corals. Proc. 8th Int. *Coral Reef Symp*. 1:107-116.

Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR (2004) the evolution of modern eukaryotic phytoplankton. *Science* 305: 354-360.

Febvre J, Febvre C, Michaels A (2000) Acantharia Haeckel, 1881. In Lee JJ, Leedale GF, Bradbury P (Eds) An Illustrated Guide to the Protozoa (2nd edition). Organisms Traditionally Referred to As Protozoa, or Newly Discovered Groups, Society of Protozoologists, Kansas, 783-803.

Febvre J (1989) Phylum Actinopoda, Class Acantharia. In Marulis L, Corliss JO, Melkonian M, Chapman DJ (Eds) Handbook of Protoctista, Jones and Bartlett Publishers, Boston, pp 363-379.

Febvre J, Febvre-Chevalier C (1989) Motility processes in the Acantharia. II. A Ca^{2+}-dependent system of contractile 2-4 nm filaments isolated from demembranated myonemes. *Biol Cell* 67: 243-249.

Febvre-Chevalier C, Febvre J (1986) Motility mechanisms in the actinopods (Protozoa): A review with particular attention to axopodial contraction/extension and movement of nonactin filamen systems. *Cell Motil Cytoskeleton* 6: 198-208.

Febvre J (1981) The myoneme of the Acantharia (Protozoa): a new model of cellular motility. *BioSystems* 14: 327-336.

Febvre J (1977) La division nucléaire chez les acanthaires. *Journal of Ultrastructure Research* 60: 279-295.

Field CB, Behrenfeld MJ, Randerson JT, Falkowski PG (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281: 237-240.

Flakowski J, Bolivar I, Fahrini J, Pawlowski J (2005) Actin phylogeny on Foraminifera. *Journal of Foraminiferal Research* 35: 93-102.

Foster RA, Collier JL, Carpenter EJ (2006a) Reverse transcription PCR amplification of cyanobacterial symbiont 16S rRNA sequences from single non-photosynthetic eukaryotic marine planktonic host cells. *Journal of Phycology* 42: 243-250.

Foster RA, Carpenter EJ, Bergman B (2006b) Unicellular cyanobionts in open ocean dinoflagellates, radiolarians, and tintinnids: ultrastructural characterization and immuno-localization of phycoerythrin and nitrogenase. *Journal of Phycology* 42: 453-463.

Frank SA (1996) Host-symbiont conflict over the mixing of symbiotic lineages. *Proc Biol Sci* 263: 339-344.

Frank AB (1877) Über die biologischen Verhältnisse des Thallus einiger Krusten-Flechten. *Beiträge zur Biologie der Pflanzen* 2: 132-200.

Galis F (2001) Key innovations and radiations. In *The Character Concept in Evolutionary Biology*, ed. GP Wagner, pp. 581−605. San Diego: Academic

Gast RJ, Caron DA (2001) Photosymbiotic associations in planktonic foraminifera and radiolaria. *Hydrobiologia* 461: 1-7.

Gast RJ, McDonnell TA, Caron DA (2000) srDNA-based taxonomic affinities of algal symbionts from a planktonic foraminifer and a solitary radiolarian. *Journal of Phycology* 36: 172-177.

Geddes P (1878) Sur la fonction de la chlorophylle avec les Planaires vertes. *Comptes rendus des séances de l'Académie des sciences de Paris*, LXXXVII 1095-1097.

Genkai-Kato M, Yamamura N (1999) Evolution of mutualistic symbiosis without vertical transmission. *Theoretical Population Biology* 55: 309-323.

Gleason DF, Danilowicz BS, Nolan CJ (2009) Reef waters stimulate substratum exploration in planulae from brooding Caribbean corals. *Coral Reefs* 28: 549-554.

Goldshmid R, Holzman R, Weihs D, Genin A (2004) Aeration of corals by sleep-swimming fish. *Limnol Oceanog* 49: 1832-1839.

Gontcharov AA, Melkonian M (2011) A study of conflict between molecular phylogeny and taxonomy in the Desmidiaceae (Streptophyta, Viridiplantae): analyses of 191 *rbc*L sequences. *Protist* 162: 253-267

Gowing MM, Coale SL (1989) Fluxes of living radiolarians and their skeletons along a northeast Pacific transect from coastal upwelling to open ocean waters. *Deep-Sea Research* 36: 561-576.

Grzebyk D, Schofield O, Vetriani C, Falkowski PG (2003) The Mesozoic radiation of eukaryotic algae: The portable plastid hypothesis. *J Phycol* 39: 259-267.

Grzymski J, Schofield OM, Falkowski PG, Bernhard JM (2002) The function of plastids in the deep-sea benthic foraminifer, *Nonionella stella. Limnol Oceanogr* 47: 1569-1580.

Haeckel E (1887a) Report on the Radiolaria collected by H.M.S. *Challenger* during the years 1873–1876. *Rep Sci Res Voy HMS Challenger Zool* 18: 1-1803.

Haeckel E (1887b) Die Radiolarien (Rhizopoda Radiaria). Eine Monographie. Zweiter Theil. Verlag von Gerog Reimer, Berlin, 248 pp

Haeckel E (1862) Die Radiolarien (Rhizopoda radiolaria). Eine Monographie. Berlin: G. Reimer.

Hamsher SE, Evans KM, Mann DG, Poulickova A, Saunders GW (2011) Barcoding diatoms: exploring alternatives to COI-5P. *Protist* 162: 405-422.

Hanlon RT, Bidwell JP, Tait R (1989) Strontium is required for statolith development and thus normal swimming behavior of hatchling cephalopods. *J Exp Biol* 141: 187-195.

Henjes J, Assmy P, Klaas C, Smetacek V (2007) Response of the larger protozooplankton to an iron-induced phytoplankton bloom in the Polar Frontal Zone of the Southern Ocean (EisenEx). *Deep Sea Res I* 54: 774-791.

Herre AE, Knowlton N, Mueller UG, Rehner SA (1999) The evolution of mutualisms: exploring the paths between conflict and cooperation. *Trends Ecol Evol* 14: 49-53.

Hoffmann LJ, Peeken I, Lochte K, Assmy P, Veldhuis M (2006) Different reactions of Southern Ocean phytoplankton size classes to iron fertilization. *Limnol Oceanogr* 51: 1217-1229.

Hollande A**,** Carré D (1974) Les xanthelles des radiolaires sphaerocollides, des acanthaires et de *Velella velella*: infrastructure-cytochimie-taxonomie. *Protistologica* 4: 573-601.

Hollande A, Martoja R (1974) Identification du cristalloïde des isospores de Radiolaires a un cristal de celestite (SrSO4). Détermination de la constitution du cristalloïde par voie cytochimique et à l'aide de la microsonde électronique et du microanalyseur à émission ionique secondaire. *Protistologica* 10 : 603-609.

Hollande A, Cachon J, Cachon-Enjumet M (1965) Les modalités de l'enkystement présporogénétique chez les acanthaires. *Protistologica* 1: 91-112.

Hörtnagl PH, Sommaruga R (2007) Photo-oxidative stress in symbiotic and aposymbiotic strains of the ciliate *Paramecium bursaria* . *Photochemical and Photobiological Sciences* 6: 813-920.

Jardillier L, Zubkov MV, Pearman J, Scanlan DJ (2010) Significant CO2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. *ISMEJ* 4: 1180-1192.

Jeffrey SW, MacTavish HS, Dunlap WC, Vesk M, Groenewould K (1999) Occurrence of UVA- and UVB-absorbing compounds in 152 species (206 strains) of marine microalgae. *Mar Ecol Prog Ser* 189: 35-51.

Kaniewska P, Anthony K, Hoegh-Guldberg O (2008) Variation in colony geometry modulates internal light levels in branching corals, *Acropora humilis* and *Stylophora pistillata*. *Marine Biology* 155: 649-660.

Keeling PJ (2009) Chromalveolates and the evolution of plastids by secondary endosymbiosis. *Journal of Eukaryotic Microbiology* 56: 1-8.

Keeling PJ (2004) Diversity and evolutionary history of plastids and their hosts. *American Journal of Botany* 91: 1481-1493.

Knowlton N, Rohwer F (2003) Microbial mutualisms on coral reefs. *American Naturalist* 162: 51-62.

Koide R (1985) The nature of growth depressions in sunflower caused by vesicular arbuscular infection. *New Phytologist* 99: 449-462.

Kremp A, Rengefors F, Montresor M (2009) Species specific encystment patterns in three Baltic cold-water dinoflagellates: the role of multiple cues in resting cyst formation. *Limnol Oceanogr* 54: 1125-1138.

LaJeunesse TC, Lambert G, Andersen RA, Coffroth MA, Galbraith DW (2005) *Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates. *Journal of Phycology* 41: 880–886.

LaJeunesse T (2001) Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus Symbiodinium using the ITS region: in search of a'' species'' level marker. *Journal of Phycology* 37: 866-880.

Lampitt RS, Salter I, John D (2009) Radiolaria: major exporters of organic carbon to the deep ocean. *Global Biogeochem Cycles* 23 GB1010, doi:10.1029/2008GB003221

Larson DW (1987) The absorption and release of water by lichens. In: Peveling, E. _Ed.., Progress and Problems in Lichenology in the Eighties. Bibliothca lichenologica 25 J. Cramer, Berlin, pp. 351-360.

Lea D, Martin P, Chan DA, Spero HJ (1995) Calcium uptake and calcification rate in the planktonic foraminifer *Orbulina universa*. *Journal of Foraminifera Research* 25: 14-23.

Lee JJ, Cervasco MH, Morales J, Billik M, Fine M, Levy O (2010) symbiosis drove cellular evolution. *Symbiosis* 51: 12-25.

Lesser MP, Weis VM, Patterson MR, Jokiel PL (1994) Effects of morphology and water motion on carbon delivery and productivity in the reef coral, *Pocillopora damicornis* (Linnaeus): Diffusion barriers, inorganic carbon limitation and biochemical plasticity. *J Exp Mar Biol Ecol* 178: 153-179.

Li YF, OuyangJH, Zhou Y,Liang XS, Zhong JY (2008) Facile fabrication of SrSO4 nanocrystals with different crystallographic morphologies via a simple surfactant-free aqueous solution route. *Materials Letters* 62: 4417-4420.

Little AF, van Oppen MJH, Willis BL (2004) Flexibility in algal endosymbioses shapes growth in reef corals. *Science* 304: 1492-1494.

Littman RA, van Oppen MJH, Willis BL (2008) Methods for sampling free-living *Symbiodinium* (zooxanthellae) and their distribution and abundance at Lizard Island (Great Barrier Reef). *Journal of Experimental Marine Biology and Ecology* 364: 48-53.

Liu H, et al. (2009) Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *Proc Natl Acad Sci USA* 106: 12803-12808.

Lombard F, Labeyrie L, Michel E, Spero HJ, Lea D (2009) Modelling the temperature dependent growth rates of planktic foraminifera. *Marine Micropaleontology* 70: 1-7.

López-García P, Philippe H, Gail F, Moreira D (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the mid-Atlantic ridge. *Proc Natl Acad Sci USA* 100: 697-702.

López-García P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409: 603-607.

Losos JB, Mahler DL (2010) Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. In *Evolution Since Darwin: The First 150 Years*, ed. M Bell, D Futuyma, Weanes, J Levinton, Chapter 15. Sunderland, MA: Sinauer.

Losos JB (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *American Naturalist* 175: 623−639.

Lovejoy C, Massana R, Pedrós-Alió C (2006) Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent seas. *Appl Environ Microbiol* 72: 3085-3095.

Lücking R, Lawrey JD, Sikaroodi M, Gillevet PM, Chaves JL, Sipman HJ, Bungartz F (2009) Do lichens domesticate photobionts like farmers domesticate crops? Evidence from a previously unrecognized lineage of filamentous cyanobacteria. *American Journal of Botany* 96: 1409-1418.

Lynch M (2008) The Cellular, Developmental and Population-Genetic Determinants of Mutation-Rate Evolution. *Genetics* 180: 933-943.

Lynch M (2006) Streamlining and Simplification of Microbial Genome Architecture. *Annu Rev Microbiol* 60: 327-349.

Mc Adam M and Balzer I (2004) Algal symbiosis in flatworms. *Symbiosis* 4: 559-574.

McShea DW (1996) Metazoan complexity and evolution: is there a trend? *Evolution* 50: 477-492.

Margulis L, Fester R (1991) Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis. The MIT Press.

Margulis L (1970) Origin of Eukaryotic Cells. Yale University Press, New Haven.

Marie D, Shi XL, Rigaut-Jalabert F, Vaulot D (2010) Use of flow cytometric sorting to better assess the diversity of small photosynthetic eukaryotes in the English Channel. *FEMS Microbiol Ecol* 72: 165-178.

Martin P, Allen JT, Cooper MJ, Johns DG, Lampitt RS, Sanders R, Teagle DAH (2010) Sedimentation of acantharian cysts in the Iceland Basin: Strontium as a ballast for deep ocean particle flux, and implications for acantharian reproductive strategies. *Limnol Oceanogr* 55: 604-614.

Massera Bottazzi E, Andreoli MG (1981) ‘‘Blooming’’of Acantharia (Protozoa-Sarcodina) in the southern Atlantic Ocean. *Quad Lab Technol* 3: 637.

Massera Bottazzi E (1973) Ulteriori ritrovamenti di cisti di Acantari (protozoa). *Mt Lomb Sci Lett* 107B: 3-26.

Mayr E (1963) Animal Species and Evolution. Belknap Press, Cambridge, MA.

Maynard-Smith J (1989) Generating novelty by symbiosis. *Nature* 341: 284-285.

Maynard Smith J (1970) Time in evolutionary process. *Studium Generale* 23: 266–272.

Medlin LK, Kaczmarska I (2004) Evolution of the diatoms: V. Morphological and cytological support for the major clades and a taxonomic revision. *Phycologia* 43: 245-270.

Mereschkowsky K (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol Centralbl* 25: 593-604 (addendum in 25, 689-691).

Michaels AF, Caron DA, Swanberg NR, Howse FA, Michaels CM (1995) Planktonic sarcodines (acantharia, radiolaria, foraminifera) in surface waters near Bermuda - abundance, biomass and vertical flux. *J Plankton Res* 17: 131-163.

Michaels AF (1991) Acantharian abundance and symbiont productivity at the VERTEX seasonal station. *Journal of Plankton Research* 13: 399-418.

Michaels AF (1988) Vertical distribution and abundance of Acantharia and their symbionts. *Marine Biology* 97: 559-569.

Mielck W (1907) Acanthometren von Neu-Pommern. Wissenschaftliche Meeresuntersuchungen, herausgegeben von der Kommission zur Untersuchung der deutschen Meere in Kiel und der Biologischen Anstalt auf Helgoland. Abteilung Kiel. Neue Folge 10: 41-105.

Moisan TA, Mitchell BG (2001) UV absorption by mycosporine-like amino acids in *Phaeocystis antarctica* Karsten induced by photosynthetically available radiation. *Marine Biology* 138: 217-227.

Moran NA (2007) Symbiosis as an adaptive process and source of phenotypic complexity. *Proc Natl Acad Sci USA.* 104: 8627-8633.

Moran NA (1996) Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc. Natl Acad Sci USA* 93: 2873-2878.

Moreira D, von der Heyden S, Bass D, López-García P, Chao E, Cavalier-Smith T (2007) Global eukaryote phylogeny: Combined small- and large-subunit ribosomal DNA trees support monophyly of Rhizaria, Retaria and Excavata. *Mol Phylogenet Evol* 44: 255-66.

Morris RM, et al. (2002) SAR 11 clade dominates ocean surface bacterioplankton communities. *Nature* 420: 806-810.

Müller J (1858) Über die Thalassicollen, Polycystinen und Acanthometren des Mittelmeeres. *Abh. Königl. Preuß. Akad.Wiss. Berlin* 1-62.

Mueller UG, Stephen A. Rehner, Ted R. Schultz (1998) The evolution of agriculture in ants. Science 281: 2034-2038.

Muscatine L (1990) The role of symbiotic algae in carbon and energy flux in reef corals. In: Dubinsky, Z. (Ed.), Ecosystems of the World: Coral Reefs. Elsevier, Amsterdam, pp. 75-87.

Nielsen SP (2004) The biological role of strontium. *Bone* 35: 583-588.

Nobre T, Aanen DK (2010) Dispersion and colonisation by fungus-growing termites. *Communicative & Integrative Biology* 3: 248-250

Norris RD (1996) Symbiosis as an evolutionary innovation in the radiation of Paleocene planktic foraminifera. *Paleobiology* 22: 461-480.

Not F, Gausling R, AzamF, Heidelberg JF, Worden AZ (2007) Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ Microbiol* 9: 1233-1252.

Nyholm SV, McFall-Ngai MJ (2004) The winnowing: establishing the squid-vibrio symbiosis. *Nat Rev Microbiol* 2: 632-642.

Pawlowski J, Burki F (2009) Untangling the phylogeny of amoeboid protists. *J Eukaryot Microbiol* 56: 16-25.

Pawlowski J, Holzmann M, Berney C, Fahrni J, Gooday AJ, Cedhagen T, Habura A, Bowser SS (2003) The evolution of early foraminifera. *Proc Natl Acad Sci USA* 100: 11494-11498.

Pedrós-Alió C (2012) The rare bacterial biosphere. *Annu Rev Mar Sci* 4: 449-466.

Perru O (2006) Aux origines des recherches sur la symbiose vers 1868-1883. *Revue d'histoire des sciences* 59: 5-27.

Perry CC, Hughes NP (1990) Transmission and scanning electron microscopic evidence for cytoplasmic deposition of strontium sulphate crystals in colonial radiolaria. *Philos Trans R Soc Lond [Biol]* 329: 81-86.

Piganeau G, Eyre-Walker A, Grimsley N, Moreau H (2011) How and Why DNA Barcodes Underestimate the Diversity of Microbial Eukaryotes. *PLoS ONE* 6(2): e16342. doi:10.1371/journal.pone.0016342

Pillec Le, de Vargas C, Pawlowski J (2011) Molecular identification of sequestered diatom chloroplasts and kleptoplastidy in foraminifera. Protist 162: 394-404.

Pochon X, Gates RD (2010) A new Symbiodinium clade (Dinophyceae) from soritid foraminifera in Hawai'i. *Mol Phylogenet Evol* 56: 492-497.

Pochon X, Montoya-Burgos JI, Stadelmann B, Pawlowski J (2006) Molecular phylogeny, evolutionary rates, and divergence timing of the symbiotic dinoflagellate genus *Symbiodinium*. *Mol Phylogenet Evol* 38: 20-30.

Popofsky A (1904a) System und Faunistik der Acanthometriden der Planktonexpedition. Inaugural-Dissertation zur Erlangung der Doktorwürde der hohen philosophischen Fakultät der Königlichen Christian-Albrecht-Universtät zu Kiel, Kiel, pp 5-77**.**

Popofsky A (1904b) Die nordischen Acantharien. Teil 1: Acanthometriden. Ergebnisse der Plankton-Expedition der Humboldt-Stiftung 3(Lf. α): 3-158.

Popofsky A (1906) Die Acantharia der Plankton-Expedition Teil II: Acanthophracta. Ergebnisse der Plankton-Expedition der Humboldt-Stiftung 3(Lf. ß): 3-160.

Porter JW, Tougas JI (2000) Reef ecosystems: threats to biodiversity. *Encyclopedia Biodivers* 5: 73-95.

Porter S (2011) The rise of predators. *Geology* 39: 607-608.

Potts M (2000) Nostoc. In The Ecology of Cyanobacteria (Whitton, B.A. & Potts, M., editors), 465–504. Kluwer, Dordrecht.

Putman HM, Stat M, Pochon X, Gates R (2012) Endosymbiotic flexibility associates with environmental sensitivity in scleractinian corals. *Proc R Soc Lond B* doi: 10.1098/rspb.2012.1454.

Quaiser A, Zivanovic Y, Moreira D, Lopez-Garcia P (2010) Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. *ISMEJ* 5: 285-304.

Reef R, Kaniewska P, Hoegh-Guldberg O (2009) Coral Skeletons Defend against Ultraviolet Radiation. *PLoS ONE* 4: e7995. doi:10.1371/journal.pone.0007995.

Reshetnyak VV (1981) Akantarii (Acantharia, Protozoa) Mirovogo okeana. Fauna SSSR 123: 1-224. (in Russian).

Reyes-Prieto A, Bhattacharya D (2007) Phylogeny of Nuclear Encoded Plastid Targeted Proteins Supports an Early Divergence of Glaucophytes within Plantae. *Mol Biol Evol* 24: 2358-2361.

Richardson SL (2001) Endosymbiont change as a key innovation in the adaptive radiation of Soritida (Foraminifera). *Paleobiology* 27: 262-289.

Rowan, R., Powers, D.A., 1992. Ribosomal RNA sequences and the diversity of symbiotic dinoflagellates, (zooxanthellae). *Proc Natl Acad Sci USA* 89: 3639-3643.

Sachs JL, Skophammer RG, Regus J (2011) Evolutionary transitions in bacterial symbiosis. *Proc Natl Acad Sci USA* 108: 10800-10807.

Sachs JL, Wilcox TP (2006) A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*. *Proc R Soc B* 273: 425-429.

Sachs JL, Mueller UG, Wilcox TP, Bull JJ (2004) The evolution of cooperation. *Q Rev Biol* 79: 135-160.

Salamini F, Ozkan H, Brandolini A, Schäfer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the near east. *Nat Rev Genet* 3(6): 429-441.

Santoferrara LF, McManus GB, Alder VA (2012) Utility of genetic markers and morphology for species discrimination within the order Tintinnida (Ciliophora, Spirotrichea). *Protist in press*.

Sapp J (1994) Evolution by association. Oxford University Press, New York.

Schiebel R (2002) Planktic foraminiferal sedimentation and the marine calcite budget. *Glob. Biogeochem Cycles* 16: 1065. doi:10.1029/2001GB001459.

Schwendener S (1868) Untersuchungen über den Flechtenthallus, *Beiträge zur wissenschaftliche Botanik*, VI 195-207.

Selosse MA (2000). La Symbiose. Vuibert, Paris.

Schaap P, Winckler T, Nelson M, Alvarez-Curto E, Elgie B, Hagiwara H, Cavender J, Milano-Curto A, Rozen DE, Dingermann T, Mutzel R, Baldauf SL (2006) Molecular phylogeny and evolution of morphology in the social amoebas. *Science* 314: 661-663.

Schewiakoff WT (1926) The Acantharia. Fauna e Flora del Golfo di Napoli 37: 1-755.

Schluter D (1988) Character displacement and the adaptive divergence of finches on islands and continents. *American Naturalist* 131: 799-824.

Shaked Y, de Vargas C (2006) Pelagic photosymbiosis: rDNA assessment of diversity and evolution of dinoflagellate symbionts and planktonic foraminiferal hosts. *Marine Ecology Progress Series* 325: 59-71.

Shi XL, Lepère C, Scanlan DJ, Vaulot D (2011) Plastid 16S rRNA gene diversity among eukaryotic picophytoplankton sorted by flow cytometry from the South Pacific Ocean. *PLoS One* 6:e18979.

Siano R champion, Montresor M, Probert I, Not F, de Vargas C (2010) *Pelagodinium* gen. nov and *P-béii* comb. nov., a dinoflagellate symbiont of planktonic Foraminifera. *Protist* 161: 385-399.

Simó R (2001) Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and evolutionary links. *TRENDS in Ecology & Evolution* 16: 287-294.

Simpson GG (1953) The Major Features of Evolution. Columbia University Press, New York.

Smetacek V (2001) A watery arm race. *Nature* 411: 744-745.

Smetacek VS (1985) Role of sinking in diatom life-history cycles: ecological, evolutionary and geological significance. *Marine Biology* 84: 239-251.

Smith DC and Douglas AE (1987) *The Biology of Symbiosis*. Edward Arnold.

Stanley GD (2006) Photosymbiosis and the evolution of modern coral reefs. *Science* 312:857-858.

Stimson J, Sakai K, Sembali H (2002) Interspecific comparison of the symbiotic relationship in corals with high and low rates of bleaching-induces mortality. *Coral Reefs* 21: 409-421.

Spero H.J (1987). Symbiosis in the planktonic foraminifer, *Orbulina universa*, and the isolation of its symbiotic dinoflagellate, *Gymnodinium béii* sp. nov.. *Journal of Phycology* 23: 307-317.

Spero H.J, Parker SL (1985) Photosynthesis in the symbiotic planktonic foraminifer *Orbulina universa*, and its potential contribution to oceanic primary productivity. *J Foram Res* 15: 273-281.

Stoeck T, Taylor GT, Epstein SS (2003) Novel eukaryotes from the permanently anoxic Cariaco Basin (Caribbean Sea). *Appl Environ Microbiol* 69: 5656-5663.

Stoecker DK, Johnson MD, de Vargas C, Not F (2009) Acquired phototrophy in aquatic protists. *Aquat Microb Ecol* 57: 279-310.

Stoecker DK, Gustafson DE and Verity PG (1996) Micro- and mesoprotozooplankton at 1400W in the equatorial Pacific: heterotrophs and mixotrophs. *Aquat Microbial Ecol* 10: 273-282.

Sunda W, Kieber DJ, Kiene RP, Huntsman S (2002) An antioxidant function for DMSP and DMS in marine algae. *Nature* 418: 317-320.

Swanberg NR and Caron DA (1991) Patterns of feeding in epipelagic oceanic plankton. *J Plankton Res* 13: 287-312.

Taylor FJR (1982) Symbiosis in marine microplankton. *Annals Institute Oceanographique, Paris* 58: 61-90.

Thompson JN (1999) The evolution of species interaction. *Science* 284: 2116-2118.

Thompson D'AW (1917) *On Growth and Form* (Cambridge Univ. Press, Cambridge).

Tillmann U (2004) Interactions between planktonic microalgae and protozoan grazers: *The Journal of Eukaryotic Microbiology* 51: 156-168.

Tregouboff G (1953) Traité de Zoologie, ed. Grasse, Masson, Paris, pp 271-320

Trench R (1993) Microalgal-invertebrate symbioses – a review. *Endocytobiosis and Cell Research* 9: 135-175.

Van Beneden PJ (1875) Les commensaux et les parasites dans le règne animal. Baillières, Paris.

Van Bergeijk SA, Stal LJ (2001) Dimethylsulfoniopropionate and dimethylsulfide in the marine flatworm *Convoluta roscoffensis* and its algal symbiont. *Mar Biol* 138: 209-216.

Van Valen L (1973) *Evol Theor* 1: 1-30.

Venn A, Tambutté E, Holcomb M, Allemand D, Tambutté S (2011) Live Tissue Imaging Shows Reef Corals Elevate pH under Their Calcifying Tissue Relative to Seawater. *PLoS ONE* 6(5): e20013. doi:10.1371/journal.pone.0020013

Vermeij MJA, Marhaver KL, Huijbers CM, Nagelkerken I, Simpson SD (2010) Coral Larvae Move toward Reef Sounds. *PLoS ONE* 5: e10660. doi:10.1371/journal.pone.0010660.

Webster NS, Taylor MW, Behnam F, Lu¨ cker S, Rattei T, Whalan S et al. (2010). Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ Microb* 12: 2070-2082.

Wilkinson DM, Sherratt TN (2001) Horizontally acquired mutualisms, an unsolved problem in ecology? *Oikos* 92: 377-384.

Wilkinson DM and Dickinson NM (1995) Metal resistance in trees: the role of mycorrhizae. *Oikos* 72: 298–300.

Wooldridge SA (2010) Is the coral-algae symbiosis really “mutually beneficial” for the partners? *Bioessays* 32: 615-625.

Yacobovitch T, Benayahua Y, Weis VM (2004) Motility of zooxanthellae isolated from the Red Sea soft coral Heteroxenia fuscescens (Cnidaria). *J Exp Mar Biol Ecol* 298: 35-48.

Yamamura N (1997) Diversity and evolution of symbiotic interactions, in ``Biodiversity An Ecological Perspective" (T. Abe, S. A. Levin, and, M. Higashi, Eds.), pp. 75-82, Springer-Verlag, New York.

Yellowlees D, Rees TAV, Leggat W (2008) Metabolic interactions between algal symbionts and invertebrates hosts. *Plant Cell Environ* 31: 679-694.

Yost DM, Mitchelmore CL (2009) Dimethylsulfoniopropionate (DMSP) lyase activity in different strains of the symbiotic alga *Symbiodinium microadriaticum*. *Mar Ecol Prog Ser* 386: 61-70.

Young JR (1994) Functions of Coccoliths by, in *Coccolithophores* (Amos Winter and William Siesser, Eds, Cambridge University Press, New York).

Yuasa T, Horiguchi T, Mayama S, Matsuoka A, Takahashi O (2012) Ultrastructural and molecular characterization of cyanobacterial symbionts in *Dictyocoryne profunda* (polycystine radiolaria). *Symbiosis* doi:10.1007/s13199-012-0174-2

Zas’ko DN, Vedernikov VI (2003) A comparative analysis of the vertical distribution of epipelagic radiolaria, chlorophyll, and zooplankton in different areas of the North Atlantic Ocean (from June to September 2001). *Oceanol* 43:63-71.

Zwirglmaier K et al (2008) Global phylogeography of marine S*ynechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. E*nv Microbiol* 10: 147-161.

Remerciements

Je tiens tout d'abord à remercier la Région Bretagne qui a financé ce projet de thèse durant ces trois années (DIPHOPE), ainsi que le GIS Europôle Mer (Nadine Reniers) qui m'a permis d'assister à certaines conférences. Merci également à tous les laboratoires d'accueil pour la collecte d'échantillons (Stazione Zoologica de Naples, l'Interuniversity Institute for Marine Sciences d'Eilat, le laboratoire d'Akajima et de Villefranche-sur-Mer).

Mes premiers remerciements vont tout naturellement à mes deux mentors et héros de cette aventure, Fabrice et Colomban. Alors que j'errais tel un plancton à la recherche d'une thèse, ils m'ont proposé d'intégrer l'équipe EPPO, de me convertir à la protistologie marine, mais aussi et surtout de participer à cette aventure inoubliable Taratata. Je leur dois tous ces voyages et souvenirs inoubliables…

Fabrice, je ferai court sinon je vais être obligé d'écrire cette thèse en plusieurs volumes. Un immense merci pour ton accompagnement, ton encadrement, ta disponibilité, ton écoute, ta présence, tes conseils, ton pragmatisme, ton énergie, ton enthousiasme et j'en passe. Merci de m'avoir laissé une si belle liberté pour explorer ces acanthaires, qui n'en ont pas l'air, qui a rendu cette thèse si légère. Merci pour avoir cru et poussé sur ce papier Acanthaire-*Phaeocystis*. Merci pour toutes ses belles aventures dans la bonne humeur aux quatre coins du monde aussi bien autour de cartilages d'oreilles de cochons, dans la limousine de Roger Milla, dans les karaokés, qu'en pleine mer au milieu de nulle part. Difficile de se remémorer à cet ultime étape de la thèse tous ces bons moments...ils resteront.

Colomban, je me souviendrai toujours de notre discussion sur le muret face à la station où tu m'as parlé pour la première fois de Tara. Déjà 5 ans. Merci infiniement pour cette épopée unique, cette aventure humaine et scientifique, un rêve de gosse. Merci également pour ces escapades philosophiques et artistiques loin des acanthaires et de la Science parfois trop terne. Merci pour tes colombaneries, tes idées folles, ton ouverture et ta générosité.

Ian, tu es aussi mon héros d'outre manche, ma plume shackespearienne, un camarade de pêche planctonique à la nage mais aussi autour de matchs de sport. J'ai adoré nos escapades viriles et exigues à Villefranche. Je te remercie vivement pour tes conseils si précieux pour l'écriture d'articles. Même si tu n'a pas relevé le défi de cultiver ces sacrés *Phaeocystis*, je dois beaucoup à ton immense talent d'isolateur fou de microalgues symbiotiques. Ne laisse pas trop traîner quand même tes bières dans les incubateurs à microalgues. Allez Derby County !

Un grand merci à toute l'équipe EPPO pour leur aide, leur sympathie et leur accompagnement au cours de cette thèse. Merci à Stéphane, Marie-Jo, Noan Le Bes, Angélique (El Tiburon), Raphael foramimi, Christophe B, Nathalie B, Camille. Merci à Mahdi pour nos discussions folles sur l'évolution du plancton et des formes. Merci à Seb qui a refait le portrait des acanthaires. Tes photos sont magnifiques !!! Un immense merci à Fred, l'homme qui manie les millions de séquences plus vite que son ombre. Ton aide, ta disponibilité, ton désir de résoudre des défis, le tout dans la bonne humeur, ont largement contribué à cette thèse. Merci à toi Lucie pour ton aide, ta collaboration, ton énergie et ta vision. Sarah, « *Franchement, on est trop bien ensemble* ». Je suis d'accord. Merci merci merci pour ton aide et disponibilité au labo, les commandes, ton initiation à la pipette et à la rigueur au-dessus d'une paillasse. Je n'ai manqué de rien. *Franchement*, tu as énormément contribué au bon déroulement de cette thèse et à l'obtentien des résultats. Ca vaut bien un petit Martini ca !!!

Noritoshi, ありがたい !!! Thank you so much for your collaboration, your help, the huge efforts to identify all the acantharian cells and to elucidate the highly confused and complex taxonomy of Acantharia. Thank you for having transmitted me your impressive knowledge on the history of these fascinating organisms. It was really a pleasure to work with you.

Je tenais également à remercier toutes les personnes de l'équipe plancton qui m'ont permis de travailler dans des conditions optimales. Un merci tout particulier à Fabienne, Priscilla, Roseline, Florence, et à Daniel V (l'escapade à Aka Jima était génial !!!). Merci à Anne-Claire pour son soutien et nos échanges passionants sur *Phaeocystis*. Merci à Céline Manceau, tata, qui m'a également accompagné durant cette thèse et a donné tellement d'énergie et de temps pour toutes les tâches administratives. Ca vaut bien quelques kilos de chocolat suisse à la pistache. Merci à François Lallier, qui m'a fait découvrir ce bout du monde. Merci pour sa confiance et son support. Un merci à l'ensemble du personnel de la Station Biologique, en particulier Brigitte. Merci à Nicole Guyard pour la

recherche de vieux articles apparus avant l'ère numérique qui prennent la poussière dans les bibliothèques de France et de Navarre. Tous ces articles ont beaucoup compté dans ce travail.

Un grand merci à Eric Karsenti, et à toute l'équipe Tara (Agnes, Etienne, Romain, Loic, Hervé, François, Myriam, Eloise, Julien, Vincent, Daniel…), qui ont rendu ce projet fou et cette magnifique aventure réalisable, et ont permis une collecte merveilleuse de plancton qui fait fantasmer plus d'un biologiste marin …
Steffi, danke vielmals !!! Thank you so much for your support and friendship. I'm very grateful for taking care of me like a mum !!! Merci à Colette pour son partage des conaissances sur les acanthaires et nos entrevues sur la côte d'Azur. Merci à John Dolan pour son accueil chaleureux chaque année dans son laboratoire les pieds dans l'eau qui a permis de constituer très largement la belle collecte d'acanthaires. Merci à Jan Pawlowski.Merci à tous les auteurs qui ont contribué à ces travaux de thèses.

Je tiens également à saluer les amis qui ont aussi participé sans le savoir à cette thèse: Clément, Marine, Doudou, Camille, Bruno, Toti, Polo, Xavier, Gäelle C, Cécile, Loic le Pillec, Gäelle P, Steph…et ceux que j'oublie. Merci à mon grand champion zio Raph (merci de me laisser gagner au tennis), et à mon champion Christophe.

Mille grazie a Gabriella e Ernesto, e tutta la familia italiana. Grazie tante !!!

Je dédie cette thèse à mes parents, mes frères, mes grands-parents, ma grand-mère et à Michel parti trop tôt…
Enfin, Lucia, pour cette force incroyable qu'elle m'a donnée durant ces 3 années, nos échanges, sa générosité, sa patience, son aide précieuse, sa cuisine et son amour…

Annexes

Annexe 1
Liste des présentations orales et posters

Annexe 2
Planches taxonomiques des différents clades d'acanthaires

Annexe 3
Essai sur l'Histoire de la découverte des radiolaires

Annexe 4
Publications en collaboration

Annexe 1 : Liste des présentations orales et posters

Communications orales:

-7th Symposium Morphometry and Evolution of Forms - Mai 2012 - Lyon, France: « *L'évolution des formes dans le plancton marin unicellulaire: les Acanthaires véritables architectes du minéral* »

- 7ème réunion du Réseau Ecologie des Interactions Durables - Février 2012 - Rennes, France: « *Dr Jekyll and Mr Hyde: harmful and bloom-forming phytoplankton in symbiosis in worldwide oceans* »

- European Congress of Protistology - Juillet 2011 - Berlin, Germany: «*Molecular phylogeny and morphological evolution of the Acantharia*»

- The 18th meeting of the International Society for Evolutionary Protistology - Juillet 2010 - Kanazawa, Japan: « *Unveil the diversity and symbiosis in planktonic Acantharia*»

Posters:

- *7th International Symbiosis Society Congress* - Juillet 2012 - Krakow, Poland: « *An original mode of symbiosis in open ocean plankton*».

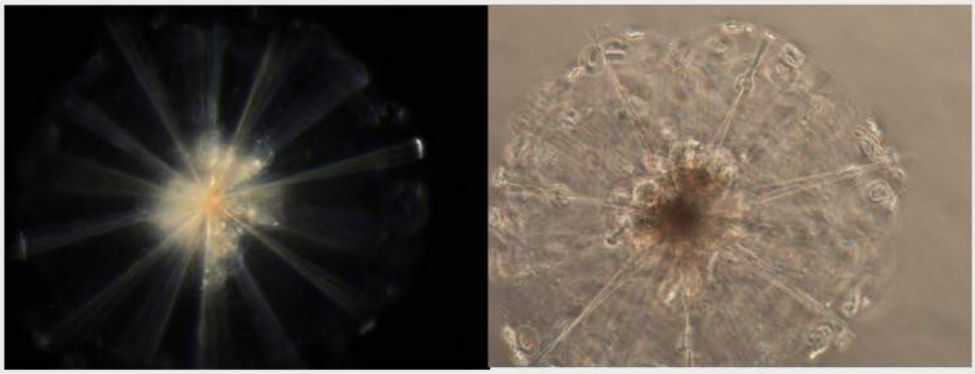
- Microorganisms facing their environment. - Octobre 2011 - Paris, France: « *Molecular phylogeny and evolutionary history of the Acantharia*»

- 7th International Congress of Systematic and Evolutionary Biology (ICSEB VII)- Février 2011- Berlin, Germany: « *Molecular phylogenetics and evolutionary history of planktonic Acantharia (Radiolaria)* ». ***Prize*** *of the best poster presentation*

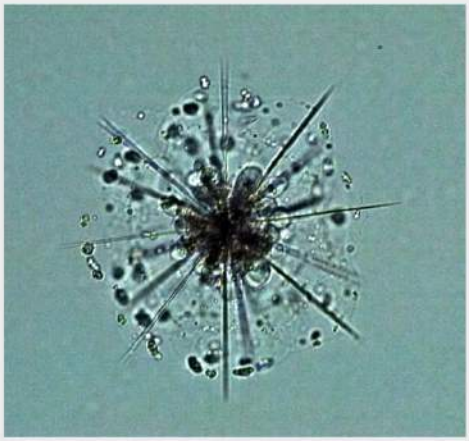
Annexe 2: Planches taxonomiques des différents clades d'acanthaires

Clade A

Order: Holacanthida
Family: Acanthoplegmidae



Ei 59 *Acanthoplegma* sp.



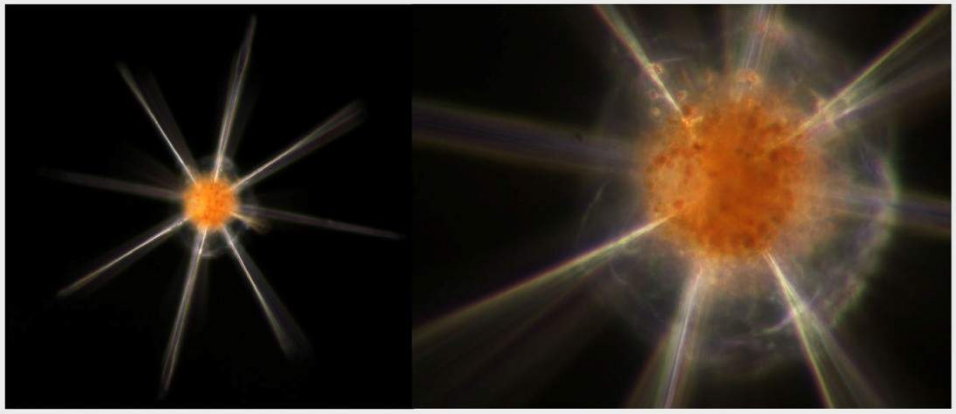
SES 25 *Acanthoplegma* sp.

Main morphological characters:

- 10 diametral spicules very thin and of equal length, simply crossing the cell center
- Presence of many hyaline plates (or spherical granules) on the surface of the periplasmic cortex
- No endosymbiotic microalgae
- Presence of bubble-like structures in the ectoplasm
- Axopodes and myonemes are not visible

Clade B1

Order: Holacanthida
Family:*incertae sedis*
Genus: *Acanthonidium*



Ei 68 - *Acanthonidium* sp.

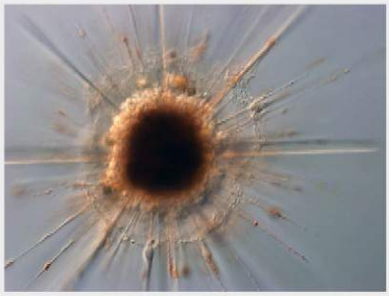
Elongated cysts found in mesopelagic and bathypelagic sediment traps

Main morphological characters:

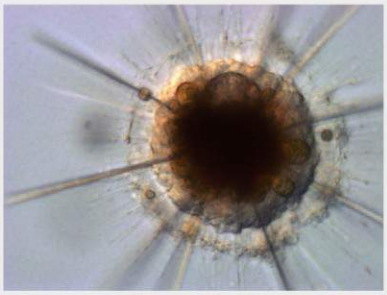
- Central junction not visible because of a dark reddish endoplasm
- Very long and thin spicules, of equal length
- Red inclusions in the small endoplasm (lipid storage?)
- No endosymbiotic microalgae

Clade B2

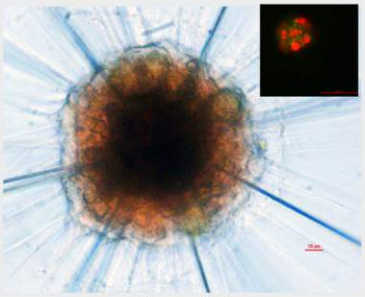
Order: Holacanthida
Family: Acanthochiasmidae
Genus: ***A****canthochiasma*
Species: *Acanthochiasma* sp.



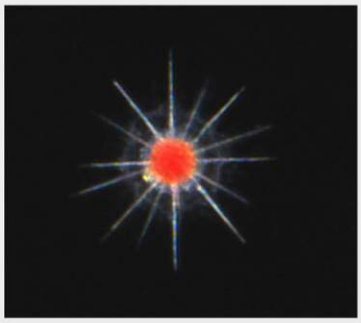
Pec 18



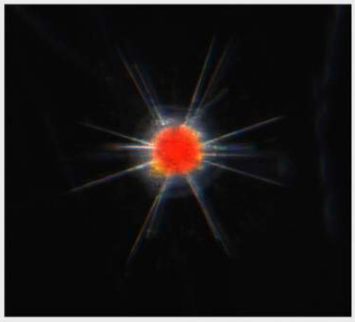
Acanth 23



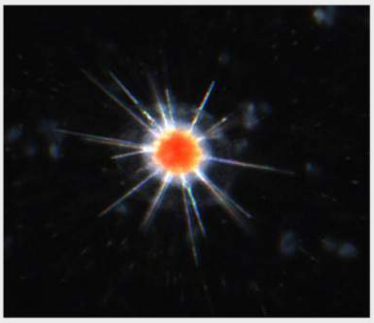
Pec 19



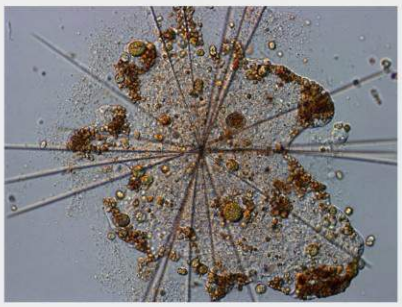
Vil 51



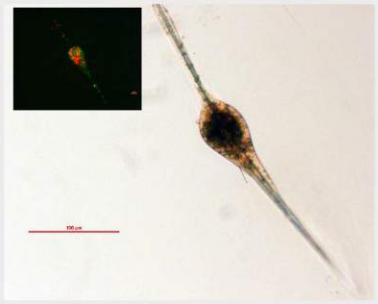
Vil 60



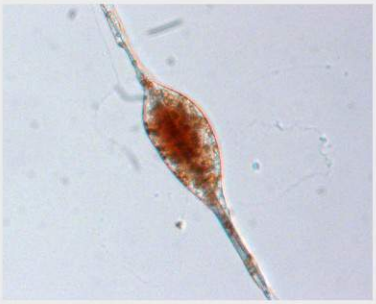
Vil 64



Pec 18



Pec 19



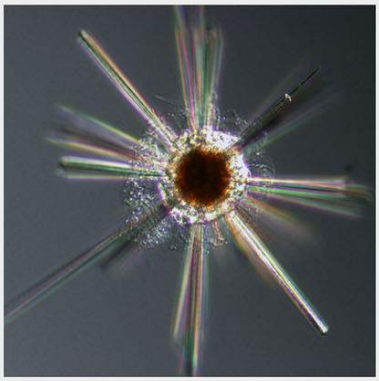
Pec 20

Main morphological characters:

- No central junction
- 10 diametral, long and very thin spicules of equal length
- Presence of many axopods, but myonemes are not visible
- Reddish endoplasm with microalgal cells inside of different sizes. Small insets of PEC 19 show the red autofluorescence of the chlorophyll
- Cells are very fragile and can trigger cyst formation when manipulated too severly (eg. Pec19)

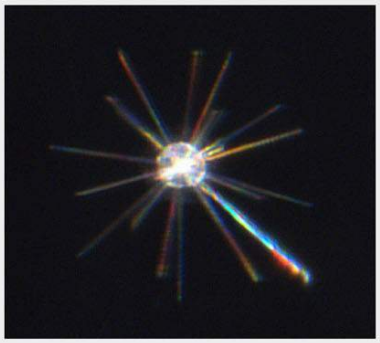
Clade C

Order: Chaunacanthida
Family: Stauraconidae, Litholophidae, Incertae sedis

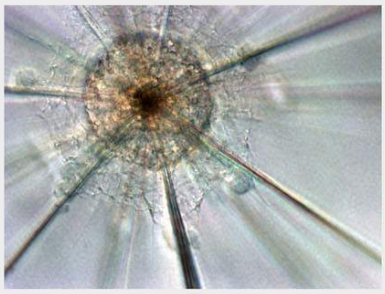


Vil 110 - sub-clade C1
Litholophus sp?

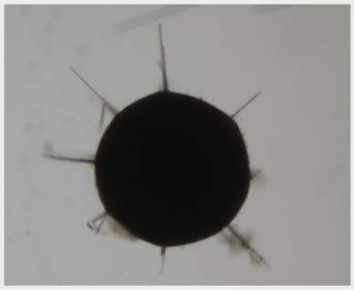
Pec 13 - sub-clade C2
Stauracon palida



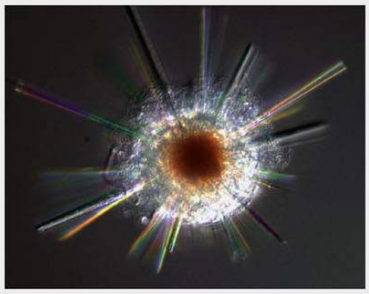
Vil 47 - sub-cladeC3
Heteracon biformis



Pec 16 - sub-clade C3
Heteracon biformis



Round cyst - subclade C3



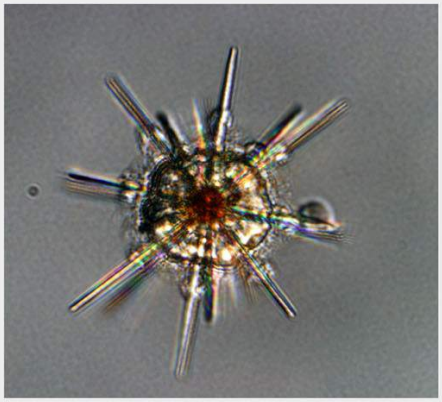
Vil105 - sub-clade C4
Litholophus muelleri

Main morphological characters:

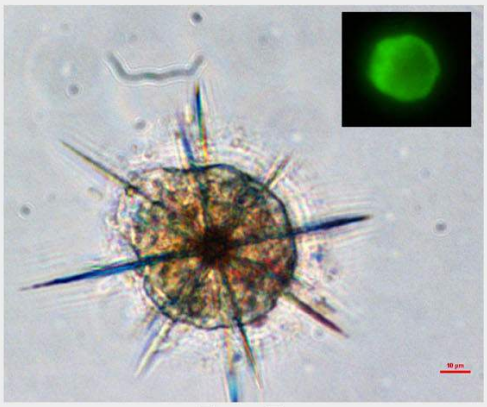
- Spicules can be of different length
- Central junction
- No apophyses (lateral extension) on each spicule but some species serrated spicules
- Cell center black or reddish (*Litholophus*) or very clear (*Stauracon*, *Heteracon*)
- Myonemes very long
- No endosymbiotic microalge

Clade IV

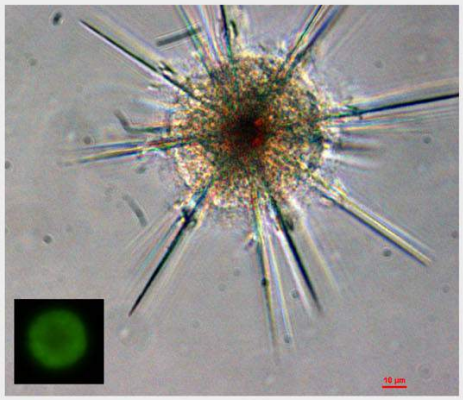
Order: New suborder?



Vil 119



Pec 5



Pec 10

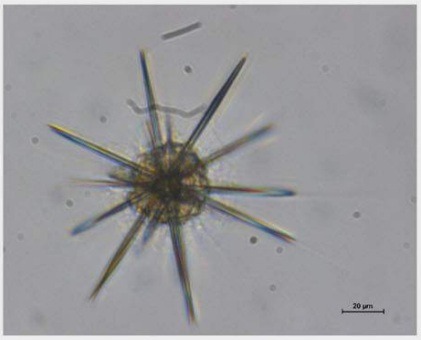
Main morphological characters:

- Central junction not visible because of a reddish zone in the endoplasm
- Simple and short spicules with no apophyses
- Presence of myonemes
- Capsular membrane well visible
- No symbiotic microalgae (no red autofluorescence of the chlorophyll shown by the small insets)

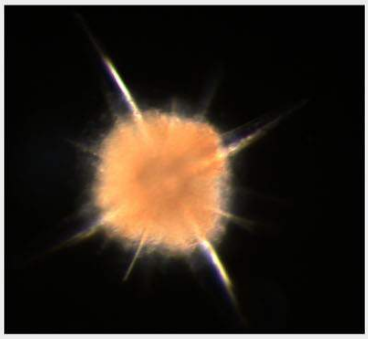
Incertae sedis: It was not possible to define the taxonomic position of these specimens (new species?)

Clade D

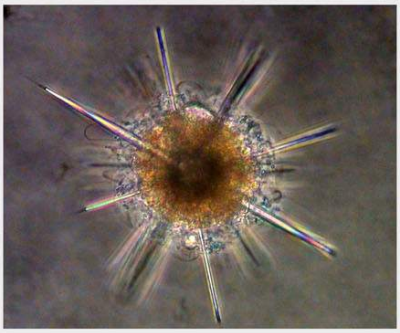
Order: Trizonacantha n. subord.
Family: Trizonidae and *Incertae sedis*
Genus: *Trizona* and *Staurolithium*



Ros 6 - clade D1
Trizona brandtii



Ei 84 - clade D2
***Staurolithium* sp.**



Oki 47 - clade D2
***Trizona* sp. A**

Pec 6 - clade D2
***Staurolithium* sp. B**

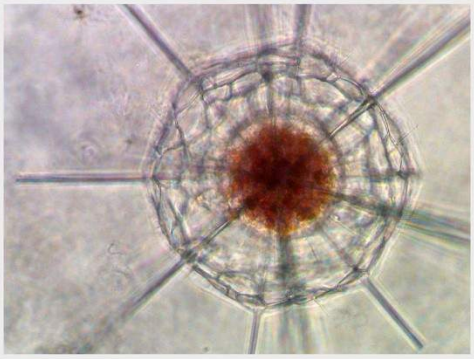
Main morphological characters:

- Ball-like central junction, simple and short spicules with no apophyses
- Spicules can be of equal length (*Trizona brandtii*) or 4 spicules can be longer and thicker (*Staurolithium* sp).
- Spicules can have lanceolate, sharp distal ends, sometimes with longitudinal keel
- No endosymbiotic microalgae
- Brown endoplasm and absence or very thin ectoplasm.
- Axopods

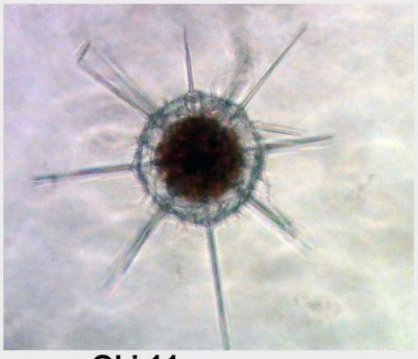
Clade E

Sub-clade E1-E2

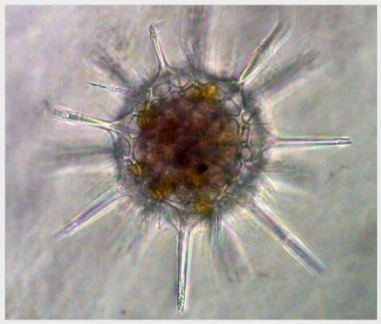
Order: Arthracanthida
Family: Tessaropelmidae, Aspidommidae, Lychnaspinae



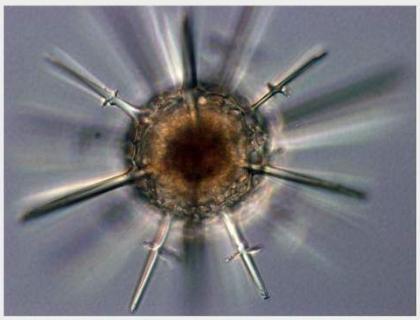
Oki 49 - sub-clade E1
Tessaropelma arachnoides



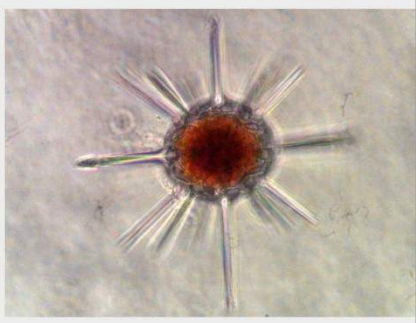
Oki 11 - subclade E1
Tessaropelma arachnoides



Oki 46 - sub-clade E2
Lychnaspis giltschi



Pec 3 - sub-clade E2
Larcidium dodecanthum



Oki 73 - sub-clade E2
Larcidium dodecanthum

Main morphological characters:

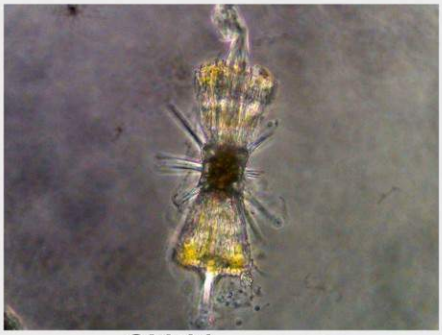
- Presence of a latticed shell, sometimes double shells. The pore-frame is a diagnostic character.
- Spicules can be of different length and thickness; Frizzy accessory spines (*Lychnaspis giltschi*)
- Presence of endosymbiotic microalgae
- Endoplasm generally brownish to reddsish

Clade E

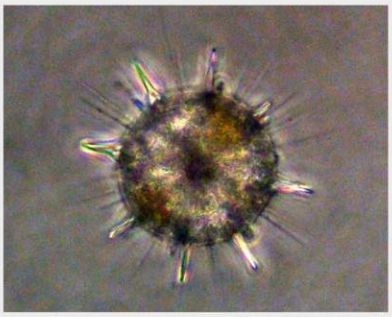
Sub-clade E3-E4

Order: Arthracanthida
Family: Diploconidae and Dorataspidae

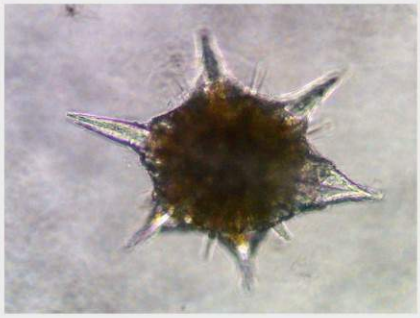
Ta 61 - sub-clade E3
Diploconus tridentatus



Oki 14 - subclade E3
Diploconus sp.

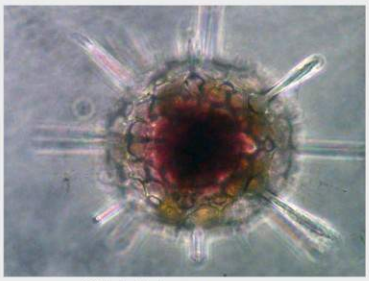


Oki 4 - sub-clade E4
Dorataspis loricata



Oki 32 - sub-clade E4
Coleaspis ciliatus

Vil 86 - sub-clade E4
Coleaspis ciliatus



Oki 74 - sub-clade E4
Dorataspis ? typica

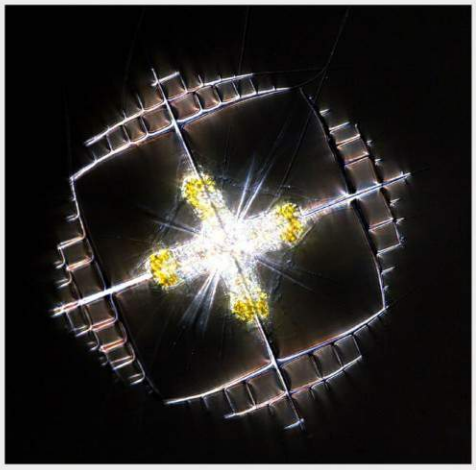
Main morphological characters:

- 20 radial spicules of different shape and length (2 can be longer and thicker like in *Diploconus* sp.)
- Very thick and heavy shell of different shapes (cylindrical, round) with no pores, except *Dorataspis* sp.
- Ectoplasm not visible, many axopods, 6-12 myonemes per spicule
- Presence of endosymbiotic microalgae

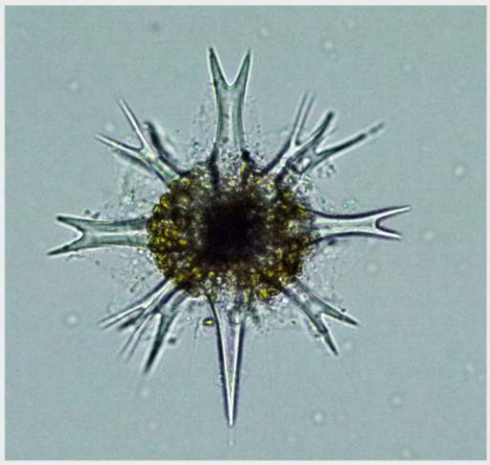
Clade F

Sub-clade F1 - F2

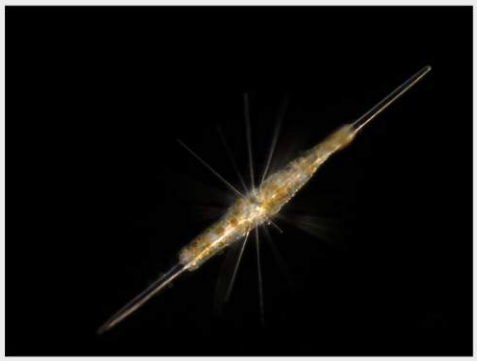
Order: Arthracanthida
Family: Lithopteridae, Amphilonchidae and Incertae sedis



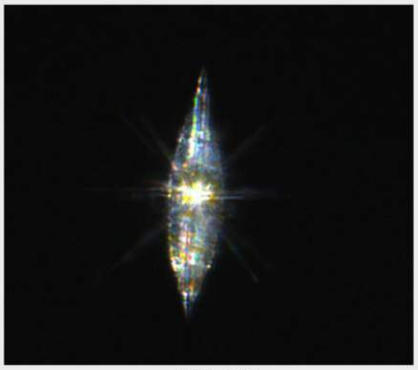
Vil 160 - sub-clade F1
Lithoptera fenestra



SES 29
Zygostaurus amphitectus.



Ei 54
Amphilonche elongata



Vil 70
Amphibelone hydrotomica

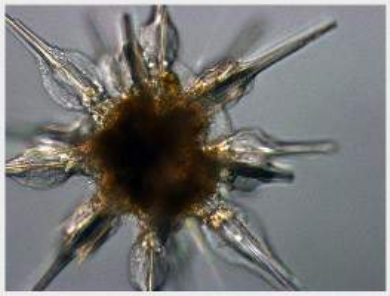
Main morphological characters:

- Small central junction
- Presence of many endosymbiotic microalgae (up to 100 in *Lithoptera*)
- Spicules of different forms and length, presence of apophyses (*Lithoptera* and *Zygostaurus*)
- 2 main spicules thicker and longer in *Amphilonche* sp. and *Amphibelone* sp.
- Many axopods and pseudopods

Clade F
Sub-clade F3a

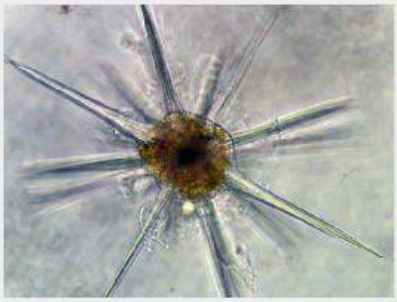
Order: Arthracanthida
Family: Acanthometridae

Ei 43 - sub-clade F3a
Xiphacantha murrayana



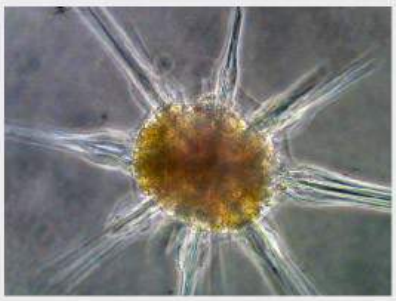
Vil 96 - sub-clade F3a
Acanthonia alata

Oki 98 - sub-clade F3a
Xiphacantha quadridentata



Oki 53 - sub-clade F3a
Acanthonia longispina

Ei 16 - sub-clade F3a
Acanthonia tetracopa



Oki 85 - sub-clade F3a
Acanthometra multispina

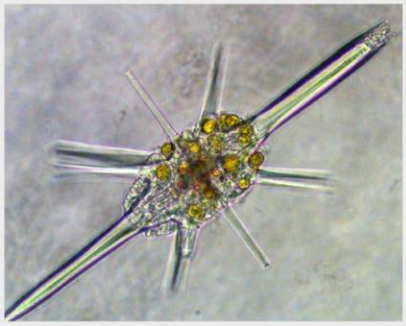
Main morphological characters:

- The 20 spicules can be of equal length and shape
- Presence of apophyses (lateral extension) on each spicule. Wing-shaped or cross-shaped
- up to 50 myonemes on each spicule and many axopods
- Brownish endoplasm with presence of endosymbiotic microalgae

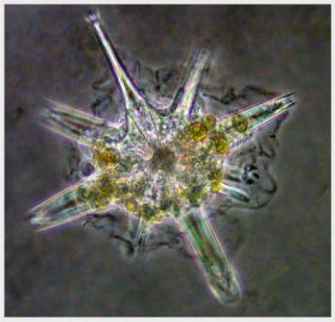
Clade F

Sub-clade F3b

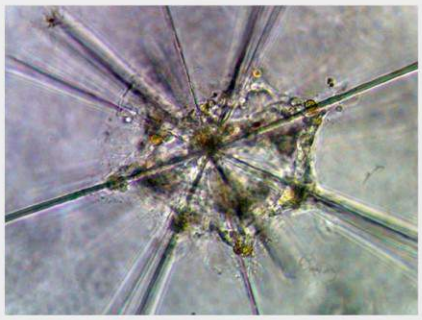
Order: Arthracanthida
Family: Acanthostauridae



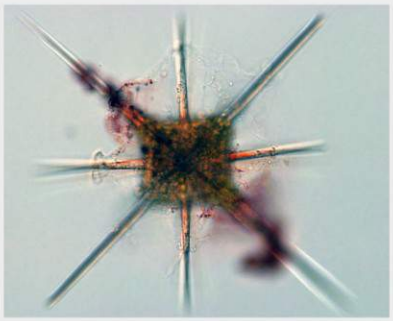
Oki 29 - sub-clade F3b
Amphistaurus complanatus



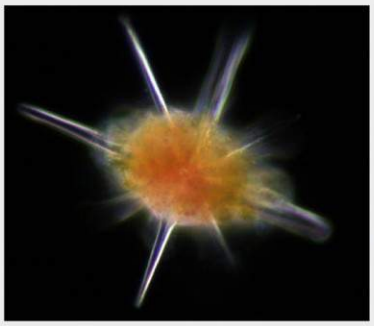
Oki24 - sub-clade F3b
Acanthostaurus conacanthus



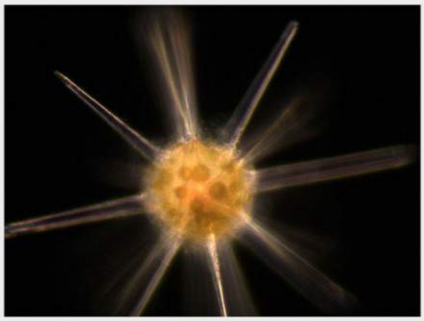
Oki 67 - sub-clade F3b
Acanthometra pellucia



PEC 17 - sub-clade F3b
Acanthostaurus purpurascens



Ei 52 - sub-clade F3b



Ei 39 - sub-clade F3b
Phyllostaurus siculus

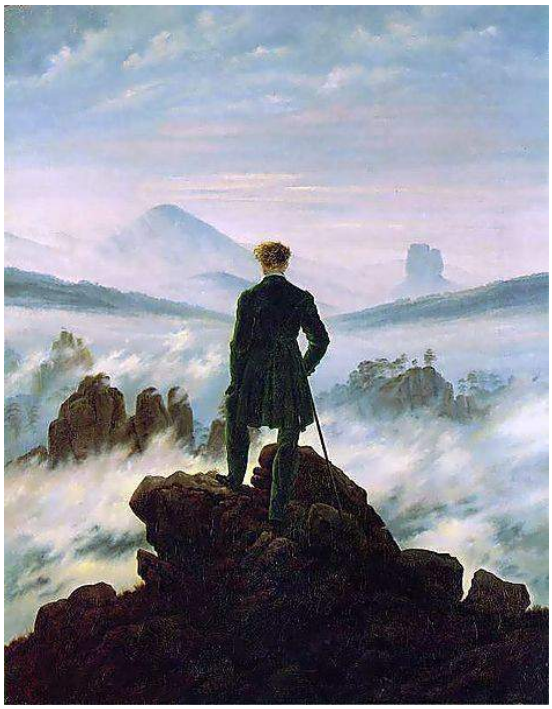
Main morphological characters:

- Central junction can be very large and robust (*Acanthostaurus*) or small (*Acanthometra*)
- Clear endoplasm with many endosymbiotic microalgae
- The 20 spicules can be of different length and shapes, bladed
- Axopods and about 20 myonemes
- Ectoplasm can be well-developped

Annexe 3

Du romantisme à la découverte des radiolaires

Johan Decelle



Caspar David Friedrich: « Le voyageur contemplant une mer de nuages » (1818).

Les radiolaires sont des organismes marins unicellulaires (protistes) capable de construire un squelette minéral et, pour certains, de vivre en symbiose avec des algues. Ces créatures planctoniques dérivent au grès des courants dans tous les océans depuis plusieurs centaines de millions d'années. Notre connaissance sur la biologie des radiolaires reste très maigre. Pourtant, au milieu du 19ème siècle, ils font partie des organismes marins unicellulaires les mieux étudiés. D'abord explorés dans les fossiles, ils font l'objet d'une grande attention dans le monde planctonique. Leur morphologie, symétrique et à la fois très complexe, dégage un certain esthétisme qui fascine et ne laisse pas sans interrogation. Les radiolaires furent au départ au cœur des études sur la morphologie, science partie du grand courant romantique. Poètes, philosophes et scientifiques s'interrogent sur la multitude des formes du vivant qui les entourent, et essayent d'en comprendre l'origine. Johannes Wolfang von Goethe fut l'un deux, le pionnier. Aussi habile avec les mots qu'avec les sciences, il parle pour la première fois de morphologie et promulgue les premières théories. Il aura une grande influence pendant plus d'un siècle sur l'étude des radiolaires, considérés comme modèles empiriques pour comprendre et confirmer les théories de l'époque.

L'histoire de la découverte des radiolaires est associée à de magnifiques voyages et aventures humaines où le dépassement de soi et la soif de connaissance étaient les principaux moteurs. Les premiers scientifiques à approcher ces organismes avaient pour la plupart une formation de médecin et d'anatomiste, et ils ont été très rapidement subjugués par ces formes planctoniques, vues nul par ailleurs. Ces explorateurs de la connaissance ont consacré leur vie à étudier, enquêter, chercher, parfois d'une manière boulimique et mettant leur santé à rude épreuve. Bourreaux du travail et passionnés ardents, ils ne supportaient aucune impasse de la connaissance. Ils accordaient également une grande importance à l'enseignement afin de transmettre avec un enthousiasme débordant leurs connaissances et leurs théories. Leur enfance était émaillée de textes philosophiques et de récits d'aventuriers en terres exotiques qu'ils dévoraient. Devenant à leur tour des explorateurs intrépides, ils ne manquaient pas une occasion pour sauter dans un bateau, telle une fuite spirituelle. Lorsqu'ils partaient collecter des échantillons, ils menaient une vie quasi monacale, seuls devant la mer et ses créatures aux « formes harmonieuses ». S'il y avait des montagnes à l'horizon, certains se lançaient dans une folle montée pour atteindre les hauts sommets. Abandonnant leur collecte et observations le temps d'un jour, ils voulaient atteindre de nouvelles cimes, quitte à frôler parfois la mort, tel était leur défi.

Parce qu'étudier les radiolaires sans connaître toute l'épopée vers leur connaissance, serait manquer une partie de l'histoire. C'est aussi une occasion pour connaître et saluer de grands noms scientifiques, tels que Goethe immense philosophe, poète et scientifique, Ehrenberg père de la micropaléontologie, Johannes Müller, grand physiologiste, Thomas Henry Huxley, défenseur ardent et proche ami de Darwin, et enfin Ernst Haeckel père de l'écologie et de la phylogénie. Cette postface de thèse s'arrêtera sur la vie de ces hommes, leurs influences et leurs œuvres artistiques et scientifiques. Ces lignes n'auront pas pour objectif de juger certaines de leurs théories mais, tel un récit de voyage, elles nous emmèneront à la découverte de mondes académiques, de villes, de pays et bien sûr d'océans où dérive une myriade de créatures microscopiques.

Au commencement était le romantisme

On ne pouvait parler de l'histoire de la découverte des radiolaires sans mentionner l'esprit universel qu'est **Johann Wolfgang Goethe**. A la fin du 18ème siècle, le poète Schelling le fit entrer dans le cercle des romantiques à Jena en Allemagne. On connaît Goethe pour ses poésies et ses œuvres littéraires comme « *Faust* » et « *Les souffrances du jeune Werther* », véritables prémices du mouvement romantique. Mais ses travaux pionniers en science et plus particulièrement sur la morphologie sont beaucoup moins connus. C'est d'ailleurs lui qui formula pour la première fois le mot « morphologie » en 1796 dans une lettre destinée au poète Schiller. Ce génie multiforme avait un goût prononcé pour l'horticulture, et passait beaucoup de temps à flâner dans son jardin ou dans les jardins botaniques de Weimar et Jena pour observer la forme et la croissance des plantes. Son voyage en Italie, qui dura 3 années, fut une renaissance spirituelle et influença grandement ses pensées littéraires et scientifiques. Dès sa prime jeunesse,

il fut attiré par ce pays par l'intermédiaire de son père qui avait pérégriné auparavant sur la péninsule. Le 3 septembre 1786, muni d'une simple valise, il « *s'enfuit* » en toute discrétion de Carlsbad au beau milieu de la nuit afin que personne ne le retienne. A son plus grand bonheur, il explora les principales villes de l'Italie comme Venise, Bologne, Rome, Naples. Il s'intéressait à l'architecture des palais et des églises et appréciait plus particulièrement aller au théâtre le soir. Il côtoyait chaque jour de nombreux littéraires et peintres allemands qui furent une grande source d'inspiration (comme Johann Heinrich Wilhelm Tischbein qui dessina son portrait dans la campagne romaine). Ensemble, ils exploraient et apprenaient la peinture italienne de la Renaissance et l'art de l'Antiquité romaine. Goethe s'échappait parfois de ce réseau intellectuel le temps d'une journée pour arpenter en solitaire les flancs du Vésuve et de l'Etna, et dominer ainsi la belle nature enchanteresse italienne. En Sicile, à dos de cheval avec un ami peintre, son observation scientifique s'intensifia et devint son occupation principale. Il récoltait, crayonnait, étudiait, comparait aussi bien les roches que les plantes. Il s'interrogeait sur le développement de la vie organique en général, et cherchait une loi universelle. C'est à ce moment là que se développa et émergea sa théorie de la plante originelle « Urpflanze », la forme primitive qui donna naissance à toutes les plantes.

Portrait de Goethe par son ami allemand Tischbein lors de son voyage en Italie (*Goethe dans la campagne romaine, 1787)*

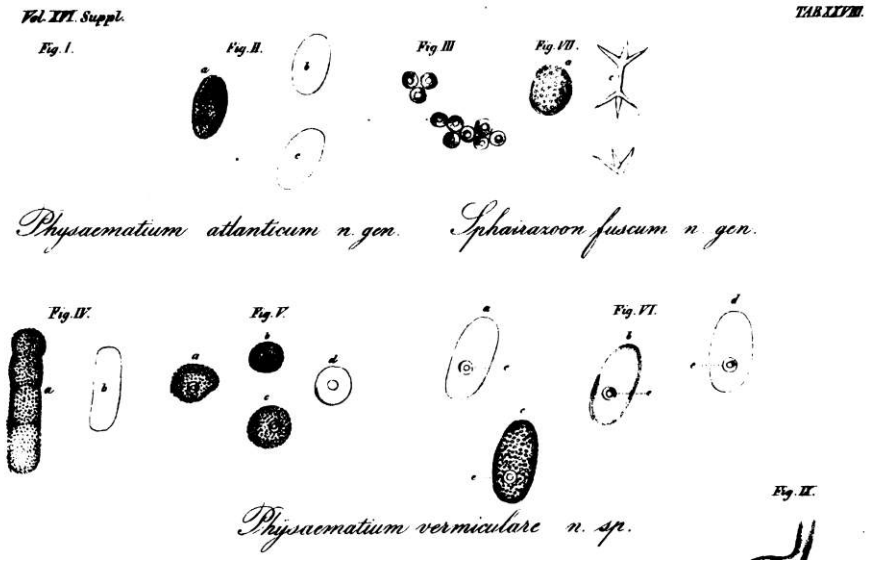
A son retour en Allemagne, carnet rempli d'observations et d'écrits diverses, il examina de plus près les plantes de son jardin et celles de l'Université de Jena. Il écrivit son premier ouvrage sur la morphologie en 1790: « *Metamorphosis of plants* ». Il développa ainsi la notion de l'archétype où les différentes parties d'une plante sont la résultante d'une feuille (l'archétype) après transformations successives. « *Tout est feuille, et à travers cette simplicité, la plus grande multiplicité est possible* ». Lors de son deuxième voyage en Italie, il tomba sur un crâne de

mouton qu'il le frappa. Il suggéra qu'il existe aussi chez les animaux un archétype. A l'instar d'une feuille, une vertèbre donne naissance après modifications aux différentes parties d'un squelette animal. L'archétype est retrouvé chez toutes les plantes et les animaux et constitue un modèle par lequel on peut comprendre les diverses structures et le développement des formes. Implicite et abstrait, sa découverte ne passe pas par l'œil seul mais par un travail d'esprit. Goethe est un des premiers scientifiques à comparer les structures entre de nombreux animaux et à pratiquer l'expérimentation en répétant ses observations sous différentes conditions. Il reconnu déjà à l'époque que les formes ne sont pas fixes et transmutent dans le temps selon un mouvement constant. Ces formes du monde organique résulteraient de l'interaction entre l'archétype et les forces extérieurs de l'environnement. Il impose ainsi la téléologie divine à une causalité naturelle.

Fin 1794, Goethe, qui regrettait l'absence d'un compagnon avec qui il pouvait échanger ses idées scientifiques, rencontra à Jena les très influents frères Wilhelm et Alexander von Humboldt. Il leur fit part de ses travaux sur la morphologie plus particulièrement au naturaliste Alexander, qui encouragea grandement l'homme de lettres à poursuivre vers cette nouvelle science. Adaptant les idées de Kant et son « *Troisième critique* », Goethe et Humboldt pensaient que le jugement esthétique était nécessaire à la compréhension scientifique, car il permettait de capturer les lois de la nature et dévoiler ainsi l'unité parmis une immense variété. L'étude de la morphologie était pour Goethe une manière aussi de comprendre l'art et l'esthétisme.
A la fin du 18ème siècle, Goethe lança ainsi une nouvelle science, l'étude de la morphologie, faisant le lien entre anatomie et histoire naturelle. Il écrivit plusieurs essais, aussi bien en botanique qu'en zoologie, qui auront une grande influence au cours du siècle suivant sur la taxonomie et la classification du monde vivant. Goethe n'a vraissemblablement pas observé des microorganismes marins comme les radiolaires. En revanche, ses théories et sa vision sur les formes de la vie marqueront la génération suivante de scientifiques comme Johannes Müller et Ernst Haeckel, qu'ils le considéraient comme son idole, et partageront ce même désir de découvrir les lois universelles de la morphologie.

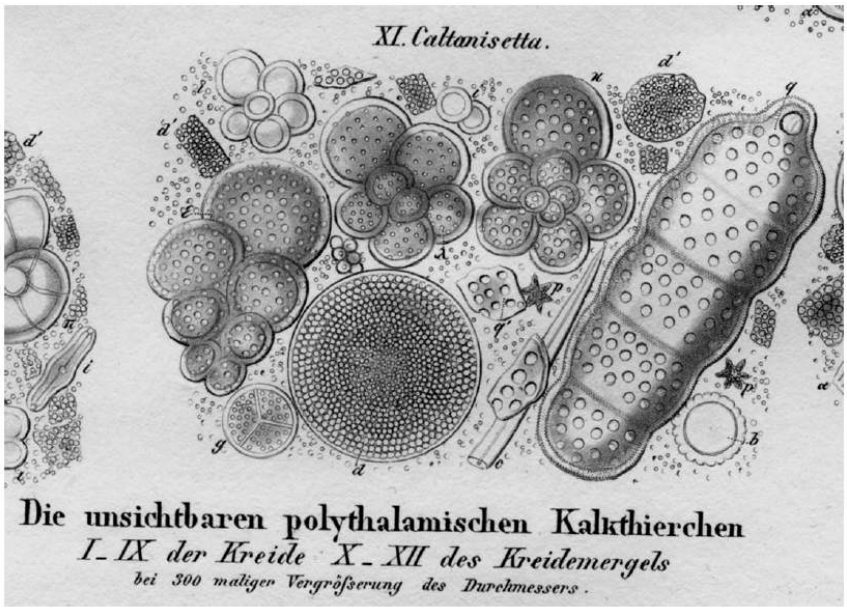
Découverte des premiers radiolaires des pierres

En 1834, un scientifique allemand **Meyen** décrit pour la première fois un radiolaire vivant *Physematium atlanticum* au large des îles Canaries. C'est surement grâce à sa grande taille que ce radiolaire a pu être récolté et observé à travers un microscope de l'époque. En effet, il fait partie d'un groupe de radiolaires, appelé *Sphaerozoum* par Meyen, qui forment des colonies de milliers d'individus et ainsi peuvent atteindre plusieurs centimètres. Les observations de Meyen restent néanmoins anecdotiques et il faudra attendre les années 1850 pour avoir des descriptions plus nombreuses et plus détaillées de radiolaires vivants.



C'est d'abord à travers les fossiles que les premières connaissances sur la classification et la distribution géographique des radiolaires ont été faîtes. Ami et compagnon de voyage d'Alexander von Humboldt, **Christian Gottfried Ehrenberg** fut professeur de médecine à l'Université de Berlin et est considéré aujourd'hui comme le père de la micropaléontologie. Ehrenberg fut un grand voyageur, avec plus d'une vingtaine de pays parcourus. Pendant 5 ans (1820-1825), il foula les terres du Moyen-Orient pour recueillir des milliers d'échantillons de plantes et d'animaux. L'Égypte, le désert de Libye, la vallée du Nil et la côte nord de la mer Rouge étaient ses terrains de jeu où il collecta plus particulièrement des coraux, puis ensuite il poursuivit son épopée en Syrie, Arabie et Abyssinie. Des 17 personnes que composaient cette expédition, seuls Ehrenberg et un peintre italien revinrent vivants. Ses mésaventures en terres étrangères ne l'empêchèrent pas de repartir en 1829, à travers la Russie, l'Ukraine, la Sibérie jusqu'à la frontière chinoise pour accompagner son ami Humboldt. Pionnier dans l'utilisation du microscope, il s'intéressa après ce dernier voyage à tous les microorganismes qu'il observait aussi bien dans le sol, dans les fonds marins, dans l'eau mais aussi dans la poussière de l'air (l'aérobiologie). Il devint l'expert en Europe de tout ce monde invisible d'animalcules et recevait régulièrement des échantillons de nombreux scientifiques, comme Charles Darwin qui revenait de son tour du monde sur le HMS Beagle. Des milliers d'espèces sont décrites pour la première fois, appartenant à des groupes très différents comme les Euglènes, les ciliés (les infusoires à l'époque comme *Paramecium*), les diatomées, les haptophytes mais aussi les foraminifères et les radiolaires. En 1835, Ehrenberg était un des premiers à venir sur l'île d'Helgoland en Mer du Nord à environ 20 miles des côtes allemandes. Il voulait étudier la phosphorescence de la mer, phénomène inconnu à l'époque qu'il observa lors des ses précédentes expéditions. L'île d'Helgoland, que l'Angleterre céda à l'Allemagne en 1890, était un lieu incontournable pour tous les scientifiques allemands en botanique et zoologie au 19ème siècle. Ses côtes rocheuses, quasi absentes sur le continent allemand, abritent une grande diversité d'algues, au plus grand

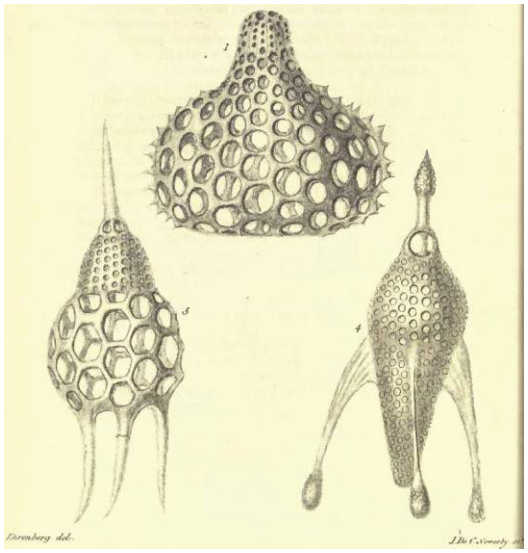
plaisir des naturalistes. Ehrenberg continuait de recevoir une pléthore d'échantillons d'autres scientifiques et explorateurs revenant des quatre coins du globe, et notamment des sédiments contenant des microfossiles. Il examinait, dessinait et décrivait dans le moindre détail ces microorganismes inconnus, dont les radiolaires.



Les premiers radiolaires fossiles illustrés, par Ehrenberg (1839)

Ces premières descriptions seront le début d'une grande série, qui constituera son impressionnante collection à Berlin. Ehrenberg présenta pour la première fois les radiolaires devant l'Académie des Sciences de Prusse le 20 décembre 1838, et rédigea la première publication l'année suivante. En 30 ans environ, il écrivit 40 publications sur les radiolaires où il défini sur la base du squelette plus de 500 espèces dans 77 genres. Ces radiolaires viennent de l'Atlantique Nord, de Méditerranée, de l'Océan Indien, de l'Antarctique, du Golf du Mexique et du Pacifique. En plus de leur répartition géographique, il étudia la distribution verticale des fossiles de radiolaires dans les années 1850, en comparant les communautés vivant dans les fonds marins et à différentes profondeurs. Il classa ces animalcules dans l'ordre des *Polygastrica* et leur donna le nom de *Polycystina* comprenant les divisions «*Spumellaria* et *Nassellaria* . Il inclue par erreur dans cette grande famille d'autres organismes siliceux comme les diatomées (*Coscinodiscus*), silicoflagéllés et des spicules d'éponge. Les acanthaires, un autre groupe de radiolaires, ne font pas partie des échantillons collectés et décrits par Ehrenberg, très probablement parce qu'ils ne sont pas présents dans les microfossiles. Selon lui, les radiolaires sont multicellulaires, dont les organismes les plus proches sont les échinodermes. Il n'a surement jamais observé des radiolaires vivants à travers son microscope. C'est peut être la raison pour laquelle il ne sera pas d'accord avec les observations faites par d'autres scientifiques sur les radiolaires fraîchement collectés. A la fin de sa vie, il devenait presque aveugle et c'est sa fille

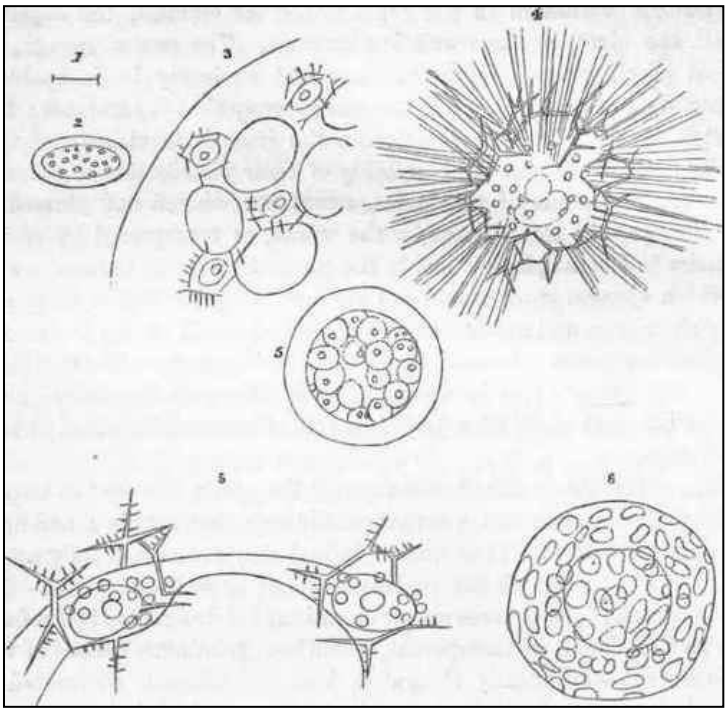
Clara qui lui décrivait ce qu'elle voyait à travers le microscope. Ehrenberg fut salué par ses pères et reçu au cours de sa carrière de nombreux prix prestigieux comme la médaille d'or de Leeuwenhoek et le Prix Cuvier. Il rentra en 1860 à l'Académie française en prenant la place de son ami Alexander von Humboldt. Il meurt à Berlin le 27 juin 1876 à l'âge de 81 ans.



Dessins de radiolaires polycystines par Ehrenberg

En Angleterre, 10 après les premières observations de radiolaires fossiles d'Ehrenberg, **Thomas Henry Huxley** embarqua à bord du HMS Rattlesnake en 1847 en tant que jeune chirurgien pour explorer la Nouvelle Guinée et l'Australie. Physiologiste de formation, il ne souhaitait pas collecter et décrire de nouvelles espèces comme son homologue Darwin, mais il voulait comprendre les différents mécanismes des animaux marins en les comparant un à un. Il s'intéressa tout particulièrement aux mollusques et méduses. Passionné par le Romantisme, Huxley emporta de nombreux livres à bord dont ceux de Johann Wolfgang Goethe. Le londonien Edward Forbes, naturaliste marin expérimenté et auteur de la théorie azoïque, l'encouragea à prendre également un filet pour collecter les organismes à la surface de la mer. Durant l'expédition, Forbes recevra également toutes les observations et descriptions de Huxley, qui feront l'objet de nombreuses publications. Contrairement au Beagle, le Rattlesnake était très peu équipé et offrait des conditions misérables de vie et de travail pour le jeune Huxley. Il traînait lui-même ses filets le soir à l'arrière du bateau et amenait sa pêche dans la minuscule cabine des cartes marines, gentiment prêtée par le capitaine Stanley. Une bonne partie de la nuit, alors que le calme régnait sur le pont, il observait à travers son microscope et égratignait son cahier de dessins d'organismes. Parmi ceux-là, dans les eaux tropicales entre l'Australie et la Nouvelle Guinée, il tombait constamment sur des créatures étranges, agglomérés dans une gelée et flottant à la surface de l'eau. Il décrivit minutieusement l'organisation de ce nouveau zoophyte, observé par Meyen en 1834, et donne le nom de *Thalassicolla*, confiture des mers. Huxley discuta de leur position dans le règne vivant et reconnu pour la première fois que ces masses gélatineuses,

Thalassicolla et *Sphaerozoum*, ne sont en fait que des colonies d'unicellulaires (protistes). Selon Huxley, ils seraient le lien entre les éponges et les foraminifères mais leur bizarrerie et le manque de lien avec tout autres organismes le découragèrent à entreprendre de plus longues investigations. Dans ses colonies, Huxley observa aussi des « *petites cellules jaunes claires et sphériques* ». C'est peut être la première fois qu'est observée la symbiose animale avec des microalgues, terme qui d'ailleurs n'existe pas encore à l'époque. Le phénomène de symbiose dans le monde vivant n'est pas du tout connu et il faudra attendre une trentaine d'années pour voir les premières expériences sur ces cellules jaunes, trouvés aussi chez d'autres organismes, qui montrent que ce sont en fait des microalgues, entités indépendantes de leur hôte.



Dessins de radiolaires coloniaux (*Thalassicolla*) d'après TH Huxley

Les radiolaires coloniaux (Collodaria) furent sans doute les toutes premières observations de radiolaires vivants collectés dans les océans, et ont constitué un modèle pour étudier la symbiose animale-végétale. Le mot « radiolaire » n'existe nul par ailleurs encore dans les publications de l'époque, jusqu'aux travaux d'un physiologiste allemand, Johannes Müller.

Johannes Müller

Johannes Müller, physiologiste allemand de renom, étudia les radiolaires pendant une grande partie de sa vie. C'est lui qui décrivit pour la première fois les acanthaires, ou *Acanthometren*, ainsi il les nomma. Müller a été lui aussi grandement influencé par l'œuvre de Goethe. Il admirait chez le poète aussi bien ses travaux d'anatomie comparative que ses œuvres littéraires auxquelles il se referait souvent au cours de sa carrière scientifique. En octobre 1828, les deux hommes se rencontrèrent et échangèrent leurs théories et dernières trouvailles.

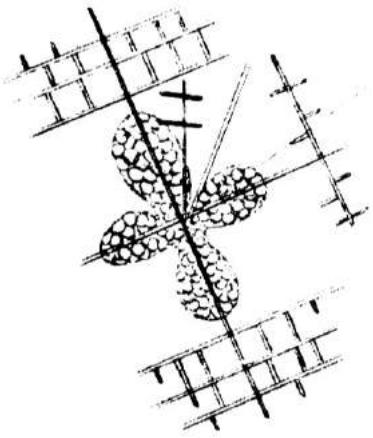
Au départ, Müller, fils de cordonnier, choisit la faculté de médecine à Bonn après avoir hésité avec la théologie. Il se passionna très vite pour l'observation et le dessin à travers le microscope de structures animales, ce qui fut très apprécié par ses professeurs. A l'époque, la Naturphilosophen était très présente dans le monde académique. Inspirée par la philosophie de Kant et de Schelling, la Naturephilosophen admettait que l'ordre de la nature correspondait à la structure de la conscience humaine, et essayait de discerner des systèmes dans les structures naturelles, parfois classifiant les animaux en les comparant avec les systèmes sensitifs de l'Homme. Après un diplôme sur le mouvement animal en 1822, Müller souhaitait vivement rejoindre Paris pour étudier aux côtés de Georges Cuvier, l'un des fondateurs de l'anatomie comparée. Mais le directeur de l'université de Bonn l'envoya plutôt à Berlin, où il rencontra Asmund Rudolphi qui lui donna un microscope (Frauenhofer microscope) et l'encouragea à poursuivre ses études en microscopie. Müller retourna ensuite à Bonn où il enseigna l'anatomie et la physiologie, et entreprit en parallèle de nombreuses observations avec son nouvel instrument. Il travaillait à un rythme effréné, et commençait à souffrir de premières dépressions. Il dormait très mal et ses nuits se résumaient souvent par des insomnies.

Au printemps 1833, commença sa nouvelle vie à Berlin où il devint professeur d'anatomie et de physiologie et dirigea à la fois l'Institut d'Anatomie et le Muséum d'Anatomie. Il habitait avec son élève Theodor Schwann non loin de l'université dans l'ancien appartement du philosophe Hegel. Complices, les deux hommes passaient beaucoup de temps derrière leur précieux microscope pour disséquer et dessiner des animaux. Sous l'impulsion de son professeur, Schwann fera des découvertes majeures comme la théorie de la cellule ("*All living things are composed of cells and cell products*") et les cellules nerveuses qui portent son nom. Avec toutes ses responsabilités, Müller ne lui restait que peu de temps pour sa recherche. De plus, l'Université à l'époque ne fournissait pas de laboratoires et de bureaux pour les expériences, ce qui rendait l'activité de recherche inconfortable. Müller et ses quelques 200 élèves disséquaient des corps de suicidés et de criminelles dans des petites salles sombres et puantes. Quant au Muséum d'Anatomie de l'Université, il donna énormément de son temps et parfois même son propre argent pour développer ce lieu unique. C'était pour lui comme une arche de Noé où il voulait rassembler, parfois de manière obsessionnelle, le plus d'animaux possible pour offrir une représentation de la vie animale. En comparant toutes ces formes et structures multiples, il souhaitait découvrir le « grand plan » d'organisation de la vie. Müller ne gagnait pas beaucoup d'argent malgré ses heures de travail et ses responsabilités académiques. Il craignait la pauvreté.

Son père mourut en 1832, ses 2 frères tombèrent dans l'alcoolisme, et Müller essaya de supporter sa famille tant bien que mal.

Depuis sa venue à Berlin, il se concentra sur l'anatomie comparative. Il fit énormément de travaux sur les systèmes visuels, endocriniens, circulatoires et reproductifs sur un grand nombre d'animaux pour comprendre le fonctionnement du vivant. Cette approche comparative fut pionnière à l'époque, et lui permit de faire de nombreuses découvertes. Il fonda aussi un nouveau journal en 1834 *The archive for anatomy, physiology and scientific Medecine*., qui devint fameux en Europe. Müller encouragea la microscopie comme méthode clef pour les études d'anatomie et de physiologie mais aussi en cas clinique pour diagnostiquer des pathologies. Selon lui, un œil sans microscope est un œil « *non armé* ».

Mais le projet qui l'occupa le plus dans la seconde moitié de sa vie, fut la classification des organismes marins. Cet intérêt pour le monde marin a sûrement été déclenché par les découvertes de son collègue Christian Gottfried Ehrenberg. Ses responsabilités à Berlin, dont celle de doyen de la faculté de Médecine, ne lui permettaient pas de voyager comme il le souhaitait, et l'été était la seule période où il pouvait se consacrer à la recherche. Appréciant la compagnie et le partage des observations, il proposait à ses élèves les plus doués de venir avec lui. Il parcouru la Mer du Nord et la Méditerranée pour observer et étudier les créatures marines. Il partait à la rencontre des pêcheurs locaux et vagabondaient sur les marchés pour récupérer des spécimens frais. Il fut particulièrement frappé par la symétrie étonnante de certains organismes comme les étoiles de mer et les oursins. En 1838, Müller se rendit à Marseille, puis à Gêne, pour collecter des élasmobranches (requins et raies). En 1840, il voyagea à Trieste pour étudier les échinodermes, les cyclostomes (lamproies) et les poissons à écailles. Müller se concentrait sur des organismes qui selon lui étaient aux frontières du système de classification. Il développa un nouveau système de classification pour de nombreux groupes de poissons.



J. Müller. ger.

Durant son premier séjour sur l'île d'Helgoland en 1845, Müller construisit un filet, le « *Müller's net* », similaire à celui pour capturer les papillons, mais avec en plus un collecteur à l'extrémité du filet. Méthode plus douce que le dragage utilisé à l'époque, le filet à plancton aurait été pensé et utilisé pour la première fois en 1828 par un chirurgien de l'armée anglaise, John Vaughan Thompson. Le naturaliste François Perron utilisait également de grand filets pour écumer la surface de l'eau dans la baie de Villefranche-sur-Mer ou lors de son expédition en Australie, dirigée par le capitaine Nicolas Baudin. Johannes Müller, d'une embarcation de pêcheurs, tenait son *« petit filet à papillon* » à demi immergé dans l'eau et le ramenait ensuite délicatement sur le pont. Le filet était brièvement tourné et immergé dans un seau d'eau de mer

afin de libérer les animaux accrochés. Microscopiste averti, Müller pouvait collecter ces créatures et les transférer rapidement sur des lames de microscopes avec une pipette. Cette technique de collecte, qu'il définissait comme « *pêche pélagique* », était pour lui un moyen pour résoudre l'énigme du développement animal. Il était fasciné par les formes des larves de poissons, d'échinodermes et de méduses, mais aussi les radiolaires, qui resteront ses observations préférées. En 1854, il écrit un article « *On the Different Forms of Marine Animals* ». Selon lui, l'ordre de la nature, *le système*, peut être découvert non pas par une simple description de la forme extérieure des organismes mais par l'étude de leur développement et organisation interne. Il se cramponnait nuit et jour, avec ardeur et passion à son filet et à son microscope. Müller n'avait jamais peur d'avoir plus de travail et ses étudiants pouvaient à peine suivre le rythme effréné de leur professeur.

A la fin des années 1840, cherchant un nom pour rassembler toutes ces nouvelles créatures marines microscopiques, Müller consulta le philologue allemand Jakob Grimm (1785-1863), qui venait de compiler un dictionnaire de la langue allemande, en plus de sa fameuse collection de contes avec son frère. Ensemble, le scientifique et le linguiste inventèrent le terme *pelagischer auftrieb* (*Pelagischer* : oceanic/pelagic upwelling ; trieb :deriver ; auf : vers le haut). Il faudra attendre près d'un demi-siècle plus tard, pour voir apparaître le terme plancton, instauré par ce même Grimm et Victor Hensen, élève de Müller.

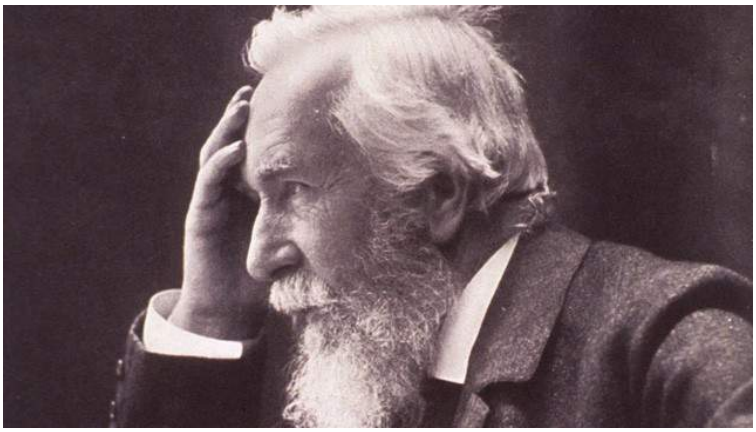
Müller poursuivit son travail sur les radiolaires et partit collecter en France à Nice puis à St Tropez. Lors de sa troisième visite sur l'île Helgoland en 1854, accompagné de son fils, Müller rencontra un peu par hasard d'anciens étudiants, La Valette St.George et un certain Ernst Haeckel qui étudiait à l'époque avec Koelliker et Virchow à Wurzburg. Toutes ses collectes et observations sous le microscope l'amenèrent à rédiger une courte monographie qui se publiera un an après sa mort en 1858: « *Uber die Thalassicollen, Polycystinen und Acanthometren des Mittelmeeres* ». Cet ouvrage représente la première étude sur les radiolaires vivants et est une première tentative de classification. Müller appela tous ces nouveaux organismes *Rhizopodia radiaria* ou *Radiolaria*, et décrivit entre autres les acanthaires (*Acanthometren*). Il remarqua que ces derniers n'avaient pas un squelette en silice comme les autres Radiolaires. Les Radiolaires sont pour Müller un modèle formidable pour étudier la morphologie mais leurs rôles écologiques dans les océans restent très abstraits. Il appelle d'ailleurs ces créatures microscopiques « *la poussière philosophique* ».

A la fin de sa vie, il souffra de plusieurs épisodes dépressifs, suite au décès de sa mère, et au choix de son fils Max qui décida de devenir physicien plutôt que biologiste. Lassé, il pensait que sa « *pêche pélagique* » ne lui ramènerait plus aucune découverte. « *Parce que j'exploite cette pêche tellement intensément depuis plusieurs années, je vais devoir bientôt abandonner tout ça puisque j'ai maintenant presque déjà tout rencontré parmi les animaux de la mer* ». Une tragédie survint en septembre 1855. Alors qu'il retournait d'une campagne d'échantillonnage en Suède avec des étudiants, son bateau se fit heurter par un autre et coula soudainement dans la Mer Baltique. Il survécu en s'accrochant à un morceau de bois dérivant, mais un de ses étudiants n'eut pas cette chance et succomba. Müller se sentit unique responsable de cette disparition et

tomba dans une dépression telle qu'il n'en sortira jamais. Il peinait à poursuivre l'enseignement et la recherche et sa santé empirait. Il prit de l'opium pour soulager ses douleurs abdominales et combattre l'insomnie, mais rien n'y fera. Müller s'éteindra mystérieusement à 56 ans (pour certains d'une overdose d'opium), l'année avant la publication de "*The Origin of Species*" de Charles Darwin.

En plus de ses travaux colossaux et grandes découvertes sur la physiologie, la cellule mais aussi en biologie marine, Johannes Müller aura formé les premiers anatomo-pathologistes et physiologistes de l'Allemagne de cette époque : Du Bois-Reymond, Helmholtz, Virchow, Henle, Haeckel, Schwann. En fabricant son propre filet et en utilisant un microscope destiné au départ à la médecine, Müller peut être considéré comme le premier scientifique à avoir étudié le plancton vivant. D'une manière générale, l'illustre professeur a permis de faire entrer la biologie marine comme nouvelle discipline à l'Université et a encouragé également la création de stations marines pour étudier de plus près la faune marine. Il transmit ses premières découvertes avec une grande largesse, ce qui inspira une génération de scientifiques pour explorer les merveilles de la vie marine. Ernst Haeckel fut l'un d'eux. Grâce à son professeur, Haeckel se découvrit une passion ardente pour la biologie marine, la morphologie du plancton et les radiolaires. Müller et Goethe seront ses deux idoles.

Ernst Haeckel



Inutile de présenter Ernst Haeckel, un des plus grands intellectuels du 19ème siècle. Homme d'encyclopédie aux multiples facettes, Haeckel reste le père de plusieurs disciplines en biologie comme l'écologie et l'embryologie, et invente aussi les termes « protiste » et « phylogénie ».

Aventurier et rêveur

Ernst Haeckel est né dans la petite ville allemande de Postdam en 1834, dans une famille très religieuse. Durant sa jeunesse, ses parents lui firent découvrir la poèsie de Schiller et de Goethe,

et son frère Karl lui transmit sa passion pour les voyages. Attiré par les pays lointains et l'exotisme, il dévora rapidement les livres d'aventures d'Alexander von Humboldt et de Charles Darwin, personnages qui le suivront durant toute sa vie. Il se passionnait aussi très vite pour les plantes et constitua un herbier des espèces locales collectées lors de ses promenades à travers la nature. La peinture était également un de ses passe-temps préférés, qu'il pratiquait avec un certain talent, remarqué lors de ses cours au lycée. La personnalité d'Haeckel est déjà très bien définie : voyageur, rêveur, poétique et observateur.

Après avoir hésité avec la botanique, il rentra à l'université de médecine à Meresburg. Il aimait ce qu'il étudiait mais disséquer les macchabés n'était pas ce qu'il affectionnait le plus. Il se passionna plutôt pour certains livres comme celui du botaniste allemand Schleiden. L'auteur y explique le rôle primordial du « développement » dans les sciences naturelles, qui permet de passer des formes simples aux formes complexes. Selon Schleiden, la compréhension scientifique doit être accompagnée par un raisonnement esthétique, et la connaissance scientifique doit inclure la peinture et le voyage pour représenter et voir respectivement la nature. Cette union des sens artistiques et scientifiques est clairement un héritage de Goethe. Haeckel connaissait parfaitement chaque page de ce livre, et les idées développées façonneront sa manière de pensée. Haeckel fut aussi marqué par les cours de Samuel Weiss, qui cherchait à comprendre les formes des cristaux avec des formules mathématiques. Ce n'étaient pas les mathématiques qui l'attiraient mais plutôt cette tentative de décrypter des formes complexes.

Rencontre déterminante avec Johannes Müller

Au printemps 1854, il se rendit à Berlin pour prendre des cours avec le renommé professeur Johannes Müller, qui enseignait la zoologie mêlant physiologie et morphologie. Müller, qui ne concevait pas la recherche scientifique sans la philosophie, restera un model durant toute la carrière scientifique d'Haeckel. En Août, Haeckel partit avec son ami La Valette St George sur l'Ile d'Helgoland et rencontra, un peu par hasard, ce même Johannes Müller. Son professeur l'invita à collecter des organismes marins en partant chaque jour en mer, et l'initia à la pêche pélagique et à l'identification de la faune comme les échinodermes et les méduses. Müller communiqua sa fascination pour cette pléthore de formes marines: *« As soon as you have entered into this pelagic wonderland, you will see that you cannot leave it* ». L'étudiant sortira transformé de son séjour sur l'île, véritable berceau de la recherche sur le plancton à cette époque, et écrivit à ses parents pour leur dire toute son excitation et sa joie devant cette faune marine. Cette période en compagnie de Müller restera comme un de ses plus beaux souvenirs et ne cessera de le répéter lors de ses cours et rencontres. Sa décision d'abandonner la médecine pour s'orienter plutôt vers la zoologie est de plus en plus claire.

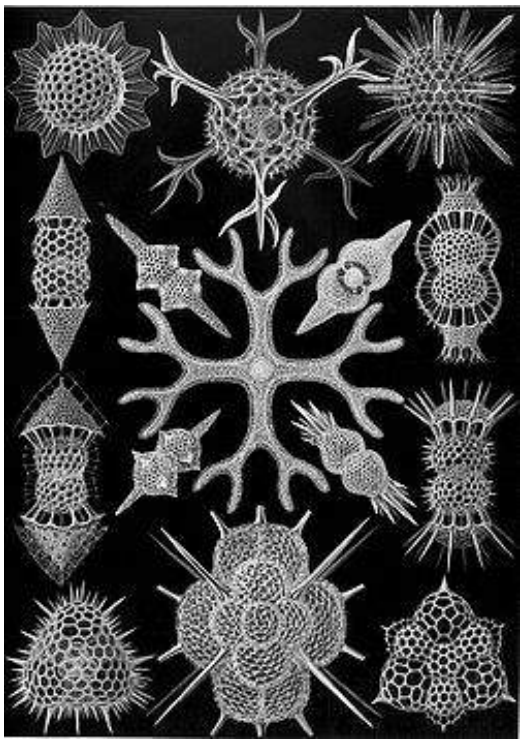
A l'été 1856, Haeckel se rendit à Nice et dans la « très jolie baie » de Villefranche-sur-Mer avec Kolliker pour collecter et étudier la faune pélagique. C'est là qu'Haeckel vit de nombreux radiolaires, les acanthometra et les polycystinea, qu'il trouvait « énigmatiques » et

« fantomatiques ». Au même moment, Müller était aussi à Nice pour poursuivre son étude sur les radiolaires. En Mars 1857, il obtint son doctorat de médecine, décerné par le doyen de l'Université de Berlin Christian Ehrenberg. Haeckel commença à pratiquer la médecine dans la maison de son père mais très peu de patients vinrent se faire soigner. Il était persuadé que sa vocation était plutôt la recherche en biologie marine. Haeckel avait en tête un souhait omniprésent de travailler avec le professeur Müller. La mort soudaine de celui-ci fut un grand bouleversement et une profonde tristesse, tant il aimait et respectait son professeur. Dégouté par la chirurgie, il souhaita rentrer dans le monde académique pour enseigner plutôt la zoologie. Carl Gegenbaur, un autre protégé de Müller et professeur d'anatomie, invite Haeckel dans la paisible ville de Jena (Iéna). Située dans une vallée entre des montagnes boisées, Iéna était à l'époque réputée pour son université qui accueillait de grands intellectuels, tels que Schelling et Hegel, et était le berceau du romantisme. Goethe y venait régulièrement pour visiter le jardin botanique. Haeckel fit la rencontre de plusieurs scientifiques dont Max Schultze, celui qui traduisit en allemand les ouvrages de Charles Darwin et de Thomas Henry Huxley. Haeckel se décida à passer son habilitation pour rentrer à l'université en tant que professeur de biologie (habilitationsschrift).

Voyage en Italie et la découverte d'une fascination pour les radiolaires

A l'instar de son idole Goethe, il voulait aller faire un voyage en Italie pour libérer son esprit et trouver de nouvelles sources d'inspiration. En 1859, il obtint un financement de l'université et commença son périple le long de la péninsule. Il se rendit à Gène puis à Florence, où il acquit un microscope d'Amici. Il découvrit Rome et ses peintures durant plus d'un mois, puis Naples avec l'intention de commencer son travail scientifique. Mais il fit une rencontre, qui sera déterminante dans sa vie et sa carrière scientifique, avec un poète et peintre allemand romantique, Hermann Allmers. Tous deux menèrent une vie de bohème sur les côtes méditerranéennes, arpentant le vésuve, se baignant dans les bains thermaux d'Ischia, flânant sur l'ile de Capri. Une amitié très forte se lia entre les deux hommes et Haeckel se sentait au paradis, tel un Robinson Crusoe, un Goethe, un Humboldt, ses héros de jeunesse. Allmers initia Haeckel à la peinture, qui développa une sensibilité romantique pour l'esthétisme dans la nature. Les deux amis, inséparables, se posaient en pleine nature et sortaient tout leur matériel pour esquisser quelques aquarelles au dessus du Golf de Naples. L'esprit d'Haeckel était à ce moment très loin de la science et de la médecine, et l'idée d'abandonner la science pour la peinture lui vint. Mais son talent n'était pas à la hauteur pour poursuivre, et son père, politique important de Meresburg, ne souhaitait pas que son fils s'engage dans une vie artistique. Cependant, la rencontre avec Allmers va permettre à Haeckel de développer grandement son sens de l'observation, qui sera très utile pour sa future

carrière. Pour comprendre les formes, rien de tel que de pouvoir les dessiner. Le temps passait et le jeune médecin avait dans l'obligation de trouver rapidement un sujet d'habilitation s'il voulait commencer une carrière académique. Il décida de laisser son ami Allmers et se rendit à Messina en Sicile, où Müller était venu passer de paisibles jours, pour explorer les créatures marines de la Méditerranée. Il souhaitait poursuivre les travaux de son professeur sur le développement des échinodermes mais il ne récolta que très peu de larves. Au contraire, il collecta beaucoup de radiolaires grâce aux filets de Müller. Sa récolte se faisait soit de la côte ou soit plus au large avec l'aide de pêcheurs. Haeckel fut frappé par la beauté, le caractère unicellulaire et l'extraordinaire diversité des formes chez les radiolaires. A la fin de son voyage en Italie, il se décida ainsi de faire son habilitation sur ces organismes, et suivre ainsi les traces de son vénéré professeur. Il décrivit environ 130 espèces de radiolaires dans un premier ouvrage « *De Rhizopodum finibus* » en 1862 (en latin). Il devint ainsi privatdozen (équivalent de maître de conférences privé) à l'Université de Jena. Haeckel investissait énormément de temps pour préparer ses cours, qui, au début, n'étaient suivis que par très peu d'élèves et donc n'était pas rentable économiquement.



Il fut ensuite nommé professeur avec la publication de « *Die Radiolarien* » en 1887 où il décrit environ 750 genres et 4300 espèces d'acanthaire, de collodaire, de nassellaire et de spumellaire, provenant de sa collecte méditerranéenne mais aussi de l'expédition du HMS Challenger. Basé sur le développement du squelette, Haeckel souhaite construire le système naturel des radiolaires, énigme soulevé par Linné, en décrivant les concepts de famille, genre, espèces. Il déclara « *Les radiolaires sont un exemple parfait illustrant comment la vraie évolution des formes vivantes peut être conçue comme descendance de formes géométriques abstraites* ». Ernst Haeckel envoya une copie de son « *Die Radiolarien* » à Huxley, qui le remercia et lui envoya en retour des radiolaires des Barbades et des fonds marins de l'Atlantique. Bien que différent d'un point de vue philosophique, les deux hommes échangeront tout le long de leur carrière.

Malgré de nombreuses responsabilités administratives et dans l'enseignement, Ernst Haeckel prenait le temps de se promener avec son ami Gegenbaur au jardin botanique ou dans les montagnes environnantes pour contempler la nature. Il faisait également beaucoup de sport comme la natation dans les cours d'eau de Jena, ou alors le saut en longeur où il remporta une compétition. Jena, berceau du romantisme à l'époque, était pour lui le paradis et un magnifique

endroit pour travailler. Il avait tout près de chez lui l'usine de Carl Zeiss où il pouvait venir autant qu'il le voulait pour adapter les objectifs de son microscope pour l'observation des radiolaires. Le grossissement optique atteignait celui des microscopes d'aujourd'hui et explique l'incroyable précision avec laquelle Haeckel dessinait les radiolaires. De plus, toutes ses planches étaient imprimées à deux pas de chez lui par l'imprimeur des biologistes de Jena. A l'époque, la lithographie était peu développée et il fallait donc un professionnel pour produire les planches. C'était un processus très long plus particulièrement quand les planches contenaient de nombreuses légendes. Il retournera plus de 20 fois en Italie tant il appréciait les paysages et le climat. Il retourne aussi à Helgoland avec un de ses étudiants Anton Dohrn, passé aussi par Jena, qui fondera quelques années plus tard la Stazione Zoologica à Naples. Ernst Haeckel continuera de voyager dans les pays tropicaux qu'il affectionnait tant durant son enfance. Il continuera à collecter, à dessiner et à décrire une pléthore d'organismes marins. Ses idées et concepts sur les formes, sur l'écologie, sur le développement et sur l'évolution auront une influence majeure dans la science mais aussi dans l'art et l'architecture.

A suivre….

Annexe 4: Travaux en collaboration

- **Lejzerowicz** J, Majewski W, Szczuciński W, **Decelle** J, Obadia C, Martinez P, Pawlowski J - Ancient DNA complements microfossil record in deep-sea subsurface sediments **- to be *submitted***

- **Sierra** R, Matz MV,Aglyamova G, Pillet L, **Decelle** J, Not F, de Vargas C and Pawlowski J. Deep relationships of Rhizaria revealed by phylogenomics: a farewell to Haeckel's Radiolaria - ***submitted to Molecular Phylogenetics and Evolution***

- **Karsenti** E, González-Acinas S, Bork P, Bowler ., de Vargas C, Raes J, Sullivan MB, Arendt D, Benzoni F, Follows M, Olivier Jaillon, Gabriel Gorsky, Pascal Hingamp, Iudicone D., Kandels-Lewis S., Krzic U., Not F., Ogata H., Pesant S., Reynaud E.G., Sardet C., Sieracki M.E., Speich S., Velayoudon D., Weissenbach J., Wincker P. & Tara Oceans Consortium (Abergel C, Arslan D, Audic S, Aury JM, Babic NS, Beaufort L, Bittner L, Boss E, Boutte C, Brum J, Carmichael M, Casotti R, Chambouvet A, Chang C, Chica C, Claverie J-M, Clerissi C, Colin S, Cornejo-Castillo FM, d'Ortenzio F, d'Ovidio F, Da Silva C, De Monte S, **Decelle** J, Desdevises Y, Dimier C, Dolan J, Duhaime M, Durrieu de Madron X, Ferrera I, Garczarek L, Garet-Delmas MJ, Gasol JM, Grimsley N, Heilig R, Ignacio-Espinoza J, Jamet JL, Karp-Boss L, Katinka M, Khalili H, Le Bescot N, Le Goff H, Lima-Mendez G, Mahé F, Mazzocchi MG, Montresor M, Morin P, Noel B, Pedrós-Alió C, Pelletier E, Perez Y, Picheral M, Piganeau G, Poirot O, Poulain J, Poulton N, Prejger F, Prihoda J, Probert I, Rampal J, Reverdin G, Romac S, Romagnan JB, Roullier F, Rouviere C, Samson G, Santini S, Sarmento H, Sciandra A, Solonenko S, Stemmann L, Subirana L, Sunagawa S, Tanaka A, Testor P, Thompson A, Tichanné-Seltzer V, Tirichine L, Toulza E, Veluchamy A, Zingone A (2011) - A holistic approach to marine eco-systems biology. ***PLoS Biology***, 9(10): e1001177, doi:10.1371/journal.pbio.1001177, IF: 12,9.

- **Pawlowski** J, Adl S, Audic S, Bass D, Belbahri L, Berney C, Bowser SS, Cepicka I, **Decelle** J, Dunthorn M, Fiore-Donno AM, Gile GH, Holzmann M, Jahn R, Jirků M, Keeling PJ, Kostka M, Kudryavtsev A, Lara E, Lukeš J, Mann DG, Mitchell EAD, Nitsche F, Romeralo M, Saunders GW, Simpson AGB, Smirnov AV, Spouge J, Stern RF, Stoeck T, Zimmermann J, Schindel D, de Vargas C. CBOL Protist Working Group: Barcoding eukaryotic richness beyond the animal, plant and fungal kingdoms **- *accepted in PLoS Biology***

- **Guillou** L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, de Vargas C, **Decelle J**, del Campo J, Dolan JR., Dunthorn M, Edvardsen B, Holzmann M, Kooistra W HCF, Lara Enrique, Lebescot N, Logares R, Mahé F, Massana R, Montresor M, Morard R, Not F, Pawlowski J, Probert I, Sauvadet AL, Siano R, Stoeck T, Vaulot D, Zimmermann P and Christen R - The Protist Ribosomal Reference database (PR^2): a catalog of unicellular eukaryote Small SubUnit rRNA sequences with curated taxonomy **– *accepted in Nucleic Acids Research***

Ancient DNA complements microfossil record in deep-sea subsurface sediments

Franck Lejzerowicz[1], Wojciech Majewski[2], Witold Szczuciński[3], Johan Decelle[4], Cyril Obadia[1], Pedro Martinez Arbizu[5], Jan Pawlowski[1]

[1]*Department of Genetics and Evolution, University of Geneva, CH 1211 Genève 4, Switzerland*

[2] *Institute of Paleobiology, Polish Academy of Sciences, 00-818 Warszawa, Poland*

[3] *Institute of Geology, Adam Mickiewicz University in Poznań, 61-606 Poznań, Poland*

[4] *CNRS UMR7144 EPPO & UPMC Université Paris 6, Station Biologique de Roscoff, France*

[5] *Pedro*

Keywords: Ancient DNA, deep-sea sediments, paleogenomics, microfossils, environmental sequencing

Abstract

Deep-sea subsurface sediments are the most important archives of marine biodiversity. Until now, these archives were based exclusively on microfossil record, disregarding large amounts of DNA accumulated on deep-sea floor. Accessing ancient DNA (aDNA) preserved down-core would offer unique insights into the history of marine biodiversity, including both fossilized and non-fossilized taxa. Here, we recover aDNA of eukaryotic origin across four cores collected at abyssal depths in the South Atlantic, in up to 32.5 thousand-year-old sediment layers. Our study focuses on foraminifera and radiolaria, two major groups of marine microfossils comprising also diverse non-fossilized taxa. We describe their assemblages in down-core sediment layers applying both micropaleontological approaches and environmental DNA sequencing. Short fragments of the foraminiferal and radiolarian SSU rRNA gene recovered from sedimentary DNA extracts provide evidence that eukaryotic aDNA is preserved in deep-sea sediments encompassing the Last Glacial Maximum. Most aDNA were assigned to non-fossilized taxa that also dominate in molecular studies of modern environments. Our study reveals the potential of aDNA to better document the evolution of past marine ecosystems and opens new horizons for the development of deep-sea paleogenomics.

To be submitted

1. Introduction

Deep-sea cores are commonly used to study paleoceanographical changes based on rich microfossil record composed of agglutinated, calcareous or siliceous skeletons of foraminifera and radiolaria. Both groups also comprise non-fossilized taxa, such as allogromiid foraminifera or Acantharia that have been shown to be particularly diverse and abundant according to recent metagenetic surveys [1,2]. These and other environmental DNA studies [3, 4] confirm that the deep-sea floor is a molecular repository for virtually all groups of organisms living in the open ocean. Most of this DNA is mainly extracellular and its concentration is high, reaching 0.3 g m^{-2} [5]. Theoretically, the stable abiotic conditions at the deep-sea bottom, including low temperature, great pressure and small disturbance enhance DNA preservation [6], as shown by 125,000-year-old eukaryotic aDNA sequences successfully amplified from eastern Mediterranean sapropels [7]. However, the preservation of aDNA in the deep sea was demonstrated only in exceptional, anoxic settings [8]. Here, we show that eukaryotic aDNA can also be recovered from oxic, subsurface sediments at abyssal depths. We focused on foraminifera and radiolaria, comparing their fossil record to aDNA sequences obtained from down-core sediment samples. Our results bring evidence for the preservation of aDNA in deep-sea sediments, and demonstrate that subsurface sampling of abyssal sediments could be envisioned for paleogenomic studies.

2. Material and Methods

Deep sea sediment from the South Atlantic abyssal plain was sampled using a multicorer in two stations at ~4470 and ~5180 meters water depth during the *RV Meteor* cruise M79-1 (electronic supplementary material, table S1). For each station, two cores (550/552 and 600/601) showing no evident signs of bioturbation were selected and 20-30 g of sediment subsampled every 10 cm. The 30-centimeter layer of cores 550 and 601 were sampled for planktonic foraminifera for age determination using accelerator mass spectrometry (AMS) ^{14}C. All the sediment samples were analyzed for volumetric grain size distribution with a Malvern Mastersizer 2000 Particle Analyser and the microfossil assemblage described. For methodological details, see the 14C dating, grain size and micropaleontological analysis sections in the electronic supplementary material.

Several routine precautions were taken to preserve DNA and avoid contamination by surface material: samples were taken from the central part of the core using sterilized spoons and flame-decontaminated spatulas; gloves were worn at all time; sub-sampling took place in the cold room and samples were directly frozen at -20°C. For each layer, total sedimentary DNA was extracted (blank extraction ratio 1:2) from ~0.5 g of sediment using the MoBio Power Soil Kit in laboratories with no history of previous radiolaria or foraminifera molecular work (DTAMB and PALGENE platforms; Lyon, France). DNA extracts were PCR amplified in dedicated hoods using either radiolarian or foraminiferal-specific primers in order to target small subunit ribosomal RNA gene fragments of ~270 bp or ranging from 68 to 196 bp, respectively (electronic supplementary material, table S2). Controlled PCR amplifications (blank ratio 1:2 to 1:3) were cloned and sequenced as in [9]. Foraminiferal-specific 37f hypervariable sequences were assigned as in [9] and radiolarian-specific sequences were assigned as in [10]. Sequences were clustered into operational taxonomic units (OTUs) at 3% divergence using the average neighbor method in mothur [11] based on pairwise Needleman-Wunsch distances and OTUs assignations were corrected according to the consensus. Additional PCR amplifications were conducted using primer sets targeting foraminiferal fragments of ~400 bp and ~1000 bp (electronic supplementary material, table S2).

3. Results

Sediment type and age

The sediments are deep-sea muds composed mainly of silt. They contain 30 to 55% of very fine silt and clay fractions, as well as up to 4% of sand, mainly of biogenic origin. They are classified as clay-bearing and clayey nanofossil oozes. In all cores the grain size is slightly decreasing with increasing depth (electronic supplementary material, table S3, figures S1 and S2).

The age of sediments at 30 cm core depth was 11,870 ± 70 ^{14}C years BP (12,790-13,245 calibrated years BP) and 28,950 ± 260 ^{14}C years BP (32,050-33,230 calibrated years BP) for cores 550 and 601 respectively (electronic supplementary material, table S4). The calculated averaged linear sediment accumulation rates are in order of 2.3 cm/kyr in core 550 and 0.9 cm/kyr in core 601.

Microfossils

In all cores, diverse and well-preserved radiolarian and foraminiferal microfossil assemblages were encountered (electronic supplementary material, figures S2 and S3). A clear down-core trend of decreasing frequencies and abundances of radiolaria and increasing foraminifera is observed (electronic supplementary material, tables S5 and S6). It was well confirmed by qualitative inspection of the glass slides with the <20 μm fraction (electronic supplementary material, figure S4). Regarding the planktonic foraminifera (Globigeriniidae), none was found at the sediment surface except in core 600 but they appear in considerable numbers below 20 cm. Moreover, in the lower parts of cores 600 and 601, a coarser mineral material (> 20 μm) is dominated by disintegrated tests of planktonic foraminifera (figure S4D) while in the higher latitude cores 550 and 552, mineral particles are not biogenic (figure S4B).

Molecular data

DNA sequences were obtained from almost all down-core samples, but the DNA showed a clear pattern of degradation with depth. The size of successfully amplified fragments with diverse foraminifera-specific primers decreased from ~1000 bp at the surface to ~350 bp in the 10 cm layer and to less than 100 bp in lower layers (electronic supplementary material, figure S5).

In total, we obtained 627 foraminiferal sequences, of which 361 were from samples below 10 cm (figure 1). All sequences clustered in 110 OTUs, belonging mainly to non-fossilized monothalamous taxa (49%). Below the 10-centimeter layer, the assemblages of foraminifera inferred from older samples display lower diversity. In agreement with micropaleontological observations, the sequences of calcareous benthic species (Rotaliida) were nearly absent at the surface, but also rare in deeper layers in spite of the abundance of rotaliids tests. Two species of Globigeriniidae, *Globorotalia inflata* and *Globigerinita uvula* were detected in the molecular data, but only with very short (82 and 89 bp, respectively) sequences. Additionally, the DNA sequence of the deep-sea rotaliid *Pullenia subcarinata* abundant in the fossil record of core 600 was retrieved from the corresponding layer using species-specific primers also targeting a short fragment (84 bp).

In the case of planktonic radiolarians, we obtained 109 sequences, representing 21 OTUs assigned to major radiolarian groups: Taxopodida (10 OTUs), Acantharia (9 OTUs) and Spumellaria (2 OTUs). Since the primers used were not absolutely specific, 46 sequences were assigned to other eukaryotic groups (e.g. Cercozoa. Stramenopiles). There is a clear decrease in radiolarian diversity, from 15 OTUs at the surface to less than 5 OTUs below 10 centimeters. The non-fossilized Taxopodida appeared to be a major component in radiolarian sequences, particularly at the surface (8 OTUs). Distinct Taxopodida sequences were also found down to 10 and 20 centimeters in station 552 (2 OTUs). While Acantharia was absent in

micropaleontological observations, 52 sequences corresponding to 9 OTUs were found in every studied layer. None of the 6 OTUs found at the surface was detected below 10 centimeters in core 600, and 2 other acantharian OTUs were obtained down to the 40 cm in station 552, i.e. below the layer dated at 13,017 ± 227.5 calibrated years BP. One OTU assigned to Spumellaria, the major group of radiolarian observed in microfossils (figure 1), was detected in the oldest radiocarbon dated layer of the study (core 600, 30 cm).

4.Discussion

Authenticity of DNA sequences is the most recurrent problem in the aDNA studies, particularly those investigating ancient environmental diversity. We tackle this problem by applying several precautions during sampling procedure, including a dedicated paleogenomic platform, remote laboratories for DNA extraction and amplification, and frequent blank controls. We avoid eukaryotic contamination by using specific PCR primers that enrich selected groups of marine protists. Finally, the absence of high molecular weight DNA fragments in down-core samples excludes cross-contamination or leaking of DNA from the core surface as well as the vertical migration of benthic specimens. The PCR amplification pattern confirms the absence of living foraminiferans below 10-15 cm, considered as the lower limit of infaunal species distribution in deep sea sediments [12]. In addition to all these precautions, the finding of sequences assigned to well-known planktonic and benthic species from the fossil record supports the authenticity of fragmented but still amplifiable deep-sea aDNA sequences.

Our study shows that aDNA is a promising tool to detect both fossil and non-fossilized taxa in oxic subsurface abyssal sediments. The dominance of sequences from non-fossilized taxa confirms the results of environmental DNA surveys of modern samples [1,2]. Some of these taxa, like Acantharia, produce heavy mineralized cysts rapidly sinking to the sea floor without being grazed or decomposed, increasing the chance of DNA preservation in ancient sediments. Another factor that might be of importance regarding the preservation of DNA is the composition of the sediment substrate. Clay minerals present the highest capacity to adsorb and hence protect DNA from nucleases [13] but in our study, the clay and very fine silt fraction represents less than 50% of the poorly sorted sediments. We therefore suppose that amplifiable aDNA could be recovered from much older deep-sea sediment, provided a high content of phyllosilicates (e.g. deep-sea “red clays”).

Further investigations of aDNA in deep-seas will require larger sediment volumes and high-throughput sequencing in order to increase the sequence sampling coverage. Prior to sequencing, highly specific primers could be used to detect marine cryptic species and refine past distribution patterns of paleoceanographical indicators (e.g. planktonic foraminifera [14]), while universal eukaryotic primers will better document faunal changes by providing extensive data on fossilized and non-fossilized taxa. In the context of recent paleoclimatic studies, analyses of micro-eukaryotic aDNA from deep-sea subsurface sediment could complement micropaleontological research on the past hundred thousand years climate changes.

The authors thank captain and crew of R/V Meteor for help in collecting samples, and Michał Rzeszewski for grain size analysis. This study was supported by the Swiss National Science Foundation grant 31003A-140766.

REFERENCES

1. Lecroq, B., Lejzerowicz, F., Bachar, D., Christen, R., Esling, P., Baerlocher, L., Østerås, M., Farinelli, L. & Pawlowski, J. 2011 Ultra-deep sequencing of foraminiferal microbarcodes unveils hidden richness of early monothalamous lineages in deep-sea sediments, *Proc. Natl. Acad. Sci. USA.* 108, 13177–13182. (doi:10.1073/pnas.1018426108)

2. Not, F., Del Campo, J., Balagué, V., De Vargas, C. & Massana, R. 2009 New insights into the diversity of marine picoeukaryotes. *PLoS One* 4, e7143. (doi:10.1371/journal.pone.0007143)

3. Bik, H. M., Sung, W., De Ley, P., Baldwin, J. G., Sharma, J., Rocha-Olivares, A. & Thomas, W. K. 2011 Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. *Mol. Ecol.* 21, 1048–1059. (doi:10.1111/j.1365-294X.2011.05297.x)

4. Pawlowski, J., Christen, R., Lecroq, L., Bachar, L., Shahbazkia, H. R., Amaral-Zettler, L. & Guillou, L. 2011 Eukaryotic richness in the abyss: insights from pyrotag sequencing. *PLoS One* 6, e18169. (doi:10.1371/journal.pone.0018169)

5. Dell'Anno, A. & Danovaro, R. 2005 Extracellular DNA plays a key role in deep-sea ecosystem functioning. *Science* 309, 2179. (doi:10.1126/science.1117475)

6. Corinaldesi C., Barucca, M., Luna, G. M. & Dell'Anno, A. 2011 Preservation, origin and genetic imprint of extracellular DNA in permanently anoxic deep-sea sediments. *Mol. Ecol.* 20, 642–654. (doi:10.1111/j.1365-294X.2010.04958.x.)

7. Boere, A. C., Rijpstra, W. I. C., de Lange, G. J., Sinninghe Damsté, J. S. & Coolen, M. J. L. 2011 Preservation potential of ancient plankton DNA in Pleistocene marine sediments. *Geobiology* 9, 377–393. (doi:10.1111/j.1472-4669.2011.00290.x)

8. Coolen, M. J. L. 2011 7000 Years of *Emiliania huxleyi* Viruses in the Black Sea. *Science* 333, 451–452. (doi:10.1126/science.1200072)

9. Lejzerowicz, F., Voltsky, I. & Pawlowski, J. 2012 Identifying active foraminifera in the Sea of Japan using metatranscriptomic approach. *Deep-Sea Res. in press.* (doi:10.1016/j.dsr2.2012.08.008)

10. Decelle, J. J., Suzuki, N. N., Mahé, F. F., de Vargas, C. & Not, F. F. 2012 Molecular phylogeny and morphological evolution of the Acantharia (Radiolaria). *Protist* 163, 435–450. (doi:10.1016/j.protis.2011.10.002)

11. Schloss, P. D. *et al.* 2009 Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. (doi: 10.1128/ AEM.01541-09)

12. Corliss, B. H. 1985 Microhabitats of benthic foraminifera within deep-sea sediments. *Nature* 314, 435–438. (doi:10.1038/314435a0)

13. Cai, P., Huang, Q., Zhang, X. & Chen, H. 2006 Adsorption of DNA on clay minerals and various colloidal particles from an Alfisol. *Soil Biol. Biochem.* 38, 471–476. (doi:10.1016/j.soilbio.2005.05.019)

14. Darling, K. F., Kucera, M. & Wade, C. M. 2007 Global molecular phylogeography reveals persistent Arctic circumpolar isolation in a marine planktonic protist. *Proc. Natl. Acad. Sci. USA.* 104, 5002–5007. (doi: 10.1073/pnas.0700520104)

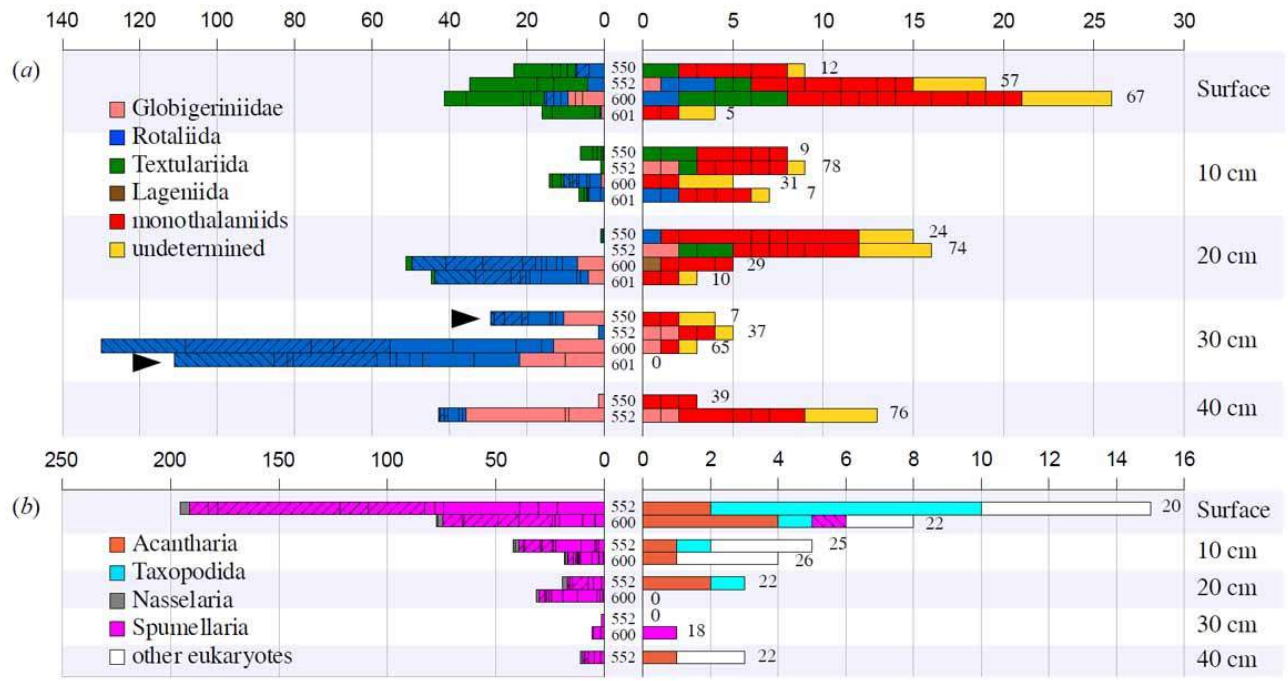


Figure 1. Taxonomic composition of foraminifera (a) and radiolaria (b) based on classical micropaleontological counts (left panels) and analysis of DNA sequences (right panels). For each sediment layer, the cores are presented in following order, from top to down: 550, 552, 600 and 601. Only cores 552 and 600 were investigated for radiolaria and only cores 550 and 552 have a 40 cm layer. Black arrows in top left panel correspond to radiocarbon-dated layers (core 550: 12,790-13,245 ^{14}C calibrated years BP and core 601: 32,050-33,230 ^{14}C calibrated years BP). On right panels, the number of DNA sequences obtained for each layer are indicated and each block of the bars corresponds to an assigned taxa. On the left panel, the proportions of extremely abundant fossil Rotaliida were reduced (hatched bars; *Oridorsalis*, *Pullenia* and *Osangularinella*: 50% displayed, *Epistominella*: 10% displayed). Among the Spumellaria, the Spongodiscidae are displayed in hatched, pink bars. Globigeriniidae: planktonic foraminifera.

Deep relationships of Rhizaria revealed by phylogenomics: a farewell to Haeckel's Radiolaria

Roberto Sierra[a,*], Mikhail V. Matz[b], Galina Aglyamova[b], Loïc Pillet[a], Johan Decelle[c], Fabrice Not[c], Colomban de Vargas[c] and Jan Pawlowski[a]

[a]Department of Genetics and Evolution, University of Geneva, Geneva, Switzerland

[b]The University of Texas at Austin, Austin, Texas, United States of America

[c]CNRS & UPMC, UMR-7144, *Evolution du Plancton et PaléoOcéans*, Station Biologique, BP 74, 19682 Roscoff, France

* Corresponding author. Address: Quai Ernest Ansermet 30, University of Geneva, 1211 Geneva 4, Switzerland Fax: +41 22 379 3340.

E-mail address: Roberto.Sierra@unige.ch

Abstract

Rhizaria is one of the six supergroups of eukaryotes, which comprise the majority of amoeboid and skeleton-building protists living in freshwater and marine ecosystems. There is an overall lack of molecular data for the group and therefore the deep phylogeny of rhizarians is unresolved. Molecular data are particularly scarce for the clade of Retaria, which include two prominent groups of microfossils: foraminiferans and radiolarians. To fill this gap, we have produced and sequenced EST libraries for 14 rhizarian species including seven foraminiferans, *Gromia* and six taxa belonging to traditional Haeckel's Radiolaria: Acantharea, Polycystinea, and Phaeodarea. A matrix was constructed for phylogenetic analysis based on 109 genes and a total of 56 species, of which 22 are rhizarians. Our analyses provide the first multigene evidence for branching of Phaeodarea within Cercozoa, confirming the polyphyly of Haeckel's Radiolaria. It confirms the monophyly of Retaria, a clade grouping Foraminifera with other lineages of Radiolaria. However, contrary to what could be expected from morphological observations, Foraminifera do not form a sister group to radiolarians, but branch within them as sister to either Acantharea or Polycystinea depending on the multigene data set. While the monophyly of Foraminifera and Acantharea is well supported, that of Polycystinea, represented in our data by Spumellaria and Collodaria is questionable. In view of our study, Haeckel's Radiolaria appears as both, a polyphyletic and paraphyletic assemblage of independent groups that should be considered as separate lineages in protist classification.

Submitted to Molecular Phylogenetics and Evolution

1. Introduction

Rhizaria is one of the six supergroups of eukaryotes established based exclusively on molecular characters (Cavalier-Smith, 2002; Nikolaev et al., 2004). The rhizarians comprises various flagellate and amoeboid protists, including the two most important groups of skeleton-building microfossils: foraminiferans and radiolarians. These large skeletonized organisms are key players in marine ecosystems as grazers, primary producers through endosymbiosis with microalgae and carbon exporters to deep oceans. They are able to mineralize and build elaborate skeletons of calcium carbonate in the case of some groups of Foraminifera, silica in Polycystinea or strontium sulfate in Acantharea. Phylogenetic relationships between these two groups inferred from molecular data have always been controversial (Pawlowski and Burki, 2009). Based on analyses of rRNA genes, some authors placed them together in a new assemblage named Retaria, composed of Foraminifera and Radiozoa (Cavalier-Smith, 2002; Moreira et al., 2007). Recently, phylogenomic analyses confirmed the Retaria hypothesis but the multigene data were available for one group of radiolarians only (Burki et al., 2010; Burki et al., 2012).

Remarkably, the ribosomal trees with broad taxon sampling show Foraminifera branching within radiolarians, usually as sister group to Polycystinea (Krabberod et al., 2011; Moreira et al., 2007). This unusual branching suggesting the paraphyletic nature of radiolarians could also be observed in the actin phylogeny (Burki et al., 2010). However, the correct branching of Foraminifera in ribosomal RNA phylogenies is impeded by extremely rapid evolutionary rate of their rRNA genes (Pawlowski et al., 1996). This variation of rDNA substitution rate also hampers a statistically reliable inference of phylogenetic relationships between different radiolarian groups. In ribosomal RNA trees, Polycystinea are usually separated into Collodaria, Spumellaria and Nassellaria (Kunitomo et al., 2006; Yuasa et al., 2005). Depending on the type of analyses and taxon sampling, Collodaria branch with Nassellaria, while Spumellaria branch as a discrete lineage or sister to Acantharea. Recently, based on combined SSU and LSU rDNA data, it has been proposed that Acantharea cluster with Taxopodida, forming together a group named Spasmaria, but Collodaria were not included in this analysis (Krabberod et al., 2011).

To resolve the phylogeny of radiolarians and to test their phylogenetic relationships to Foraminifera, we have collected a data set of 109 genes and 22 rhizarian species, which constitutes the largest and most complete rhizarian phylogenomic data set available to date. Our study included the ESTs of Foraminifera, Phaeodarea, Acantharea and Polycystinea (Spumellaria and Collodaria). The phylogenetic analyses confirm the polyphyly of Haeckel's Radiolaria (Haeckel, 1887, 1904) and indicate the paraphyletic character of radiolarian lineages that branch with Foraminifera.

2. Materials and methods

2.1 Collecting and isolation of specimens

Seven species of Foraminifera were examined in this study. Six of them (*Elphidium* sp., *Globobulimina turgida*, *Brizalina* sp., *Bulimina marginata*, *Nonionellina* sp., and *Ammonia* sp.) were collected in European coastal waters in different localities indicated in Table S1. One species (*Reticulomyxa filosa*) was maintained in the lab and handled as previously described (Burki et al., 2006). Collected specimens were picked from algal and sediment samples immediately after sampling, thoroughly cleaned with a fine brush and washed with filtered seawater. Clean specimens were stored in RNA*later* (Ambion, Austin, TX) until further processing. The number of sorted specimens varied from 40 for *Nonionellina* sp. to 3000 for *Elphidium sp*. (Table S1).

Three species of radiolarians (*Amphilonche elongata*, *Collozoum* sp., *Spongosphaera streptacantha*) and *Aulacantha scolymantha* were collected in Red or Mediterranean Seas (Table S1). The specimens were sorted from planktonic samples under a stereo microscope and stored in RNA*later* until further processing. The number of cells used for library preparation varied between 50 and 150, with the exception of *Collozoum* sp., for which a single large colony was used for RNA extraction (Table S1).

2.2 Preparation of cDNA libraries for Sanger sequencing

RNA was isolated from *Elphidium* sp., *G. turgida* and *R. filosa* at Vertis Biotechnology AG (Germany), using the mirVana miRNA isolation kit (Ambion). From total RNA, polyA+ was prepared, and cDNA was synthesized according to the Vertis Biotechnology AG standard protocol for full-length enriched cDNA using an oligo(dT)-linker primer for first-strand synthesis. Before cloning, the cDNA was amplified with 15 cycles of PCR. For directional cloning, cDNA was subjected to a limited exonuclease treatment to generate *Eco*RI overhangs at both ends of the cDNAs. Size-fractioned cDNA fractions >0.5 kb were ligated into *Eco*RI and *Bam*HI sites of plasmid vector pBS II SK+ and subsequently transformed via electroporation into T1 phage-resistant NEB 10-beta electrocompetent *E. coli* cells (New England Biolabs). The transformants were added glycerol to a final concentration of 12.5% (v/v) and stored at −70 ºC. End-sequencing was performed on plasmid DNA isolated from ~20,000 clones of the cDNA libraries by a single pass sequence from the 5′ end with a primer specific for the pBS II SK+ vector at Genoscope (Evry, France) using an ABI 3730 automatic capillary sequencer and the ABI BigDye Terminator v.3.1 sequencing kit.

2.3 Preparation of cDNA libraries for 454 sequencing

Total RNA was isolated from *B. marginata*, *Brizalina* sp., *Nonionellina* sp., *Ammonia* sp., *Collozoum* sp., *A. elongata*, *S. streptacantha*, *A. scolymantha* samples using the RNAqueous-Micro kit (Ambion, Austin, TX) or NucleoSpin RNA XS (Macherey-Nagel, Germany). Immediately after RNA isolation, cDNA synthesis and amplification was performed using the SMARTer RACE cDNA amplification kit (TaKaRa BIO/Clontech, Mountain View, CA) followed by the library preparation procedure modified from Meyer, 2009 #4433(the current version of the protocol is available on Matz lab website: http://www.bio.utexas.edu/research/matz_lab/).

Approximately 2.5 µg of the cDNA pool from previously prepared libraries of *P. siculus*, *A. serrata* and *G. sphaerica* described in (Burki et al., 2010) and the newly prepared libraries were used for a titration run using one-quarter of a plate for each sample on the Roche 454 Genome Sequencer FLX using GS-FLX Titanium series reagents.

2.4 Contig assembly and sequence alignment

For the Sanger data, we obtained base calls and quality values using phred (Ewing and Green, 1998; Ewing et al., 1998), we removed cloning vector sequences and assembled the ESTs using phrap implemented in the Bioportal (http://www.bioportal.uio.no/). Adaptor sequence trimming and assembly of the 454 data were performed using the Newbler software v. 2.6 (Roche) implemented in the Vital-IT (http://www.vital-it.ch/). Our 109-protein phylogenomic data set was constructed using as a starting point an existing 167-protein data set (Burki et al., 2010). Rhizarian sequences were assigned and added to the preexisting alignments of 167 proteins based on homology searches using blastp. To assign new

sequences to each gene, we used the previously published alignments in amino acids as a query and the translated EST data, in the 6 possible reading frames, as the database. These alignments included representative species for all major groups of eukaryotes. All sequences that had a plant or animal as a best hit with an e-value $\leq 10^{-3}$ were excluded from the protein alignments. The homologous sequences to each gene and for all species obtained were included in the alignment (e-value cut-off of 10^{-5}). If there was more than one homologous sequence of the same species for a particular gene, the longest aligned sequence was kept for further analyses. The missing data were reduced for the preexisting species (*R. filosa*, *A. serrata*, *P. siculus* and *G. sphaerica*) and new sequences were added and automatically aligned to the data set using MAFFT v. 6.847b (Katoh et al., 2002). Ambiguously aligned positions were removed using Gblocks v. 0.91b (Castresana, 2000) allowing half of the gapped positions, 50% of the number of taxa plus one as the minimum number of sequences for a flanking position, maximum number of contiguous non-conserved positions was set to 12, and 5 amino acids as a minimum block. Each protein alignment was manually examined to remove misaligned sequences and all short sequences (less than ~30% of the total alignment length).

2.5 Phylogenetic analyses

The single-gene maximum likelihood (ML) trees were performed on RAxML using the PROTCATLGF setting with 100 bootstrap replicates. All trees were carefully examined in order to discard any sequence that would branch with plants or animals. We also removed sequences that branch with other groups of eukaryotes if this branching was supported by a bootstrap value of 70 or higher, paying special attention to known symbionts. We also took into consideration that clearly monophyletic groups such as Foraminifera or Acantharea should be retrieved in this analysis even though Rhizaria would not be monophyletic. If, after cleaning the alignments by inspecting the single gene trees, rhizarians were underrepresented (less than 3 species in two different groups) the complete protein alignment was not used for further analyses. This cautious purging of the data resulted in 109 genes (SAR109). Additionally, a subset of 36 genes (SAR36) that met the above criteria and also showed a monophyletic Rhizaria was analyzed separately. The super-matrices were constructed using SCaFoS (Roure et al., 2007).

The ML analyses were performed using RAxML v.7.2.8 (Stamatakis, 2006). The best ML tree was determined with the PROTGAMMALGF implementation in multiple inferences using 20 and 30 randomized parsimony starting trees; statistical support was evaluated with the PROTCATLGF setting due to computational constraints with 500 and 1000 bootstrap replicates for the 109 and 36 genes matrices, respectively. The bootstrap values of the consensus tree were mapped on to the highest scoring ML tree. In order to assess the best fit model for Bayesian Inference (BI) analysis, CAT and LG models were tested on the data using the cross-validation implemented in PhyloBayes v.3.2f (Lartillot et al., 2009), as described in the manual of the program. The MCMC for each learning set ran under a fixed tree topology estimated by the model itself on the full dataset. BI were carried out under the CAT-Poisson model with 4 independent chains for ~18,000 and ~49,000 cycles for the 109 and 36 genes matrices, respectively. Besides, BI was also carried out under the CAT-GTR model with 4 independent chains for ~6,000 cycles for the 36 gene matrix. For post-analysis of the independent chains a 20% burnin was used. The lack of convergence (maxdiff > 0.3) suggests that the chains did not run for enough generations, however we assessed each chain individually and they illustrate that the topology were congruent, in respect to the taxa of interest.

For the phylogenetic analysis based on actin, the best-fit models were calculated using Mega5 (Tamura et al., 2011). ML was assessed using the WAG+Γ model (Whelan and Goldman, 2001), and

statistical analysis was obtained with 1000 bootstrap replicates on RAxML. BI were carried out using the WAG model as implemented in PhyloBayes with 4 independent chains and 19,038 cycles and MrBayes v.3.2 (Ronquist and Huelsenbeck, 2003) for 7,935,000 generations until convergence of the chains was reached with a maxdiff <0.25 and average standard deviation of split frequencies <0.01 for PhyloBayes and MrBayes, respectively. Bayesian consensus posterior probabilities of post-burnin (20%) bipartitions were mapped to the corresponding best ML tree.Topology comparisons were conducted using the approximately unbiased (AU) (Shimodaira, 2002) and SH (Shimodaira and Hasegawa, 1999) tests. A set of four alternate plausible hypothesis plus the 500 and 1000 bootstrap trees for SAR36 and SAR109, respectively were used to calculate the site likelihoods using the PROTGAMMALGF model implemented in RAxML. The AU and SH tests were performed using CONSEL (Shimodaira and Hasegawa, 2001).

3. Results

3.1 Sequencing and assembly of data set

Approximately 20,000 Sanger sequenced clones were used for the assembly of 3 libraries and quarter-plate runs for the eleven 454 libraries constructed. The average contig lengths ranged between 597 and 1,358 bp and N50s calculated for each dataset (contigs + singletons) ranged between 317 and 1,385 bp (Table S2). The complete data set comprised 22 rhizarian, 12 stramenopile, 15 alveolate taxa that form the SAR group, as well as 7 haptophytes used as outgroup. Two matrices of respectively 24,682 and 9,825 amino acid positions were constructed, one with 109 proteins and 56 taxa (SAR109) and the other with a subset of 36 proteins and 54 taxa. The missing data was 58% for SAR109 and 59% for SAR36 (Fig. 1).

3.2 Phylogeny of Rhizaria

The two matrices were analyzed using ML with LG+Γ4 model. The cross validation results favored CAT over LG model with a likelihood score of 276.64 ± 31.9 and 1007.12 ± 59.9 for the SAR36 and SAR109 matrices, respectively and BI analyses were carried out under the best-fit model. In the ML analyses (Figs. 2 and S1), we recovered the same relationships between the three groups composing SAR (Stramenopiles + Alveolata + Rhizaria) with Rhizaria branching as sister to Alveolata with 82 bootstrap (BS) support value in SAR109 but not supported in SAR36. The BI consensus tree for SAR109 recovered the same topology as the ML analysis with a posterior probability (PP) of 0.99 but for SAR36 we recovered Rhizaria sister to Stramenopiles and Alveolata at the base (Suppl. Figs. 2 -9). The individual chains of the Bayesian analyses were carefully inspected to note if any other major topological differences were recovered but only finding minor differences within derived groups.

The analyses of relationships within Rhizaria showed different topologies depending on the number of analyzed proteins. In the case of SAR109, the rhizarians were split into two clades: the clade of Retaria that was recovered with maximum support and a second clade grouping Cercozoa (including Phaeodarea), Plasmodiophorida and *Gromia + Filoreta* but this clade was much less supported (86BS/-PP, Fig. S1). However, *Gromia + Filoreta* always formed a strongly supported group. The SAR36 showed maximum support for Retaria as well, but other rhizarian groups branched independently with Plasmodiophorida at the base, followed by Cercozoa and *Gromia + Filoreta* clade (Fig. 2).

The relationships within Retaria were slightly different in both 109 and 36 proteins data sets. All analyses recovered strongly supported monophyly of Foraminifera and Acantharea. However, the clade of

Collodaria + Spumellaria (Polycystinea) was strongly supported only in 36 proteins analysis. The branching order of radiolarian groups was also different. In the SAR109 data set, the Acantharea branched at the base followed by Polycystinea and Foraminifera, while in the SAR36 data set, the branching order of Acantharea and Polycystinea were reversed. Remarkably, the Foraminifera never branched at the base of Retaria (Fig. 2).

3.3 Actin phylogeny

Because of much better taxon sampling, we performed a separate phylogenetic analysis of actin gene, including Haplosporidia and more numerous sequences of Foraminifera, Spumellaria, Collodaria and Cercozoa. The alignment contained 315 amino acid positions for 96 rhizarians and 7 alveolates used as outgroup. The ML analysis shows a monophyletic group of cercozoans including Phaeodarea at the base of Rhizaria. The two actin paralogs previously described for Foraminifera (Flakowski et al., 2006) and Acantharea (Burki et al., 2010) were recovered, but only one paralog was found for most of sequenced collodarians and spumellarians, except for a sequence of spumellarian *Larcopyle butschlii* obtained by other authors (Ishitani et al., 2011). One paralog was also found in the case of *Filoreta*, *Gromia*, Plasmodiophorida and Haplosporidia, although in this later case, one sequence (*Urosporidium crescens*) branched separately at the base of plasmodiophorids (Fig. S1). When the second paralog present in Retaria was removed, the relationships within Rhizaria resembled those inferred from multigene analyses. Interestingly, Haplosporidia branched with *Gromia* and *Filoreta*, but the support was not strong.

3.4 Topology tests

A topological constraint consisting of a monophyletic Radiolaria/Radiozoa sister to Foraminifera was tested. There is weak evidence to consider the alternate topology significantly worse (p-values ≤ 0.05) than the best ML tree based on the SAR36 data set, but not for the SAR109 data set (Table 1, AU tests). We also tested which radiolarian group, Polycystinea or Acantharea, was more closely related to Foraminifera within the retarian clade. Using the SAR36 matrix, the AU topology test indicated that the sister relationship between foraminiferans and Polycystinea is significantly worse than Foraminifera as sister to Acantharea at a 1% level. The topology with Foraminifera sister to Polycystinea is very strongly rejected when using the SAR109 data set. Finally, topological contraints of the paraphyly of Polycystinea was not supported in any data set when placing *Collozoum* sp. (Collodaria) at the base of Retaria. Interestingly, the paraphyly of Polycystinea when *S. streptacantha* (Spumellaria) is at the base of Retaria was not rejected in either case.

4. Discussion

4.1 Challenges of rhizarian phylogenomics

Several factors explain the relatively low number of rhizarian genomic and transcriptomic data as compared to other eukaryotic supergroups. Rhizaria are composed of many taxonomic groups that are predominantly uncultivable, so that their DNA and RNA must be extracted from specimens collected in the field, what can pose several problems for their use in phylogenomic studies. First, taxonomic identification can be problematic, and many groups have limited or lack morphological characters to clearly distinguish between species and even genera. Second, collecting a sufficient number of cells

belonging to the same species from a single location can be challenging, and thus DNA and RNA extraction yields are frequently too low. Third, the specimens may contain many foreign organisms that live outside or inside their cells and which typically contaminate the DNA or RNA extracts. This is particularly challenging in the case of Foraminifera and Radiolaria that often build large skeletal structures that can host multiple organisms, including small size representatives of the same taxonomic groups (Lecroq et al., 2009).

To address this issue, it is very important to ensure that the analyzed genes belong to the species of interest. Indeed, in our cDNA libraries we observed sequences that branched with different groups of eukaryotes. However, the general paucity of available genomes/transcriptomes makes it sometimes difficult to attribute a given sequence to any known eukaryotic group. In order to minimize uncertainties and since our aim was not to test the monophyly of Rhizaria, we constructed a data set of 36 genes displaying significant rhizarian monophyly in the independent single-gene phylogenetic analyses. Additionally, the Foraminifera and Acantharea appeared monophyletic for almost all examined genes.

4.2 Polyphyly of Haeckel's Radiolaria

The classical view of Radiolaria, composed of Phaeodarea, Acantharea and Polycystinea, was first challenged based on SSU rDNA analysis showing that Phaeodarea branch as independent lineage within Cercozoa (Polet et al., 2004). At that time, we were unable to amplify any protein coding genes from phaeodarean DNA extracts (Nikolaev et al., 2004). Moreover, additional phaeodarean SSU rDNA sequences were reported by (Yuasa et al., 2006). Here, we provide the first multigene data for a phaeodarean species, *Aulacantha scolymantha*. The analysis confirms the separation of Phaeodarea from other radiolarians, but their position within Cercozoa remains unresolved. The phaeodarean clustering with *Bigellowiella* in 109 genes analysis is weakly supported (Fig. 2), in the same way as the sister relationship to *Spongomonas* sp. in the actin tree (Fig. S1).

4.3 Paraphyly of Radiozoa

The radiolarian groups that branched together after removing the Phaeodarea were named Radiozoa by Cavalier-Smith (2003). In most of ribosomal phylogenies, Radiozoa appeared as a monophyletic assemblage in the absence of Foraminifera (Takahashi et al., 2004). However, when the foraminiferal sequences were included, they always branched within radiozoan radiation (Cavalier-Smith, 2003; Krabberod et al., 2011; Moreira et al., 2007). Our phylogenomic analysis demonstrates that the position of Foraminifera within Radiozoa is not an artifact related to the exceptionally high substitution rates of foraminiferal ribosomal genes (Pawlowski and Burki, 2009). Depending on the multigene data set analyzed herein, the Foraminifera form a sister group to either Acantharea or Polycystinea. The actin paralog 2 also supports the Foraminifera forming a sister group to Acantharea and Spumellaria+Collodaria at the base of Retaria. Our statistical tests of alternate topologies (Table 1) favor the Foraminifera-Acantharea sisterhood. We expect that the derived position of Foraminifera within Retaria will not change by adding to the phylogenomic analysis the radiolarian lineages for which EST data is not available yet, such as Taxopodida and Nassellaria. Therefore, Radiozoa appears to be a paraphyletic group comprising various radiolarian lineages, as well as foraminiferans.

4.4 The uncertain monophyly of Polycystinea

Our study contributes with the first multigene data for one collodarian and one spumellarian species, which in traditional taxonomy belong to the class Polycystinea. The bootstrap values for their grouping in

our analyses are relatively high (Figs. 2 and 3). However, the Polycystinea monophyly is doubtful. In fact, in previous SSU rDNA-based studies, with broader taxon sampling of the three polycystine groups (Spumellaria, Collodaria, Nassellaria), these never branched together (Krabberod et al., 2011; Yuasa et al., 2005). The colonial and naked Collodaria branch either as sister group to Nassellaria (Yuasa et al., 2005) or at the base of radiolarians (Takahashi et al., 2004), while the solitary, shell bearing Spumellaria branch as sister to Acantharea (Yuasa et al., 2005). Recent analyses of combined 18S and 28S rDNA data strongly support the monophyly of Spumellaria and Nassellaria, albeit in the absence of Collodaria (Krabberod et al., 2011). Although our data supports the monophyly of Polycystinea, additional multigene data from polycystine lineages will certainly be needed to accurately challenge their monophyly.

4.5 Taxonomic conclusions

Previous classifications of Rhizaria were based exclusively on SSU rDNA analyses (Bass et al., 2009; Cavalier-Smith, 2002, 2003). The relatively large taxon sampling and gene sampling presented herein allow revisiting the fundamental relationships among rhizarian lineages. Representative species of most major lineages of Rhizaria were included, except for Haplosporidia and Vampyrellida. The splitting of Rhizaria into the phyla Cercozoa and Retaria as suggested by Cavalier-Smith (2003) is only partly confirmed. We found a good support for Retaria, but the phylum Cercozoa, composed of subphyla Filosa and Endomyxa, is not supported. The Filosa (called here Cercozoa) are monophyletic in our trees, but the Endomyxa, represented in our data by Plasmodiophorida and the *Gromia + Filoreta* clade, appears as a paraphyletic group. To fully confirm this hypothesis, two endomyxean clades for which multigene data are still missing (e.g. Haplosporidia and Vampyrellida) would need to be added to the analysis in the future.

Our study focusing on the phylogeny of Retaria clearly rejects its division into the subphyla Radiozoa and Foraminifera (Cavalier-Smith, 2003). While Foraminifera remained monophyletic in all analyses, Radiozoa are clearly paraphyletic. Multigene data are still lacking for the class Sticholonchea as well as for the Nassellaria traditionally placed within Polycystinea. However, we strongly doubt that adding these two groups will substantially change the topology of our trees making Radiozoa monophyletic. Therefore, we suggest that both Radiolaria and Radiozoa as valid taxonomic groups shall be abandoned.

Acknowledgments

The authors thank the two anonymous reviewers for their valuable comments and suggestions to improve the manuscript. We thank Fabien Burki, Silvia Restrepo and Juan Montoya for valuable comments, discussions and for encouraging this work. Elizabeth Alve for providing the means for collecting foraminiferans in the Oslofjord and sharing her knowledge on the local species. Thomas Cedhagen for providing the *G. turgida* samples. José Fahrni and Jackie Guiard for technical assistance, collecting samples and maintaining cultures. Kamran Shalchian-Tabrizi and Surendra Kumar for useful suggestions and help with analyses. The computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the Swiss Institute of Bioinformatics and Bioportal (http://www.bioportal.uio.no). Julie Poulain and Corinne Da Silva for sequencing of radiolarian species as part of the Genoscope Rhizarian genomics project (JP and CdV). The EU FP7 ASSEMBLE project "Phylogenomics of Foraminifera" for funding fieldwork. This work was supported by the Swiss National Science Foundation grant No. 31003A_140766 (RS, JP), the EU BiodivErsA project BioMarKs (Biodiversity of Marine euKaryotes, CdV, FN).

Table 1. Topological tests results of alternate phylogenetic hypothesis. Summary of tested topologies as constrained trees for the 36 and 109 genes matrices and the results obtained using CONSEL.

Phylogenetic hypothesis	SAR36		SAR109	
	AU	SH	AU	SH
Best ML tree	0.82	1	0.891	1
Monophyly of Acantharea + Polycystinea	**0.054**	0.863	0.559	0.996
((Foraminifera,Polycystinea),Acantharea)	**0.009**	0.448	N/A	N/A
((Foraminifera,Acantharea),Polycystinea)	N/A	N/A	**5.00E-40**	0.817
Paraphyly of Polycystinea, *Collozoum* sp. at the base	**0.001**	0.352	**8.00E-23**	0.55
Paraphyly of Polycystinea, *S. streptacantha* at the base	0.57	0.981	0.197	0.991

AU, approximately unbiased test; SH, Shimodaira-Hasegawa test (*p*-values). *p*-values below 0.05 are highlighted in bold.

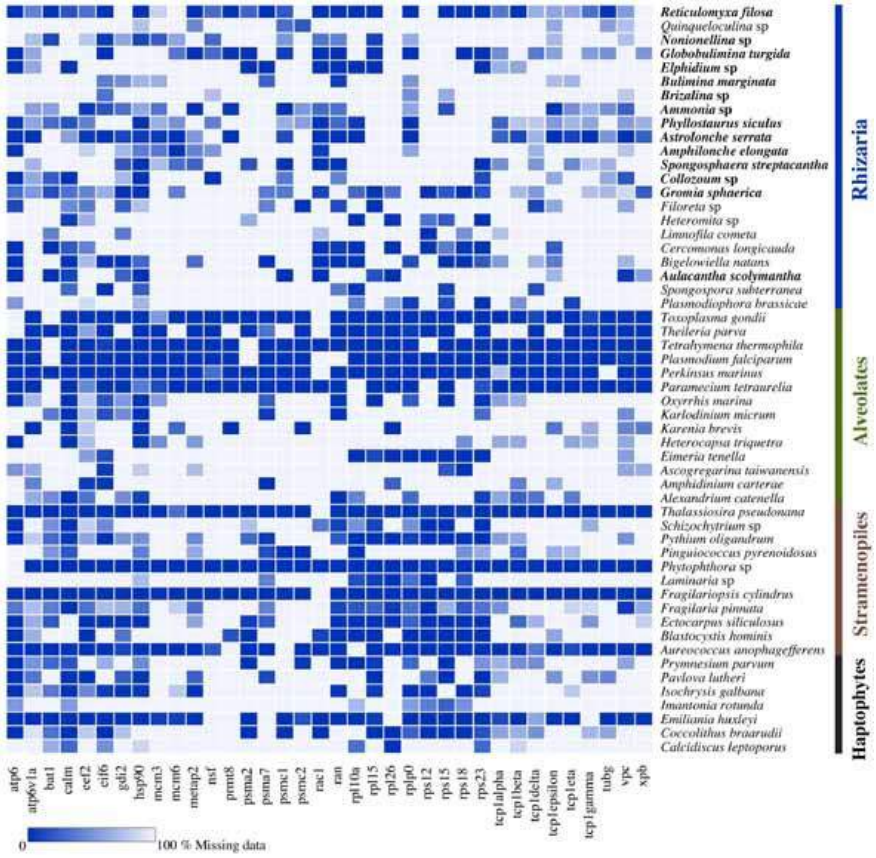


Figure 1. Heat map representing the percentage of missing data for the matrix containing 36 genes and 54 species (SAR36). Missing data is illustrated per gene (*x*-axis) and per species (*y*-axis) in a color gradient from solid red (0% missing data) to white (100% missing data). Taxa in bold represent species newly sequenced for the present study.

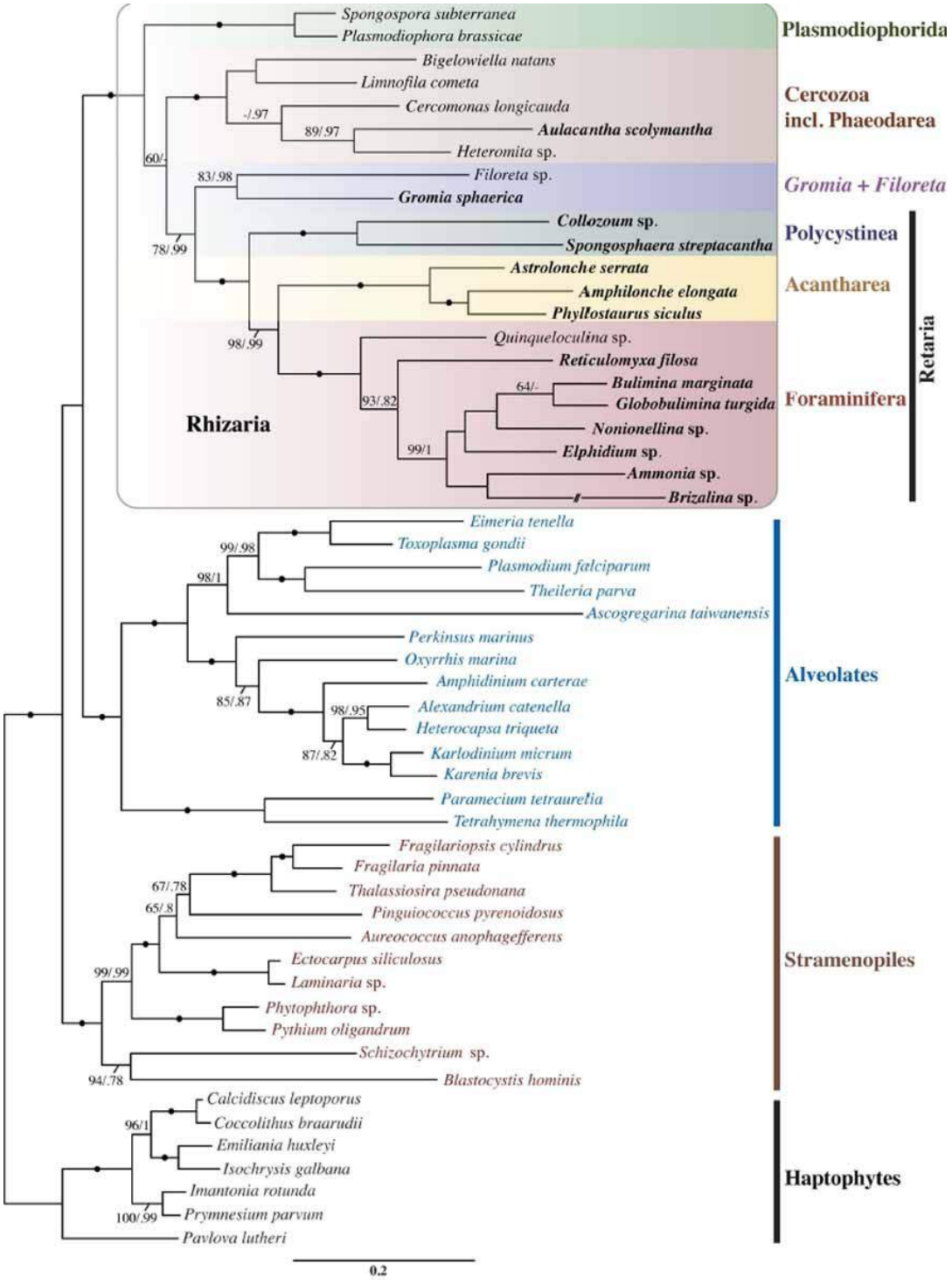


Figure 2. Phylogenetic relationships of 54 SAR (Stramenopiles, Alveolates and Rhizaria) and 7 outgroup species based on 36 genes. The tree was obtained as the highest scoring maximum likelihood tree using LG+Γ model and empirical amino acid frequencies. The numbers at nodes indicate the topological support estimated by bootstrap replicates and Bayesian consensus posterior probabilities of post-burnin bipartitions. Solid circles represent maximum support. Taxa in bold represent species newly sequenced for the present study.

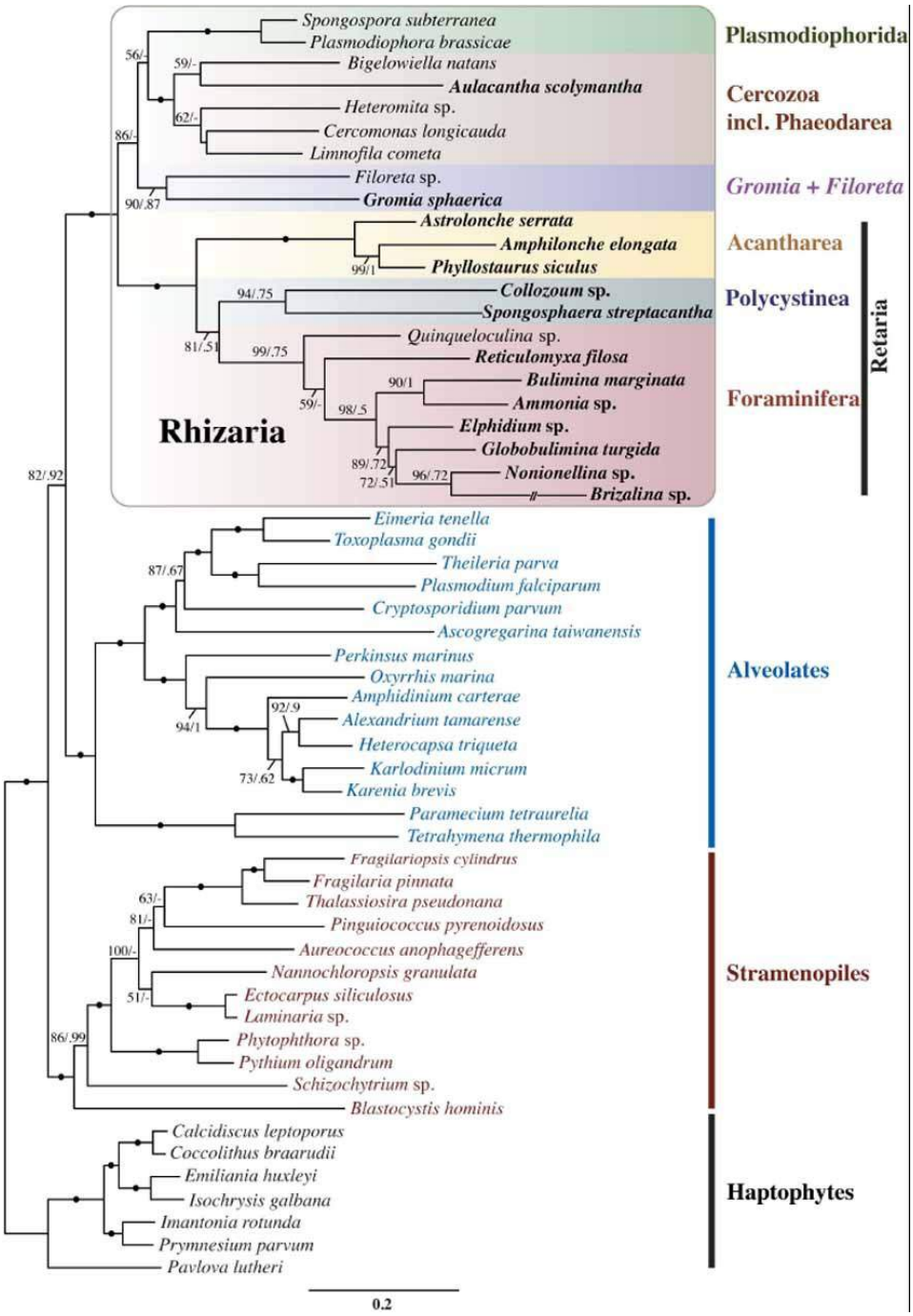


Figure S1. Phylogenetic relationships of 56 SAR and 7 outgroup species based on 109 genes. The tree was obtained as the highest scoring maximum likelihood tree using LG+Γ model and empirical amino acid frequencies. The numbers at nodes indicate the topological support estimated by bootstrap replicates and Bayesian consensus posterior probabilities of post-burnin bipartitions. Solid circles represent maximum support. Taxa in bold represent species newly sequenced for the present study.

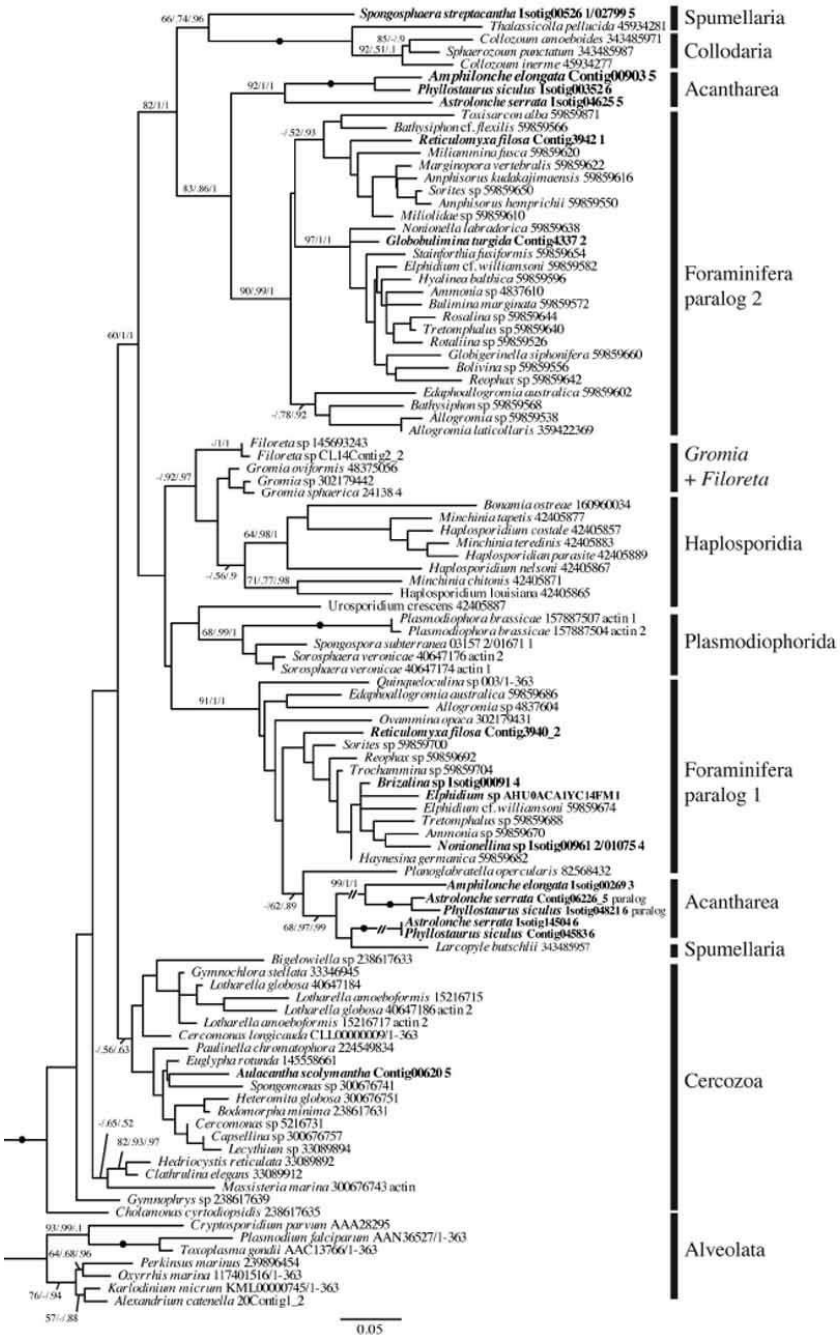


Figure S2. Actin phylogeny of 103 Rhizaria and 7 aveolates as outgroup species. The tree was obtained as the highest scoring maximum likelihood tree using WAG+Γ model and empirical amino acid frequencies. Numbers at nodes indicate the topological support for each node obtained by bootstrap replicates then by the Bayesian consensus posterior probabilities using PhyloBayes and third number using MrBayes. Only the deepest nodes and nodes with maximum support (solid circles) are shown. Taxa in bold represent species newly sequenced for the present study.

Table S1. List of collected species, origin and isolation sites of specimens.

Species	Origin	Coordinates	Depth/habitat	No. of cells
Reticulomyxa filosa	Bochum, Germany	N/A	Fresh water	plasmodium
Elphidium sp.	Biological station, Roscoff	48°43.63' N; 3°59.29' W	Intertidal zone	3000
Globobulimina turgida	Gullmar Fjord, Sweden	58°19.4' N; 11°32.7' E	117 m	400
Ammonia sp.	Camargue, France	43°33.75' N 4°6.79' E	Intertidal zone	100
Brizalina sp.	Oslofjord location Im4 , Norway	59°37.34' N; 10°37.74' E	202 m	100
Bulimina marginata	Oslofjord location Im4 , Norway	59°37.34' N; 10°37.74' E	202 m	75
Nonionellina sp.	Oslofjord location Im4 , Norway	59°37.34' N; 10°37.74' E	202 m	40
Astrolonche serrata	Villefranche Bay, France	43°41.0' N; 7°18.8' E	Water column	20
Phyllostaurus siculus	Villefranche Bay, France	43°41.0' N; 7°18.8' E	Water column	100
Amphilonche elongata	Gulf of Eilat, Israel	29°32.36' N; 34°57.85' E	Water column	150
Collozoum sp.	Villefranche Bay, France	43°41.0' N; 7°18.8' E	Water column	1 (colony)
Spongosphaera streptacantha	Villefranche Bay, France	43°41.0' N; 7°18.8' E	Water column	50
Aulacantha scolymantha	Villefranche Bay, France	43°41.0' N; 7°18.8' E	Water column	50
Gromia sphaerica	Little San Salvador Island, Bahamas	24°34.5'N; 76°00.1'W	720 m	0.5

Table S2. Summary of libraries sequenced, reads obtained and assembly details.

Species (EST library)	Sequencing system	Assembled reads	No. of contigs	Avr. Contig len.	N50	Total unitigs
Reticulomyxa filosa	Sanger	20,254	3,946	1,345	1,385	7,714
Elphidium sp.	Sanger	17,125	5,401	1,358	1,379	12,638
Globobulimina turgida	Sanger	18,444	4,634	1,340	1,374	13,677
Ammonia sp.	454	239,670	20,661	728	368	92,560
Brizalina sp.	454	98,898	3,739	671	361	62,902
Bulimina marginata	454	296,116	20,830	597	342	150,617
Nonionellina sp.	454	103,664	5,184	647	341	69,255
Astrolonche serrata	454	337,982	14,733	845	387	55,656
Phyllostaurus siculus	454	256,144	8,918	726	420	44,376
Amphilonche elongata	454	193,766	3,354	830	415	25,987
Collozoum sp.	454	213,393	5,211	643	317	83,657
Spongosphaera streptacantha	454	217,395	3,359	942	325	57,449
Aulacantha scolymantha	454	193,487	3,480	820	401	65,137
Gromia sphaerica	454	213,066	17,770	676	368	60,974

References

Bass, D., Chao, E.E.Y., Nikolaev, S., Yabuki, A., Ishida, K.-i., Berney, C.d., Pakzad, U., Wylezich, C., Cavalier-Smith, T., 2009. Phylogeny of Novel Naked Filose and Reticulose Cercozoa: Granofilosea cl. n. and Proteomyxidea Revised. Protist 160, 75-109.

Burki, F., Kudryavtsev, A., Matz, M., Aglyamova, G., Bulman, S., Fiers, M., Keeling, P., Pawlowski, J., 2010. Evolution of Rhizaria: new insights from phylogenomic analysis of uncultivated protists. BMC Evolutionary Biology 10, 377.

Burki, F., Nikolaev, S.I., Bolivar, I., Guiard, J., Pawlowski, J., 2006. Analysis of expressed sequence tags from a naked foraminiferan Reticulomyxa filosa. Genome 49, 882-887.

Burki, F., Okamoto, N., Pombert, J.F., Keeling, P.J., 2012. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. Proceedings. Biological sciences / The Royal Society.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular biology and evolution 17, 540-552.

Cavalier-Smith, T., 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. International journal of systematic and evolutionary microbiology 52, 297-354.

Cavalier-Smith, T., 2003. Protist phylogeny and the high-level classification of Protozoa. Eur J Protistol 39, 338-348.

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Research 8, 186-194.

Ewing, B., Hillier, L., Wendl, M.C., Green, P., 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Research 8, 175-185.

Flakowski, J., Bolivar, I., Fahrni, J., Pawlowski, J., 2006. Tempo and mode of spliceosomal intron evolution in actin of foraminifera. Journal of molecular evolution 63, 30-41.

Haeckel, E.H.P.A., 1887. Reports on Radiolaria Collected by H.M.S. Challenger During the Years 1873-1876. Her Majesty's Stationary Office, London.

Haeckel, E.H.P.A., 1904. Kunstformen der Natur. Verlag des Bibliographischen Instituts, Leipzig und Wien :.

Ishitani, Y., Ishikawa, S.A., Inagaki, Y., Tsuchiya, M., Takahashi, K., Takishita, K., 2011. Multigene phylogenetic analyses including diverse radiolarian species support the "Retaria" hypothesis - The sister relationship of Radiolaria and Foraminifera. Marine Micropaleontology 81, 32-42.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT version 5.25: multiple sequence alignment program. Nucl Acids Res 30, 3059-3066.

Krabberod, A.K., Brate, J., Dolven, J.K., Ose, R.F., Klaveness, D., Kristensen, T., Bjorklund, K.R., Shalchian-Tabrizi, K., 2011. Radiolaria divided into Polycystina and Spasmaria in combined 18S and 28S rDNA phylogeny. PLoS ONE 6, e23526.

Kunitomo, Y., Sarashina, I., Iijima, M., Endo, K., Sashida, K., 2006. Molecular phylogeny of acantharian and polycystine radiolarians based on ribosomal DNA sequences, and some comparisons with data from the fossil record. European Journal of Protistology 42, 143-153.

Lartillot, N., Lepage, T., Blanquart, S., 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286-2288.

Lecroq, B., Gooday, A., Cedhagen, T., Sabbatini, A., Pawlowski, J., 2009. Molecular analyses reveal high levels of eukaryotic richness associated with enigmatic deep-sea protists (Komokiacea). Marine Biodiversity 39, 45-55.

Moreira, D., von der Heyden, S., Bass, D., Lopez-Garcia, P., Chao, E., Cavalier-Smith, T., 2007. Global eukaryote phylogeny: Combined small- and large-subunit ribosomal DNA trees support monophyly of Rhizaria, Retaria and Excavata. Mol Phylogenet Evol 44, 255-266.

Nikolaev, S.I., Berney, C., Fahrni, J.F., Bolivar, I., Polet, S., Mylnikov, A.P., Aleshin, V.V., Petrov, N.B., Pawlowski, J., 2004. The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. Proc Natl Acad Sci U S A 101, 8066-8071.

Pawlowski, J., Bolivar, I., Fahrni, J.F., Cavalier-Smith, T., Gouy, M., 1996. Early origin of foraminifera suggested by SSU rRNA gene sequences. Molecular biology and evolution 13, 445-450.

Pawlowski, J., Burki, F., 2009. Untangling the phylogeny of amoeboid protists. J Eukaryot Microbiol 56, 16-25.

Polet, S., Berney, C., Fahrni, J., Pawlowski, J., 2004. Small-subunit ribosomal RNA gene sequences of Phaeodarea challenge the monophyly of Haeckel's Radiolaria. Protist 155, 53-63.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Roure, B., Rodriguez-Ezpeleta, N., Philippe, H., 2007. SCaFoS: a tool for selection, concatenation and fusion of sequences for phylogenomics. BMC Evol Biol 7 Suppl 1, S2.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Systematic biology 51, 492-508.

Shimodaira, H., Hasegawa, M., 1999. Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. Molecular Biology and Evolution 16, 1114.

Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246-1247.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688-2690.

Takahashi, O., Yuasa, T., Honda, D., Mayama, S., 2004. Molecular phylogeny of solitary shell-bearing Polycystinea (Radiolaria). Revue de Micropaléontologie 47, 111-118.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28, 2731-2739.

Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Molecular Biology and Evolution 18, 691-699.

Yuasa, T., Takahashi, O., Dolven, J.K., Mayama, S., Matsuoka, A., Honda, D., Bjørklund, K.R., 2006. Phylogenetic position of the small solitary phaeodarians (Radiolaria) based on 18S rDNA sequences by single cell PCR analysis. Marine Micropaleontology 59, 104-114.

Yuasa, T., Takahashi, O., Honda, D., Mayama, S., 2005. Phylogenetic analyses of the polycystine Radiolaria based on the 18s rDNA sequences of the Spumellarida and the Nassellarida. European Journal of Protistology 41, 287-298.

The Protist Ribosomal Reference database (PR^2): a catalog of unicellular eukaryote Small SubUnit rRNA sequences with curated taxonomy

Guillou Laure [1,2,*], Bachar Dipankar[3,4], Audic Stéphane[1,2], Bass David[5], Berney Cédric[5], Bittner Lucie[1,2,&], Boutte Christophe[1,2], Burgaud Gaetan[6], de Vargas Colomban[1,2], Decelle Johan[1,2], del Campo Javier[7], Dolan John R.[8], Dunthorn Micah[9], Edvardsen Bente[10], Holzmann Maria[11], Kooistra Wiebe H.C.F. [12], Lara Enrique[13], Le Bescot Noan[1,2], Logares Ramiro[7], Mahé Frédéric[1,2], Massana Ramon[7], Montresor Marina[12], Morard Raphael[1,2], Not Fabrice[1,2], Pawlowski Jan[11], Probert Ian[14,15], Sauvadet Anne-Laure[1,2,#], Siano Raffaele[16], Stoeck Thorsten[9], Vaulot Daniel[1,2], Zimmermann Pascal[17] and Christen Richard[3,4,*].

[1] CNRS, UMR 7144, Adaptation et Diversité en Milieu Marin, Station Biologique de Roscoff, 29682 Roscoff, France
[2] UPMC Univ. Paris 06, UMR 7144, Station Biologique de Roscoff, 29682 Roscoff, France
[3] CNRS, UMR 7138, Systématique Adaptation Evolution. Université de Nice-Sophia Antipolis. Parc Valrose. BP71. F06108 Nice cedex 02, France
[4] Université de Nice-Sophia Antipolis, UMR 7138, Systématique Adaptation Evolution. Parc Valrose. BP71. F06108 Nice cedex 02, France
[5] Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK
[6] Laboratoire Universitaire de Biodiversité et Ecologie Microbienne (EA3882), ESMISAB, Technopôle Brest-Iroise, 29280 Plouzané, France
[7] Department of Marine Biology and Oceanography, Institut de Ciències del Mar (CSIC), Barcelona, Catalonia, Spain
[8] UPMC Univ. Paris 06 et CNRS, UMR7093, Laboratoire d'Océanographie de Villefranche, Marine Microbial Ecology, Station Zoologique, B.P.28, 06230 Villefranche-sur-Mer, France
[9] Department of Ecology, University of Kaiserslautern, 67663 Kaiserslautern, Germany
[10] University of Oslo, Department of Biology, Marine Biology, NO-0316 Oslo,Norway
[11] Department of Genetics and Evolution, University of Geneva, Switzerland
[12] SZN, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy
[13] University of Neuchâtel, Institute of Biology, Laboratory of Soil Biology, rue Emile Argand 11, CP 158. CH-2009, Neuchâtel, Switzerland
[14] UPMC Univ. Paris 06, FR2424, Station Biologique de Roscoff, 29682 Roscoff, France
[15] CNRS, FR2424, Station Biologique de Roscoff, 29682 Roscoff, France
[16] Ifremer, Centre de Brest, DYNECO/Pelagos BP 70 29280 Plouzané, France
[17] Point Compétence Informatique, Rue Jean-Baptiste Say, 56850 Caudan, France

*To whom correspondence should be addressed. Laure Guillou; Tel: +33 2 98 29 23 79; Fax: +33 2 98 29 23 24; Email:lguillou@sb-roscoff.fr. Correspondence may also be addressed to Richard Christen; Tel: +33 601 830 144; Email: christen@unice.fr
[&] Present address: Department of Ecology, University of Kaiserslautern, 67663 Kaiserslautern, Germany
[#] Present address: School of Biosciences, Cardiff University, Main Building, Cardiff, CF10 3AT., UK

Accepted in Nucleic Acids Research

Abstract

The interrogation of genetic markers in environmental meta-barcoding studies is currently seriously hindered by the lack of taxonomically-curated reference datasets for the targeted genes. The Protist Ribosomal Reference database (PR^2, http://ssu-rrna.org/) provides a unique access to eukaryotic small sub-unit (SSU) ribosomal RNA and DNA sequences, with curated taxonomy. The database mainly consists of nuclear-encoded protistan sequences. However, metazoans, land plants, macrosporic fungi and eukaryotic organelles (mitochondrion, plastid and others) are also included because they are useful for the analysis of High-Troughput Sequencing datasets. Introns as well as putative chimeric sequences have been also carefully checked. Taxonomic assignation of sequences consists of eight unique taxonomic fields. In total, 136,866sequences are nuclear encoded, ,45,708 (36,501 mitochondrial and 9,657 chloroplastic) are from organelles, the remaining being putative chimeric sequences. The website allows the users to download sequences from the entire and partial databases (including representative sequences after clustering at a given level of similarity). Different web-tools also allow searches by sequence similarity. The presence of both rRNA and rDNA sequences, taking into account introns (crucial for eukaryotic sequences), a normalized eight terms ranked-taxonomy and updates of new GenBank releases were made possible by a long term collaboration between experts in taxonomy and computer scientists.

Introduction

The modern definition of the term "protist" refers to unicellular eukaryotes that are either free-living or parasitic, sometimes forming colonies, but without clear differentiation into tissues. This includes all eukaryotes other than land plants (and macro-algae), animals, and fungi with differentiated tissues. Protists are notoriously paraphyletic and include a wide range of microorganisms employing a huge variety of reproductive, nutritional and life history strategies. Nevertheless, the term protist has pragmatic uses and has recently gained in popularity. Large-scale analysis of protistan diversity is complicated by their heterogeneity, which reflects their extremely broad distribution and implication in multiple ecological and functional processes. This difficulty is exacerbated by the facts that: (1) species delineation is often obscure due to lack of clear morphological criteria and paucity of knowledge concerning processes of sexual recombination; (2) the taxonomy of protists has been radically modified in recent decades in light of new phylogenetic data; (3) a large proportion of protists are probably still not cultivable or yet unknown. Molecular bar-coding using SSU rRNA (Small Sub-Unit Ribosomal) gene sequences consequently have become extremely popular among protistologists. Environmental bar-coding has unveiled an extensive genetic diversity of protists in a wide range of ecosystems (1,2), including lineages only known by their genetic signatures (orphan environmental sequences). Recently, the use of NGS (Next Generation Sequencing) technologies targeting selected domains of the SSU rRNA gene has permitted ecological studies of complex assemblages at ever increasing scales (3-7). However, interpretation of such data is currently seriously hindered by the lack of taxonomically-curated reference datasets. Unassigned and incorrectly assigned sequences are accumulating at an increasing and alarming rate in public databases, to the extent that in early 2012, almost 20% of submitted SSU rRNA eukaryotic gene sequences had no or a very poor taxonomic assignation (see the web-site for more details). Undetected chimeric sequences (8), as well as the presence of introns in gene sequences (9), are also problematic.

In order to facilitate and increase the efficiency and accuracy of NGS datasets analyses, we here present the first comprehensive curated database that places eukaryotic SSU rRNA gene sequences within a coherent ranked taxonomic framework covering eukaryotic diversity. Every sequence was quality-checked and annotated using a multi-level taxonomic assignation. As a lot of protists are still only known by their environmental sequences, cluster names were retained when the formal taxonomy was missing (such as Syndiniales (10), and Marine STramenopiles, MAST (11)). Although curated in less detail, sequences from Metazoa, land plants and macrosporic fungi, as well as eukaryotic organelles (mitochondria, plastids, etc.), are also included in the database for their ecological interests. As example, protists may live in close association with metazoan (commensalisms, symbioses,..), and very small metazoan exists, inhabiting similar ecological niches. For example, copepods and polychaetes, as well as benthic animal larvae coexist with planktonic protists in aquatic systems. They may have also a great interest in ecological studies (as predators for example), even for protistologists. Even if this database is dedicated to protists, such outgroup sequences are of high relevance for extracting these groups in further analyses of NGS datasets when “universal” eukaryotic primers are used for PCR amplifications. Metazoan sequences in PR2 allow not identifying them wrongly as new deep lineages of protists.

Material and Methods

The construction of this database started more than 10 years ago, and our procedure has been optimized over time (recent history detailed at http://ssu-rrna.org/method.html for more details). Here, we briefly describe the present general architecture of the database.

Entries containing at least one partial SSU rRNA gene sequence of eukaryotic origin are retrieved from three public databases using keywords. Our last update retrieved 484.657, 496.462 and 123 such entries from GenBank, EMBL and WGS-EMBL, respectively. An INSDC (http://www.insdc.org/) entry as defined by its accession number in public databases may contain several rRNA gene sequences, for example in long genomic fragments containing several partial or complete ribosomal operons. In order to allow such duplicated sequences within a single entry, each sequence was given a unique identifier, acc.p1.p2, where acc is the accession number of the entry containing the sequence, and p1 and p2 are the first and last positions of the sub-sequence within the complete sequence.

A majority of extracted sequences were shorter than 100 nucleotides or around 500 nucleotides in length (63% of retrieved sequences) , likely resulting from the recent integration of short environmental sequences derived from clone libraries. Only sequences longer than 799 nt were considered.

The first step was the identification of sequences originating from organelles. A reference database of SSU-rRNA gene sequences from chloroplasts and mitochondria was constructed using entire genomes or genomic fragments that contained a SSU-rRNA gene sequence and a protein-coding gene specific either of mitochondria or of chloroplasts. For derived-organelle sequences such as apicoplasts, hydrogenosomes and nucleomorphs, databases were manually built, using information found in scientific publications. These databases were used to determine by sequence similarity the origin of every sequence in the database.These sequences were assigned to a reduced taxonomic framework, including their location (such as: |Organelle|chloro-SSU| or |Organelle|mito-SSU|). These sequences are not more detailed in the database.

Introns were found to be a major problem in eukaryotic rRNA sequences compared to prokaryotic sequences (1,536 sequences with intron(s) described, 10,644 sequences with introns found by computation). A dedicated C++ algorithm was developed in order to identify the presence of introns in the remaining sequences (9). When detected, sequences with and without the intron(s) were generated (rRNA and rDNA sequences).

Sequences in the PR^2 database are assigned an identifier in the form accession.p1.p2_X, where accession is the accession number of an entry, p1 and p2 are the positions of this sequence in a larger genomic entry, and X corresponding to introns treatment of the sequence (X=G: genomic sequence containing a described intron (rDNA); X=R: the previous genomic rRNA sequence, without the intron(s); X=U: no intron described, but intron(s) may be present; X=UC: introns were detected *in silico* and removed from the sequence (putative rRNA)).

Taxonomy of nuclear-encoded sequences

Since all SSU-rRNA genes are orthologs, a global phylogeny can be built and essential past speciation events can be evidenced. This property is essential to build a ranked taxonomy. For example at rank-1, there is a world-wide agreement to recognize three clades, Bacteria, Archaea and Eukaryota. We chose to additionally use “Organelle” as rank-1. Organelles have an eukaryote origin when they are nucleomorphs and a bacterial origin when they are mitochondrion and plastid. Because evolution of organelles and their hosts differ over time, their taxonomy are quite different too. In addition, scientists working on diversity are more interested in the identification of the cells that bear such organelles. Our choice was thus to

allow their easy identification (and filtering out) during the first step of an analysis, targeting them as "Organelle" at rank-1.

Nomenclature and terms of the following ranks mainly follows the classification of eukaryotes proposed by Adl et al. (12). Thus, the second rank describes each eukaryotic "Super-Group" or Phylum (both terms are in use in different communities): Alveolata, Amoebozoa, Apusozoa, Archaeplastida, Excavata, Opisthonkonta, Rhizaria, or stramenopiles. The taxonomic descriptions are structured by the use of eight ranks, and following ranks mainly correspond to the division, class, order, family, genus, and species.

The terms used for each rank are non-ambiguous (a term cannot be found in two different clades), contain no space (that may pose problems to computers), and whenever possible retained if monophyletic. When monophyly could not be insured, the term of rank above was used, appended with suffix _X (suffix X if rank above was already _X). As the same species name frequently occurs in different genera, the species name is composed of the genus and species, using "+" as a separator (e.g. genus=Diderma, species=Diderma+niveum). Genus and species names from public databases are stored in separate fields for comparison.

For protists and unicellular fungi, a taxonomy was proposed by the group of experts, authoring this paper. For multicellular fungi, plants and metazoans, the taxonomy was built mostly using the taxonomy assigned in NCBI's GenBank database entries. We first built a core reference database containing 23,116 manually analyzed sequences representative of eukaryotic diversity. These analyses included reading published papers as well as phylogenetic analyses done by the authors of this paper when necessary. This core reference database was subsequently used to automatically annotate the remaining sequences using different methods.

We are aware that for some clades such as Metazoa, Plants and Fungi, our eight terms taxonomy is probably not as precise as it should be. Barcoding of Metazoa and Plants using SSU-rRNA sequences is not often used (normally only to complement ITS sequences). We will therefore try in a next release to propose an extended, still ranked and unified, taxonomy for Fungi.

An outcrop of PR^2 is the web-based tool KeyDNAtools (http://keydnatools.com/). It uses 159,982 specific short (15 nt) oligonucleotide sequences (named keys) generated from the core reference database. Each key is a signature present in sequences of a given clade but not in those of other clades. Besides providing a very fast taxonomic identification, it also allows for detecting putative chimeric sequences, as when different identifications are obtained from the 5' and 3' ends of sequences.

Specific new computer programs mostly in C, C++ and Python have been developed. First, a new parallel distributed computing Needleman-Wunsch based C program allowing to compute pair-wise distances not taking into account terminal gaps (partially overlapping sequences) and long internal gaps (introns). This was coupled to a newly rewritten C average linkage clustering program. Second, a new parallel distributed computing Needleman-Wunsch based C++/Python program allowing to assign a consensus taxonomy to new sequences by comparison to a reference database (Crunch_Assign).
When a conflict between taxonomies assigned using the different methods was found, it was manually solved. In the end, each nuclear encoded sequence is assigned an identifier in the form of this example:

>AY827845.1.1765_U|Eukaryota|Apusozoa|Hilomonadea|Planomonadida|Planomonadidae|Planomonadidae_Group-1|Ancyromonas|Ancyromonas+sigmoides

Results

In total, we found 1136,866 nuclear encoded sequences, 5 pseudo-genes (FJ854546, FJ854545, D14632, AF310844, AJ404858, not included in PR^2) and 34 sequences we could only assign as putative rRNA sequences (HM538255, GU385678, AB275106, AJ628837, AY180011, CP000499, CP000499, AY256215, EU402432, AB017015, GQ330639, GU820811, JF488788, AF239231, DQ423737, DQ104596, AY835700, DQ423728, EU545797, GU072272, GU072526, GQ247249, HM174255, DQ104594, EU174762, FN598473, EU726200, EF695080, GQ483783, GQ462590, EU173354, EF567390, EF695215, HQ871039, not included in PR^2). Manual analyses of some of them allowed concluding for the presence artefactual sequence internal or at the 5' or 3' end. Among nuclear-encoded sequences, we detected 1,756 putative chimeric sequences, either using the KeyDNAtools, and/or by manual inspection (listed on the web site). For example, sequence EF023694.1.1975_U is a chimera between parent sequences of Opisthokonta, Amoebozoa, and Rhizaria in position 179-471, 623-1264, and 1536-1925, respectively. Other "18S" sequences are nucleomorphs (262 sequences). 9,657 sequences have a chloroplastic origin, 33,051 are from mitochondria, 6 from hydrogenosomes (AJ237907, AJ237908, AJ871215, AJ871217, AJ871267, Y16670), and 26 from apicoplasts (U87145, AB471801, AB471802, AB471803, AB471804, AB471805, AB471806, AB471807, AB471808, AB471809, AB471810, AB471811, AB471812, AB649417, AB649418, AB649419, AB649420, AB649421, AB649422, AB649423, AB649424, HQ110105, JQ437257, JQ437258, JQ437259, U28056).
Within nuclear-encoded sequences, 54 data entries remained unassigned at the Super-Group level (Table 1), meaning that they could not be assigned to any specific taxon group within the domain Eukaryota (Eukaryota_X). The Super-Group "Eukaryota_Mikro" was created for sequences HM563060, AF477623 and HM563061, for which no consensus has been reached for their affiliation, although Haplosporidiidae has been suggested (13). BLAST analyses conducted at NCBI against nr or at DDBJ against all, showed extremely weak sequence similarity with sequences of fungi. Using our global similarity tool (Crunch_Assign) showed no other sequence similar at 80% or more along the entire sequence. These results conducted to the creation of this new Super-Group (rank2). For unassigned nuclear-encoded sequences (Eukaryota_X), either no other similar sequence was found or similar sequences were detected but also annotated by us as Eukaryota_X. A BLAST on NCBI nr (excluding environmental sequences) and at DDBJ (all) revealed that a large number of them probably contained undescribed introns. Therefore these sequences probably require a manual curation, but again highlight the importance of intron identification in eukaryotic sequences.

For lower taxonomic ranks, there were primarily two types of cases resulting in a failure to assign a taxonomic identity:

- No agreement between experts to resolve at a given rank. For example the genus (rank 7) is assigned, the order (rank 5) is assigned but a family (rank 6) has not yet been described, or this rank is in fact polyphyletic, with no proper descriptions of the different families.

- A given sequence is similar at the family level with several sequences from different families; however they agree at the order level.

In such cases, this sequence was assigned as ...|Order| Order_X[Genus|Genus+species. If a genus was not described (i.e. uncultured), the taxonomy becomes: ...|Order| Order_X[Order_XX|Order_XX+sp.

More than 74,000 sequences (54% of total number of sequences in the PR2 database) belong to Opisthonkonta (Fig. 1). Alveolata and Archaeplastida are second in abundances (15 and 12%, respectively). Stramenopiles and Rhizaria represent 7.2 and 5.6 %, respectively. Others SuperGroups represent less than 2.2%. Only 29,4% are complete or nearly complete. In total, 63.7% of sequences include the V4 region and only 12.1% and 11.7% include the V9 region as recognized by primers Biomarks and Wamps (see the legend of Figure 1), respectively. Apusozoa, Hacrobia, Excavata and Opisthokonta have less than 10% of their sequences that include the V9 region. V9 region of Amoebozoa and Archaeplastida are better represented (34% and 25%, respectively using the Biomarks primers). DOWNLOADS

We provide several different ways of downloading the database or part of it (see more explanations at http://ssu-rrna.org/downloads_eukaryotic_main_page.html).

- The entire database, or sequences of a specific clade can be downloaded using a taxonomy browser under fasta format, with sequence identifiers as described above. Putative chimera have been removed.

- The entire database, or sequences of major groups can be downloaded under fasta format, with only the short unique identifier. The corresponding taxonomy is then downloaded as a tabulated file. This fasta format is appropriate to use in tools that do not allow for long sequence identifiers. They are also easier to use in large computations as they spare the memory required. Finally they are easier to use in pipelines or web sites (see below).

- The entire database, taxonomies and sequences under tabulated format, for easy import in relational databases.

- The entire database, or sequences of a specific clade under fasta format, with sequence identifiers as described above, but after a clustering by sequence similarity (98, 96, 92 %) and choosing only the longest sequence as representative of the cluster.

- Phylogenetic trees are available for the main groups. They were built using pair-wise distance computations (not taking introns as differences as explained above) and FastMe (14).

- Finally we provide an “arb” filter that allows to easily import a fasta file (with taxonomy in the identifier) into an arb database, separating sequences and taxonomy as required.

- *In silico* extracted domains corresponding to regions widely used in published articles and corresponding to several couples of primers.

Searching the database

We provide the following additional kinds of tools:

- A search by keywords, allowing to search according to taxonomy, accession number and PMID (PubMed ID: retrieval of sequences described in a given publication). Retrieven sequences can be

filtered according to length, quality and when containing the variable V4 of V9 domains (often used in conjunction with deep sequencing).

- A search by "sequence signature", with a link to the KeyDNAtools web site (http://keydnatools.com/). This tool provides very fast results even for files containing many sequences. It also allows for detection of putative chimera as explained above.
- A BLAST search against the database, as usually found on most sites.
- A search (Crunch_Assign) using our modified global (Needleman-Wunsch based) algorithm that returns the most similar hits based on the entire alignment of the sequences, and not based on a good local alignment (HSP for High Scoring Pair, in BLAST). As a result, the percentage of similarity computed is more in agreement with what would be found using a Multiple Sequence Alignment (Clustal (15), Muscle (16), MAFFT (17), …) before computing distances. It allows accounting or not for introns as described above.
- A search of one or two primer motifs in sequences, returning every sequence that contains the primer(s) with IUPAC encoding allowed and also the possibility of mismatches between primer and sequence (a C program).
- In silico extracted domains corresponding to regions widely used in published articles and corresponding to several couples of primers.

Both BLAST and Crunch_Assign similarity searches are coupled to BLAST2Tree or Crunch_Assign2Tree that use our Scriptree software (18). Similarity search results can simply be copied and then pasted in the "2Tree section"; a phylogenetic tree is built and displayed on the fly, with taxonomic assignations (as chosen by the user) displayed in regard of each leaf. This section also allows downloading the sequences that have been pasted as well as the taxonomy as a tabulated file (19).

Conclusion and perspectives

There are presently three databases, SILVA (20), RDP (21) and Green genes (22), offering a curated taxonomy for prokaryotic SSU rRNA sequences. Only Silva additionally provides reference sequences for SSU-rRNA sequences of eukaryotic origin, curated for sequence quality but using the NCBI taxonomy (although recently a "Silva" taxonomy is now proposed). Because our sequence identifier i.e. accession.p1.p2 is similar to that used by Silva, both databases can be easily compared.

Based on the last release 111, 1,518 of the 71,787 eukaryotic Silva reference sequences are not present in the PR^2 database. Manual checks showed that these sequences correspond to sequences extracted from entries in which no annotation allowed to identify the presence of a SSU-rRNA sequence, annotated as mRNA or annotated as prokaryotes. 670 sequences identified as mitochondria were not in PR^2; none of the Silva chloroplast sequences was absent from PR^2. Missing sequences will be soon analyzed and incorporated in PR^2.On the other hand, 53,735 / 7,774 nuclear, 31,492 / 29,763 mitochondrial, 462/18 chloroplastic and 133/80 other organelle sequences present in PR^2 were not in Silva reference sequences and Silva entire database, respectively. This can be largely explained by the use of drastic filtering steps used by SILVA both in minimal length and sequence quality. However, because we are also users of such databases to analyze NGS datasets, we detected two major reasons not to use too drastic quality filtering. First, representatives of novel environmental clades are often found within clone libraries with length of

less than 1,000 nt. Also, use of extreme quality filters may remove important sequences representatives of environmental groups, too short and/or having poor quality at one of the end of a sequence (one-step Sanger sequencing without enough noise treatment for example). In PR^2, sequence quality was indirectly inferred by the quality of the taxonomic assignation because bad-quality sequences became poorly assigned. Again, as sequence identifiers are similar between both databases, sequences can be easily compared between both databases. The PR^2 database possesses several valuable complementary tools or databases lacking in other databases.

A ranked taxonomy.

As for the PR^2 database, SILVA taxonomy for eukaryotes now offers a taxonomy based on the structure proposed by Adl et al. (12). However, contrarily to SILVA, we proposed a normalized eight terms ranked taxonomy for every sequence in the database. We proceeded to this "normalisation" from our experience in dealing with very large datasets using automated pipelines, and a depth of sequencing that revealed organisms spanning the entire spectrum of known living organisms. When considering the NCBI taxonomy for example, two sequences of Perciformes were found described using 22 ranks (AY263842 and EF470892 for Perciformes) while another Perciforme (AF112595) was described using only 15 ranks, and 10,360 sequences of Perciformes had between 16 and 21 ranks. Numerous examples exist for protists. A very good example is for the genus *Carpediomonas*. NCBI classify this genus within Eukaryota (rank1), Fornicata (rank2), Carpediomonas (rank3). However, sequence AY117416 (*Carpediemonas membranifera*, 23) has no rank-2 taxonomy in its entry. As a result it becomes extremely difficult using a computer and the lists of terms provided by a non-ranked taxonomy to identify for two different sequences, which members of the two lists indeed correspond to the same rank. This is the problem solved by our ranked-taxonomy, thanks to a worldwide list of taxonomic experts.

As an example, taxonomy of sequence AY117416 becomes Eukaryota|Excavata|Metamonada|Fornicata|Fornicata_Group-2|Carpediomonas-like|Carpediemonas|Carpediemonas+membranifera in PR^2. In SILVA, this sequence is linked to a 7 terms taxonomy, but taxonomy is seemingly not ranked and unified.

When occurring, missing ranks are automatically replaced in PR^2 (labeled as clade-i_X, where clade-i is the term for the next higher rank). This strategy allows rapidly inferring the taxonomy at the most probable higher rank and provides a rapid method for screening putative novel lineages at each taxonomic level.

Introns

Most SSU rRNA databases and biodiversity analyses of prokaryotes understandably neglect introns. Although found even in *Escherichia coli* (24,25), introns are quite rare in Bacteria and not very abundant in Archaea. Even when present, they have not yet been to our knowledge described in rRNA gene sequences. However, in Eukaryota, introns can be relatively abundant in rRNA gene sequences at least in some groups (9). This led us to incorporate in our database both the rRNA and the rDNA sequences. As most NGS (or clone library) analyses of the biodiversity are dealing with PCR amplification of extracted gDNA, introns may represent a large part of the variability observed. Having genomic sequences, in addition to the rRNA transcript, in the database is important, not only for searching by similarity but also for the *in silico* estimation of expected amplicon lengths.

Organelles

Organelles are often poorly treated in reference databases. For hydrogenosomes (AJ237907, AJ871215, AJ871217, AJ871267, Y16670), only sequence AJ871217 can be found in Silva labeled as "Unclassified". For GreenGenes, sequences were not found when searching by accession number. At RDP, the classifier resulted in every case into "unclassified_Bacteria". For the 26 apicoplast sequences, none was found in Silva reference sequences or in the "ssu-accession-parc.acs", release 111 (3,186,762 accession numbers). Even for better-known organelles, taxonomic assignation is not really better. For example, sequence AB000109 mitochondrion of *Dictyostelium discoideum* is labeled as "Unclassified" in Silva. Chloroplasts are generally well identified in Silva. However, Among the chloroplastic sequences detected in this study, 263 were found in Silva reference sequences as chloroplasts. Our approach to build independent databases for these organelles allowed us to probably reach a more precise taxonomic affiliation of organelles. Having such prokaryotic organelles in our database is essential with NGS datasets of both Prokaryotes and Eukaryotes because the use of "Bacteria" or "Eukaryota" specific primers resulted in some cases in a significant proportion of amplicons that are in fact of Organelle origin (3-7). Even if Organelle sequences are simply discarded from the final analysis, this database avoids identifying these sequences as some new deep lineages.

Chimeric sequences

Chimeric sequences are PCR-generated hybrid products between multiple parent sequences that can be falsely interpreted as novel organisms, thus inflating apparent diversity (8, 26). The two algorithms most widely used for 16S chimera detection are Pintail (27), included in RDP and SILVA databases, and Bellerophon (28) included in GreenGenes. In all cases, chimera are detected by comparing independent regions of a sequence alignment. The KeyDNAtools does not require the prior alignment of sequences, and it is particularly efficient to detect complex chimera having more than two parent sequences, or between two closely related parents. This tool can be used in concert with other detection methods. Our database, which has been screened for putative chimera, offers two possibilities of download: either including or excluding putative chimeric sequences.

Similarity searches

BLAST is a widely used tool that finds regions of local similarity between sequences. However, such search based on a good local HSP could lead to very bad results. We thus developed two independent methods of assignation. The first one, the Crunch_Assign software is using a Needleman-Wunsch algorithm. It is also faster than BLAST and returns a score computed on the entire alignment. Because we are working on Eukaryotes, we also included the possibility of ignoring putative introns (to our knowledge, this possibility is not included in any other software). The second one, the KeyDNAtools is also very fast and offers additionally chimera detection as discussed above. In more than 95% of cases, both assignations provide similar results. Sequences not annotated by the KeyDNAtools likely result from the absence of the corresponding clade in the core reference database, low quality sequences or novel variants of the gene present in newly available sequences, not yet included in the core dataset. Conversely, sequences not assigned by the Crunch_Assign software are often chimera or low quality sequences. After a search by similarity, we offer the possibility to build a phylogenetic tree on the fly, using most similar sequences found by BLAST or Crunch_Assign.

Updates

We have developed a pipeline that allows to analyze a GenBank new release within a week. Most of the time spent is indeed in manual checking of conflicts after average linkage clusterings, as explained previously. As a result, updates of the PR^2 database will be done shortly after each GenBank new release. As a result, numbers provided in this paper will probably differ from that available from PR2 at publication time of this manuscript.

Acknowledgments

We warmly thank Marion Viprey for her pioneering helps with the construction of the PR^2 database.Computations have been done on the 'Mésocentre SIGAMM' machine, hosted by Observatoire de la Côte d'Azur, Nice, France.

Funding

Financial support was provided through the European Union's Seventh Framework Programmes (FP7) BIOMARKS (2008-6530, ERA-net Biodiversa) and MicroB3 (grant agreement N° 287589) and the following ANR (France) projects: AQUAPARADOX, PARALEX, and GIME.
Conflict of interest statement. None declared.

References

1. 1. López-García,P., Rodríguez-Valera,F., Pedrós-Alió,C., Moreira,D. (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature, 409:603–607.

2. 2. Moon-van der Staay,S.Y., Watcher,R.D., Vaulot,D. (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. Nature, 409:607–610.

3. 3. Pawlowski,J., Christen,R., Lecroq,B., Bachar,D., Shahbakkia,H.R., Amaral-Zettler,A., Guillou,L. (2011) Eukaryotic richness in the abyss: insights from pyrotag sequencing. PLoS ONE, 6: e18169, 10.1371/journal.pone.0018169

4. 4. Hartmann,M., Howes,C.G., Vaninsberghe,D., Yu,H., Bachar,D., Christen,R., Henrik,N.R., Hallam,S.J., Mohn,W.W. (2012) Significant and persistent impact of timber harvesting on soil microbial communities in Northern coniferous forests. ISME J. First published on Aug 2, 2012, 10.1038/ismej.2012.84

5. 5. Lecroq,B., Lejzerowicz,F., Bachar,D., Christen,R., Esling,P., Baerlocher,L., Østerås,M., Farinelli,L. and Pawlowski,J. (2011) Ultra-deep sequencing of foraminiferal microbarcodes unveils hidden richness of early monothalamous lineages in deep-sea sediments. Proc. Natl. Acad. Sci. USA, 108:13177-13182.

6. 6. Edgcomb,V., Orsi,W., Bunge,J., Jeon,S., Christen,R., Leslin,C., Holder,M., Taylor,G.T., Suarez,P., Varela,R. and Epstein,S. (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. ISME J. 5:1344-1356.

7. 7. Behnke, A., Engel,M., Christen,R., Nebel,M., Klein,R.R. and Stoeck,T. (2011) Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. Environ. Microbiol. 13:340-349.

8. 8. Berney,C., Fahrni,J. and Pawlowski,J. (2004) How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys. BMC Biol., 2:13.

9. 9. Bachar,D., Guillou,L. and Christen,R. (2012) Detection of introns in eukaryotic small subunit ribosomal RNA gene sequences. Dataset Papers in Biology. DOI:10.7167/2013/854869.

10. 10. Guillou,L., Viprey,M., Chambouvet,A., Welsh,R. M., Massana,R., Scanlan,D.J., Worden,A. Z. (2008) Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). Environ. Microbiol.10: 3349-3365.

11. 11. Logares,R., Audic,S., Santini,S., Pernice,M. C., de Vargas,C. and Massana,R. (2012) Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. The ISME J. 1-11, 10.1038/ismej.2012.36.

12. 12. Adl,S.M., Simpson,A.G.B., Lane,C.E., Lukeš,J., Bass,D., Bowser,S.S., Brown,M.W., Burki,F., Dunthorn,M., Hampl,V., Heiss,A., Hoppenrath,M., Lara,E., Le Gall,L., Lynn,D.H., McManus,H., Mitchell,E.A.D., Mozley-Stanridge,S.E., Parfrey,L.W., Pawlowski,J.,Rueckert,S., Shadwick,L., Schoch,C.L., Smirnov,A., Spiegel,F.W. (2012) The revised classification of eukaryotes. J. Eukaryot. Microbiol. 59:429-493.

13. 13. Carnegie,R.B., Meyer,G.R., Blackbourn,J., Cochennec-Laureau,N., Berthe,F.C. and Bower,S.M. (2003) Molecular detection of the oyster parasite *Mikrocytos mackini*, and a preliminary phylogenetic analysis. Dis. Aquat. Organ. 54:219-27.

14. 14. Desper,R. and Gascuel,O. (2002) Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. J. Comput. Biol., 9:687-705.

15. 15. Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R., Thompson,J.D., Gibson,T.J., Higgins,D.G. (2007) Clustal W and Clustal X version 2.0. 23:2947-2948.

16. 16. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

17. 17. Katoh,K., Toh,H. (2010) Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics. 26:1899-900.

18. 18. Chevenet,F., Croce,O., Hebrard,M., Christen,R. and Berry,V. (2010) ScripTree: scripting phylogenetic graphics. Bioinformatics. 26:1125-1126.

19. 19. Croce,O., Chevenet,F. and Christen,R. (2010) A new web server for the rapid identification of microorganisms. J. Microbial Biochem. Technol. 2: 84-88.

20. 20. Pruesse,E., Quast,C., Knittel,K., Fuchs,B., Ludwig,W., Peplies,J., and Glöckner,F.O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 35:7188-7196.

21. 21. Cole,J.R., Chai,B., Farris,R.J., Wang,Q., Kulam,S.A., McGarrell,D.M., Garrity,G.M. and Tiedje,J.M. (2005) The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Res. 33: 294-296.

22. 22. DeSantis,T.Z., Hugenholtz,P., Larsen,N., Rojas,M., Brodie,E.L., Keller,K., Huber,T., Dalevi,D., Hu,P. and Andersen,G.L. (2006) Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl. Environ. Microbiol. 72:5069-72.

23. 23. Simpson,A.G., Roger,A.J., Silberman,J.D., Leipe,D.D., Edgcomb,V.P., Jermiin,L.S., Patterson,D.J. and Sogin,M.L. (2002) Evolutionary history of 'early-diverging' eukaryotes: the excavate taxon *Carpediemonas* is a close relative of Giardia. Mol. Biol. Evol. 19: 1782-1791

24. 24. Sunde,M. (2005) Class I integron with a group II intron detected in an *Escherichia coli* strain from a free-range reindeer. Antimicrob. Agents Chemother. 49:2512-2514.

25. 25. Ferat,J.L., Le Gouar,M. and Michel,F. (1994) Multiple group II self-splicing introns in mobile DNA from *Escherichia coli*. C. R. Acad. Sci. III. 317:141-148.

26. 26. Hugenholtz,P. and Huber,T. (2003) Chimeric 16S rDNA sequences of diverse origin are accumulating in the public databases. IJSEM, 53:289-293.

27. 27. Ashelford,K.E., Chuzhanova,N.A., Fry, J.C., Jones, A.J., Weightman, A.J. (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. Appl. Environ. Microbiol. 71:7724-7736.

28. 28. Huber,T., Faulkner,G. and Hugenholtz,P. (2004) Bellerophon; a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics, 20:2317-2319.

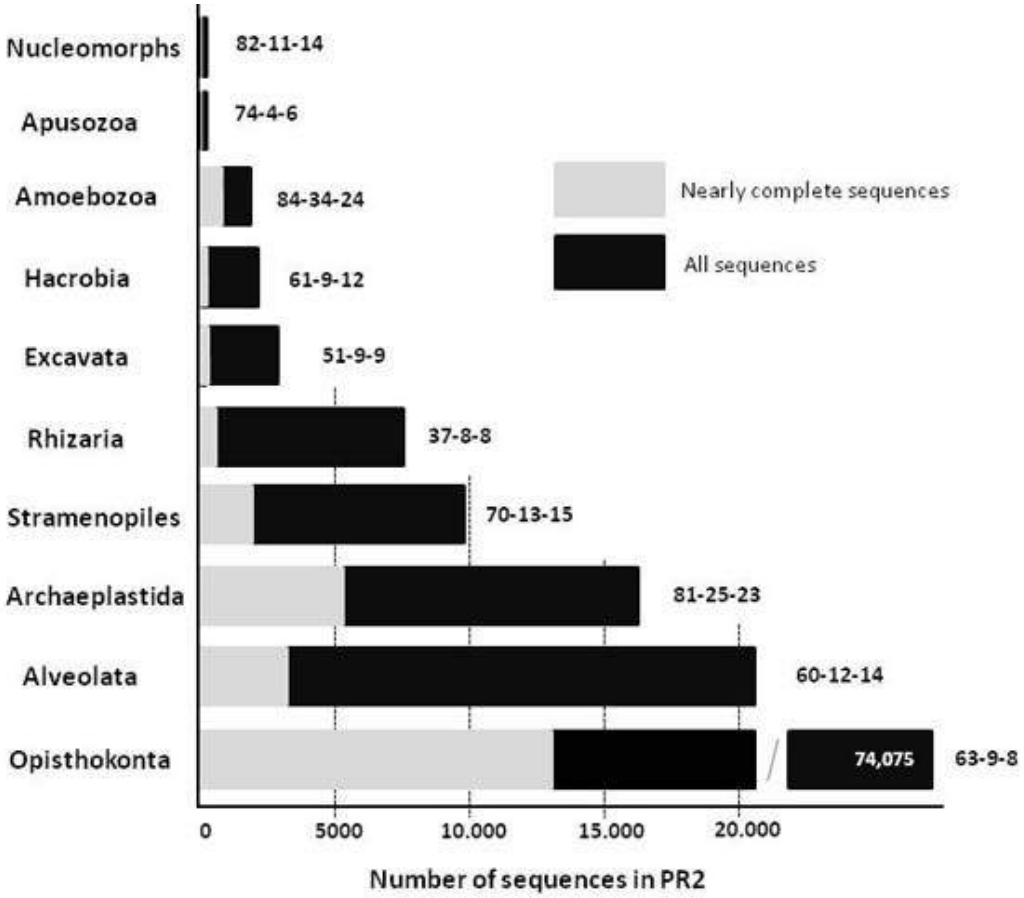


Figure 1. Total number of SSU rDNA gene sequences in the PR2 database for each main eukaryotic lineage (all sequences = grey+black, complete or nearly complete sequences in light-grey). Note that nucleomorphs were extracted from Archaeplastida. Numbers indicated after bars indicate percentages of sequences that include 1) the V4 region as defined by primers forward CCAGCASCYGCGGTAATTCC and reverse ACTTTCGTTCTTGATYRA used during the European Biomarks project, 2) the V9 region as defined by primers forward GTACACACCGCCCGTC and reverse TGATCCTTCTGCAGGTTCACCTAC used during the European Biomarks project, and 3) the V9 region defined by primers forward TTGTACACACCGCCC and reverse CCTTCYGCAGGTTCACCTAC used by the WAMPS project. For Opithokonta, number in white = total number of sequences.

Table 1 Number of nuclear-encoded sequences in PR2 as annotated at the Super-Group taxonomic level. n1: total number, n2: excluding putative chimera, Super-Group: rank-2 taxonomy.

Super-Group	n1	n2
Alveolata	20760	20255
Amoebozoa	1902	1880
Apusozoa	254	242
Archaeplastida	16309	16092
Eukaryota_Mikro	3	3
Eukaryota_X	54	54
Excavata	2871	2869
Hacrobia	2192	2132
Opisthokonta	75056	74484
Rhizaria	7581	7459
Stramenopiles	9884	9640
Total nuclear-encoded Eukaryota	136866	135110
Apicoplast	26	26
Chloroplast SSU	9657	9657
Hydrogenosome SSU	6	6
Mitochondrion SSU	36051	36051
Nucleomorph SSU (18S)	264	262