Protist 172 (2021) 125806
http://www.elsevier.de/protis
Published online date 16 April 2021

Protist

ORIGINAL PAPER

A Morpho-molecular Perspective on the Diversity and Evolution of Spumellaria (Radiolaria)

Miguel M. Sandin [a,1], Tristan Biard [b], Sarah Romac [a], Luis O'Dogherty [c], Noritoshi Suzuki [d] and Fabrice Not [a,1]

[a]Sorbonne Université, CNRS, AD2M-UMR7144 Station Biologique de Roscoff, 29680 Roscoff, France
[b]Univ. Littoral Côte d'Opale, Univ. Lille, CNRS, UMR 8187, LOG, Laboratoire d'Océanologie et de Géosciences, F 62930 Wimereux, France
[c]Facultad de Ciencias del Mar y Ambientales, Cadiz University, E-11510 Puerto Real, Spain
[d]Department of Earth Science, Graduate School of Science, Tohoku University, Sendai 980–8578, Japan

Submitted September 2, 2020; Accepted April 5, 2021
Monitoring Editor: B.S.C. Leadbeater

Spumellaria (Radiolaria, Rhizaria) are holoplanktonic amoeboid protists, ubiquitous and abundant in the global ocean. Their silicified skeleton preserves very well in sediments, displaying an excellent fossil record extremely valuable for paleo-environmental reconstruction studies, from where most of their extant diversity and ecology have been inferred. This study represents a comprehensive classification of Spumellaria based on the combination of ribosomal taxonomic marker genes (rDNA) and morphological characteristics. In contrast to established taxonomic knowledge, we demonstrate that symmetry of the skeleton takes more importance than internal structures at high classification ranks. Such reconsideration allows gathering different morphologies with concentric structure and spherical or radial symmetry believed to belong to other Radiolaria orders from the fossil record, as for some Entactinaria families. Our calibrated molecular clock dates the origin of Spumellaria in the middle Cambrian (ca. 515 Ma), among the first radiolarian representatives in the fossil record. This study allows a direct connection between living specimens and extinct morphologies from the Cambrian, bringing both a standpoint for future molecular environmental surveys and a better understanding for paleo-environmental reconstruction analysis.

Key words: Spumellaria; rDNA; molecular diversity; barcoding; molecular clock; molecular evolution.

[1]Corresponding authors;
e-mail miguelmendezsandin@gmail.com (M.M. Sandin).

https://doi.org/10.1016/j.protis.2021.125806

Introduction

Radiolaria are holoplanktonic amoeboid protists belonging to the Rhizaria lineage, one of the main branches of the eukaryotic tree of life (Burki and Keeling 2014). Along with Foraminifera they compose the phylum Retaria, characterized by cytoplasmic prolongations and a skeleton composed of different minerals (Adl et al. 2019). Spumellaria represents an important order within the Radiolaria (Suzuki and Not 2015), extensively studied across the fossil record thanks to its opaline silica skeleton (De Wever et al. 2001). Along with that of other radiolarians, the Spumellaria fossil record dates back to the early middle Cambrian (ca. 509–521 Ma; Aitchison et al. 2017; Suzuki and Oba 2015; Zhang and Feng 2019) constituting an important tool for paleo-environmental reconstructions analysis (e.g., Abelmann and Nimmergut 2005). In contemporary oceans, molecular-based metabarcoding surveys performed at global scale have shown that Radiolaria contribute significantly to plankton communities (de Vargas et al. 2015; Pernice et al. 2016) and the silica biogeochemical cycle (Llopis-Monferrer et al. 2020), although very little is known about their ecology and diversity. Living Spumellaria, as other Radiolaria, are characterized by a complex protoplasmic meshwork of pseudopodia extending radially from their skeleton (Suzuki and Aita 2011). With their pseudopodia, they actively capture prey, such as copepod nauplii or tintinnids, by adhesion (Sugiyama and Anderson 1997). In addition to predation, many spumellarian species dwelling in the sunlit ocean harbour photosynthetic algal symbionts, mainly identified as dinoflagellates (Probert et al. 2014; Yuasa et al. 2016; Zhang et al. 2018).

Radiolaria classifications are historically based on morphological criteria relying essentially on the initial spicular system, an early developed skeletal structure considered to be the foundation of the systematics at family and higher levels (De Wever et al. 2001). Traditionally, radiolarians with concentric structures and spherical or radial symmetry have been classified as Spumellaria or Entactinaria depending on the absence or presence, respectively, of such a skeletal structure. Yet these taxonomic differences, based on the initial spicular system, have been recently questioned, finding molecular evidence for the polyphyly of Entactinaria among Rhizaria (Nakamura et al. 2020). The latest spumellarian classification attempted to merge the extensive morphological criteria (De Wever et al. 2001; Noble et al. 2017; O'Dogherty et al. 2009) with recent rDNA molecular studies (Matsuzaki et al. 2015). An outcome from this later work, the current classification scheme describes nine extant superfamilies: Actinommoidea (Haeckel 1862; O'Dogherty 1994), Hexastyloidea (Haeckel 1882), Liosphaeroidea (Haeckel 1882), Lithelioidea (Haeckel 1862; Petrushevskaya 1975), Pylonioidea (Dumitrica 1989; Haeckel 1882), Saturnalioidea (Deflandre 1953), Spongodiscoidea (De Wever et al. 2001; Haeckel 1862), Sponguroidea (De Wever et al. 2001; Haeckel 1862), Stylodictyoidea (Haeckel 1882; Matsuzaki et al. 2015); and two undetermined families (Heliodiscidae, Haeckel, 1887; De Wever et al. 2001; and Spongosphaeridae Haeckel, 1862). In Matsuzaki et al. (2015) the authors stated the "overwhelmingly artificial" spumellarian classification outweighed by morphological criteria in regard to the little amount of molecular data available.

Few studies have explored the genetic diversity of Spumellaria, essentially unveiling relationships among higher rank taxonomical groups (Ishitani et al. 2012; Krabberød et al. 2011; Kunitomo et al. 2006; Yuasa et al. 2009). To date, with a total of 35 sequences from morphologically described specimens, covering 5 out of the 9 superfamilies described, mismatches between morphological and molecular characters are reported for two main groups of Spumellaria among which different innermost shell structures are found (Ishitani et al. 2012; Yuasa et al. 2009). Despite such important insights, the understanding of the inner structure relationships with the rest of the families remains elusive. Time-calibration of radiolarian phylogenies thanks to the fossil record allows a better understanding of relationships among extinct groups along with a contextualized evolutionary history (Decelle et al. 2012a; Lewitus et al. 2018; Sandin et al. 2019). In addition, the morphological description of single cell reference DNA barcodes establishes the basis for further molecular ecology surveys, allowing us to infer the actual diversity and ecology in the extant oceans through metabarcoding approaches (Biard et al. 2017; Decelle et al. 2013; Nitsche et al. 2016).

Here we present an integrative morpho-molecular classification of Spumellaria obtained by combining ribosomal DNA genes (18S and 28S partial rDNA), widely used as taxonomical markers, and imaging techniques (light and scanning electron microscopy). Phylogenetic placement of environmental sequences provided insights in the extant genetic diversity of Spumellaria in the contemporary oceans, allowing access to undescribed diversity. In addition, the extensive fossil record of Spumellaria allowed calibrating our phylogenetic analysis in time and inferring their evolutionary history contextualized with global scale geological and environmental changes from the Cambrian to present ecosystems.

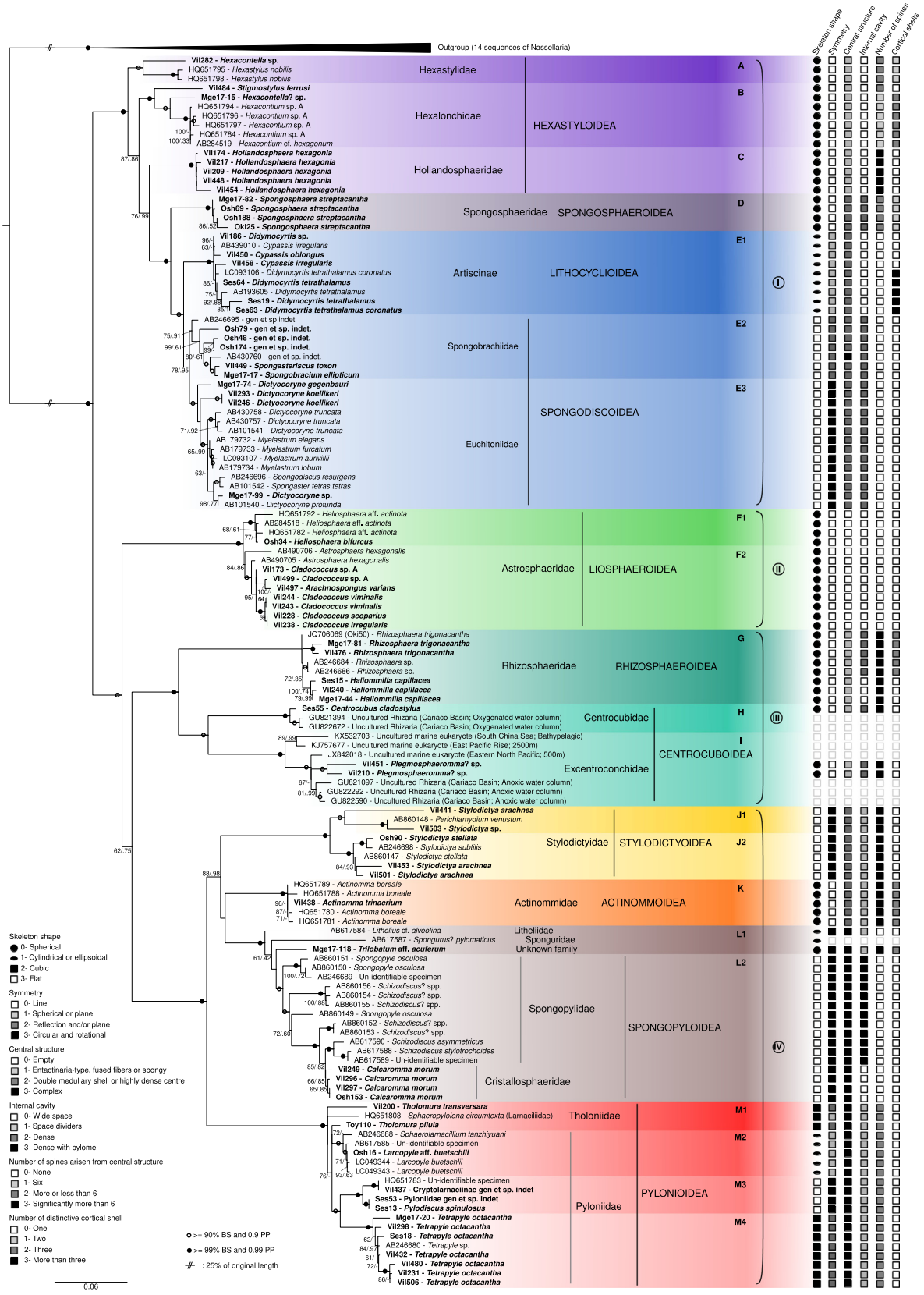
Results

Comparative Molecular Phylogeny and Morphological Taxonomy

Our final molecular phylogeny is composed of 133 distinct spumellarian sequences of the full 18S rDNA gene, completed with 55 sequences of the partial 28S (D1 and D2 regions) rDNA gene (Supplementary Material Table S1). From the 133 final sequences of the 18S rDNA, 67 were obtained in this study, 58 were previously available sequences associated to a morphologically described specimen and 8 were environmental sequences not affiliated to any morphology, added to increase phylogenetic support in poorly represented clades. Regarding the partial 28S rDNA, a total of 37 sequences were obtained in this study, where 18 were previously available. The final alignment matrix has 25.45% of invariant sites.

Phylogenetic analyses show 13 different clades (clades A, B, C, D, E, F, G, H, I, J, K, L and M) clearly differentiated by BS and PP values distributed in 4 main lineages (I, II, III, and IV) showing high BS (>97) and PP (1) and consistent position across the different phylogenetic analyses (Fig. 1). In general, morphological classification agrees with molecular phylogeny at the clade level. Considered together, morphological similarities and molecular phylogenetic support established super-families Hexastyloidea (spread in clades A, B and C), Spongosphaeroidea (clade D), Lithocyclioidea (clade E1) and Spongodiscoidea (clade E2 and E3) in lineage I; Liosphaeroidea (clade F) in lineage II; Rhizosphaeroidea (clade G) and Centrocuboidea (clades H and I) in lineage III and Actinommoidea (clade K), Stylodictyoidea (clade J), Spongopyloidea (clade L2) and Pylonioidea (clade M) in lineage IV with two unresolved families (clade L1, Litheliidae and Sponguridae). Superfamily and family definitions are based in similarity of morphological characters, shared synapomorphies and phylogenetic clustering support into the so-called morpho-molecular clades. All these morpho-molecular clades are highly supported in both the 18S rDNA and 28S rDNA gene phylogenies, except clades B and E in the 28S rDNA gene phylogeny (Supplementary Material Fig. S1). In contrast to the clade definition, the general topology and relationships between clades slightly disagree between the 18S rDNA and the 28S rDNA gene phylogenies. Such discrepancies are basically due to the variable position of clade F that in the 18S rDNA gene phylogeny appears basal to all clades and in the 28S rDNA gene appears basal to clades J, K, L and M, weakly supported as a sister group of clades H and G. Another example is, clades B, C, D and E, which are in the same highly supported group in the 18S rDNA gene phylogeny, whereas in the 28S rDNA gene phylogeny these clades appear scattered at basal positions regarding the rest of the clades. In the concatenation of the two genes clade F appears highly supported as a group with the clades G to M (Fig. 1).

Lineage I includes clades A, B, C, D and E and is phylogenetically distant to the other 3 lineages. Lineage I is characterized by the presence of one or two concentric shells or a full spongy test, sometimes non-distinguishable, where the main spines grow from the innermost shell and go through the outermost shell, when present (Fig. 2A-E). Only clade A and B shows two concentric shells with six primary spines from the center (Fig. 2A-B). Clade A is composed of 1 novel and 2 previously available sequences and a common characteristic is the fragile and hexagonal mesh constituting the outermost shell (Fig. 2A). Clade B is composed of two novel and five previously available sequences clustered with high BS and PP values despite their phylogenetic distance (Fig. 1). All specimens of clade B show a more irregular and thicker mesh of the shell compare to that of clade A, and, when present, very

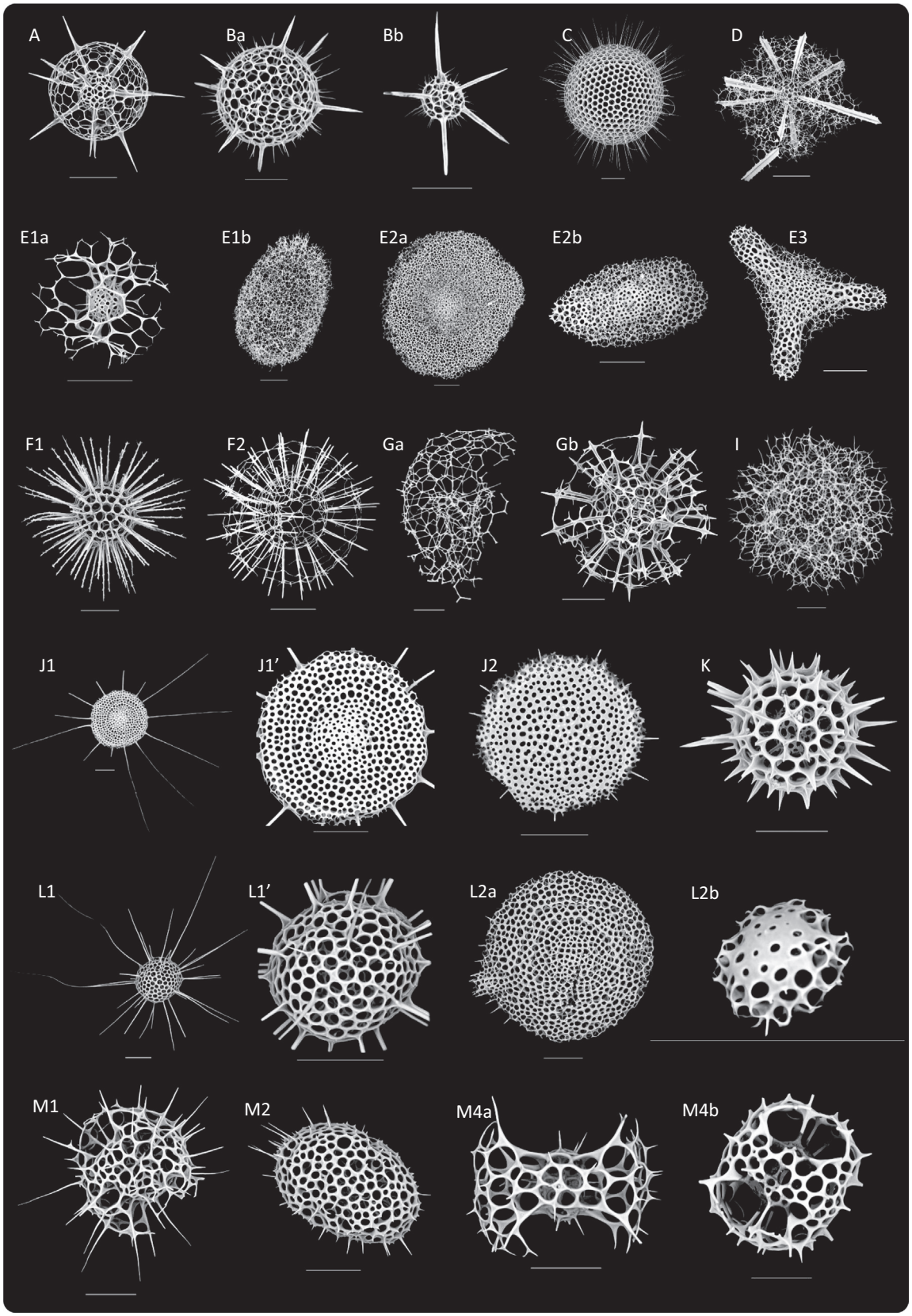

Outgroup (14 sequences of Nassellaria)
Skeleton shape
Symmetry
Central structure
Internal cavity
Number of spines
Cortical shells
Vil282 - Hexacontella sp.
HQ651795 - Hexastylus nobilis
HQ651798 - Hexastylus nobilis
Hexastylidae
A
Vil484 - Stigmostylus ferrusi
Mge17-15 - Hexacontella? sp.
HQ651794 - Hexacontium sp. A
HQ651796 - Hexacontium sp. A
HQ651797 - Hexacontium sp. A
HQ651784 - Hexacontium sp. A
AB284519 - Hexacontium cf. hexagonum
Hexalonchidae
HEXASTYLOIDEA
B
Vil174 - Hollandosphaera hexagonia
Vil217 - Hollandosphaera hexagonia
Vil209 - Hollandosphaera hexagonia
Vil448 - Hollandosphaera hexagonia
Vil454 - Hollandosphaera hexagonia
Hollandosphaeridae
C
Mge17-82 - Spongosphaera streptacantha
Osh69 - Spongosphaera streptacantha
Osh188 - Spongosphaera streptacantha
Oki25 - Spongosphaera streptacantha
Spongosphaeridae
SPONGOSPHAEROIDEA
D
Vil186 - Didymocyrtis sp.
AB439010 - Cypassis irregularis
Vil450 - Cypassis oblongus
Vil458 - Cypassis irregularis
LC093106 - Didymocyrtis tetrathalamus coronatus
Ses64 - Didymocyrtis tetrathalamus
AB193605 - Didymocyrtis tetrathalamus
Ses19 - Didymocyrtis tetrathalamus
Ses63 - Didymocyrtis tetrathalamus coronatus
Artiscinae
LITHOCYCLIOIDEA
E1
AB246695 - gen et sp indet
Osh79 - gen et sp. indet.
Osh48 - gen et sp. indet.
Osh174 - gen et sp. indet.
AB430760 - gen et sp. indet.
Vil449 - Spongasteriscus toxon
Mge17-17 - Spongobracium ellipticum
Spongobrachiidae
E2
Mge17-74 - Dictyocoryne gegenbauri
Vil293 - Dictyocoryne koellikeri
Vil246 - Dictyocoryne koellikeri
AB430758 - Dictyocoryne truncata
AB430757 - Dictyocoryne truncata
AB101541 - Dictyocoryne truncata
AB179732 - Myelastrum elegans
AB179733 - Myelastrum furcatum
LC093107 - Myelastrum aurivillii
AB179734 - Myelastrum lobum
AB246696 - Spongodiscus resurgens
AB101542 - Spongaster tetras tetras
Mge17-99 - Dictyocoryne sp.
AB101540 - Dictyocoryne profunda
Euchitoniidae
SPONGODISCOIDEA
E3
I
HQ651792 - Heliosphaera aff. actinota
AB284518 - Heliosphaera aff. actinota
HQ651782 - Heliosphaera aff. actinota
Osh34 - Heliosphaera bifurcus
F1
AB490706 - Astrosphaera hexagonalis
AB490705 - Astrosphaera hexagonalis
Vil173 - Cladococcus sp. A
Vil499 - Cladococcus sp. A
Vil497 - Arachnospongus varians
Vil244 - Cladococcus viminalis
Vil243 - Cladococcus viminalis
Vil228 - Cladococcus scoparius
Vil238 - Cladococcus irregularis
Astrosphaeridae
LIOSPHAEROIDEA
F2
II
JQ706069 (Oki50) - Rhizosphaera trigonacantha
Mge17-81 - Rhizosphaera trigonacantha
Vil476 - Rhizosphaera trigonacantha
AB246684 - Rhizosphaera sp.
AB246686 - Rhizosphaera sp.
Ses15 - Haliommilla capillacea
Vil240 - Haliommilla capillacea
Mge17-44 - Haliommilla capillacea
Rhizosphaeridae
RHIZOSPHAEROIDEA
G
Ses55 - Centrocubus cladostylus
GU821394 - Uncultured Rhizaria (Cariaco Basin; Oxygenated water column)
GU822672 - Uncultured Rhizaria (Cariaco Basin; Oxygenated water column)
Centrocubidae
H
III
KX532703 - Uncultured marine eukaryote (South China Sea; Bathypelagic)
KJ757677 - Uncultured marine eukaryote (East Pacific Rise; 2500m)
JX842018 - Uncultured marine eukaryote (Eastern North Pacific; 500m)
Vil451 - Plegmosphaeromma? sp.
Vil210 - Plegmosphaeromma? sp.
GU821097 - Uncultured Rhizaria (Cariaco Basin; Anoxic water column)
GU822292 - Uncultured Rhizaria (Cariaco Basin; Anoxic water column)
GU822590 - Uncultured Rhizaria (Cariaco Basin; Anoxic water column)
Excentroconchidae
CENTROCUBOIDEA
I
Vil441 - Stylodictya arachnea
AB860148 - Perichlamydium venustum
Vil503 - Stylodictya sp.
J1
Osh90 - Stylodictya stellata
AB246698 - Stylodictya subtilis
AB860147 - Stylodictya stellata
Vil453 - Stylodictya arachnea
Vil501 - Stylodictya arachnea
Stylodictyidae
STYLODICTYOIDEA
J2
HQ651789 - Actinomma boreale
HQ651788 - Actinomma boreale
Vil438 - Actinomma trinacrium
HQ651780 - Actinomma boreale
HQ651781 - Actinomma boreale
Actinommidae
ACTINOMMOIDEA
K
AB617584 - Litheliius cf. alveolina
AB617587 - Spongurus? pylomaticus
Litheliidae
Sponguridae
Unknown family
L1
Mge17-118 - Trilobatum aff. acuferum
AB860151 - Spongopyle osculosa
AB860150 - Spongopyle osculosa
AB246689 - Un-identifiable specimen
AB860156 - Schizodiscus? spp.
AB860154 - Schizodiscus? spp.
AB860155 - Schizodiscus? spp.
AB860149 - Spongopyle osculosa
AB860152 - Schizodiscus? spp.
AB860153 - Schizodiscus? spp.
AB617590 - Schizodiscus asymmetricus
AB617588 - Schizodiscus stylotrochoides
AB617589 - Un-identifiable specimen
Spongopylidae
SPONGOPYLOIDEA
L2
Vil249 - Calcaromma morum
Vil296 - Calcaromma morum
Vil297 - Calcaromma morum
Osh153 - Calcaromma morum
Cristallosphaeridae
IV
Vil200 - Tholomura transversara
HQ651803 - Sphaeropylolena circumtexta (Larnaciliidae)
Toy110 - Tholomura pilula
Tholoniidae
M1
AB246688 - Sphaerolarnacillium tanzhiyuani
AB617585 - Un-identifiable specimen
Osh16 - Larcopyle aff. buetschlii
LC049344 - Larcopyle buetschlii
LC049343 - Larcopyle buetschlii
M2
HQ651783 - Un-identifiable specimen
Vil437 - Cryptolarnaciinae gen et sp. indet
Ses53 - Pyloniidae gen et sp. indet
Ses13 - Pylodiscus spinulosus
Pyloniidae
PYLONIOIDEA
M3
Mge17-20 - Tetrapyle octacantha
Vil298 - Tetrapyle octacantha
Ses18 - Tetrapyle octacantha
AB246680 - Tetrapyle sp.
Vil432 - Tetrapyle octacantha
Vil480 - Tetrapyle octacantha
Vil231 - Tetrapyle octacantha
Vil506 - Tetrapyle octacantha
M4
Skeleton shape
0- Spherical
1- Cylindrical or ellipsoidal
2- Cubic
3- Flat
Symmetry
0- Line
1- Spherical or plane
2- Reflection and/or plane
3- Circular and rotational
Central structure
0- Empty
1- Entactinaria-type, fused fibers or spongy
2- Double medullary shell or highly dense centre
3- Complex
Internal cavity
0- Wide space
1- Space dividers
2- Dense
3- Dense with pylome
Number of spines arisen from central structure
0- None
1- Six
2- More or less than 6
3- Significantly more than 6
Number of distinctive cortical shell
0- One
1- Two
2- Three
3- More than three
>= 90% BS and 0.9 PP
>= 99% BS and 0.99 PP
: 25% of original length
0.06

thin spines (generally called by-spines) coming out from all the surface of the shell (Fig. 2Ba and Bb). Clade C is composed of five novel sequences showing a large single double layered shell with no main spines but a large abundance of long and thin spines (Fig. 2C). These three clades agree with the definition of the Superfamily Hexastyloidea (Haeckel 1882; Matsuzaki et al. 2015) displaying an innermost shell that, if present, always shows pyriform with six radial beams. Three different families within Hexastyloidea correspond to: Hexastylidae (Haeckel 1887) in clade A, Hexalonchidae (Haeckel 1882) in clade B and Hollandosphaeridae (Deflandre 1973) in clade C. The innermost shell if present always shows pyriform with six radial beams. Clades D and E are the more distal clades in this lineage and highly supported together (100 BS and 1 PP). The presence of spongy structures differentiates these two clades from the others of lineage I (Fig. 2D-E). Clade D is composed of four novel sequences obtained in this study. All specimens of this clade show a single, very small spherical shell where several long and three bladed spines grow interconnected by a spongy mesh (Fig. 2D). This definition agrees with the genus *Spongosphaera* belonging to the Superfamily Spongosphaeroidea (Haeckel 1862). The last clade of lineage I (clade E) is composed of 30 sequences of which 15 are novel. Members of this clade tend to lose the concentric symmetry towards a cylindrical to ellipsoidal (sub-clade E1; Fig. 2E1a and E1b) or flat symmetry (sub-clade E2, Fig. 2E2a and Eb; and sub-clade E3, Fig. 2E3), and so does the spines and the shell. Spongy structures take more importance complicating the inner structure. These morphologies agree with the definition of the Family Artiscinae (Haeckel 1882) for E1 and the Superfamily Spongodiscoidea (De Wever et al. 2001; Haeckel 1862) for E2 and E3. Within the latest clades there are representatives of two families Spongobrachiidae (Haeckel 1882) with moderate BS and PP values in sub-clade E2 and Eucthitoniidae (Haeckel 1882) highly supported in sub-clade E3.

Lineage II, including clade F alone, is composed of 13 sequences of which five were previously available and eight are novel. All specimens within this clade are characterized by one large and hollow shell where long spines grow from its surface (Fig. 2F1 and F2), that agree with the definition of the Superfamily Liosphaeroidea (Haeckel 1882). Members of the sub-clade F1 (Fig. 2F1) share a robust shell compared to that of sub-clade F2 (Fig. 2F2) where short spicules come out of the spines. Those specimens agree with the definition of the family Astrosphaeridae (Haeckel 1887; Matsuzaki et al. 2015) in the large hollow and fragile shell with the presence of long or short spicules coming out at the far end of the spines.

Within lineage III, including clades G, H and I, clade G is represented by five novel and three previously available sequences. Members of this clade have a unique innermost shell, named "rhizosphaerid-type microsphere" or "microbursa" (Dumitrica et al. 2010) with a large and fragile spherical skeleton (Fig. 2Ga) or several cortical shells with a robust skeleton (Fig. 2Gb). Either morphology agree with the definition of the family Rhizosphaeridae (Dumitrica 2017; Hollande and Enjumet 1960). Clade H is composed by one novel sequence with a morphological reference (Ses55, Supplementary Material Fig. S2) and two environmental sequences not morphologically described. The picture available for this specimen lacks taxonomic resolution, although the skeletal structure resembles that shown in Aita et al. (2009, pl. 23, Fig. 3) specified as *Centrocubus cladostylus* (Haeckel 1887). This

Figure 1. Molecular phylogeny of Spumellaria inferred from the concatenated complete 18S and partial 28S (D1-D2 regions) rDNA genes (135 taxa and 2459 aligned positions). The tree was obtained by using a phylogenetic Maximum likelihood method implemented using the GTR + γ + I model of sequence evolution. PhyML bootstrap values (100 000 replicates, BS) and posterior probabilities (PP) are shown at the nodes (BS/PP). Black circles indicate BS $\geq$ 99% and PP $\geq$ 0.99. Hollow circles indicate BS $\geq$ 90% and PP $\geq$ 0.90. Sequences obtained in this study are shown in bold. Thirteen main clades are defined based on statistical support and morphological criteria (A, B, C, D, E, F, G, H, I, J, K, L, M), divided in 4 main lineages (I, II, III, IV). For each morpho-molecular clade, capital letters represent the Superfamily name and small letters the Family name. Main skeletal traits are identified on the right for each spumellarian specimen. Fourteen Nassellaria sequences were assembled as outgroup. Branches with a double barred symbol are fourfold reduced for clarity.


A
Ba
Bb
C
D
E1a
E1b
E2a
E2b
E3
F1
F2
Ga
Gb
I
J1
J1'
J2
K
L1
L1'
L2a
L2b
M1
M2
M4a
M4b

species has characteristic interconnected spines and the highly supported position between clade G and I agrees with the definition of the family Centrocubidae (Hollande and Enjumet 1960; emend. Dumitrica 1994). The last clade of this lineage (clade I) is represented by two novel sequences obtained in this study and 6 environmental sequences. Specimens of this clade show only spongy structures with radiated fibers and there is no evidence of either shell or main spines (Fig. 2I).

In lineage IV, clade J, K and L cluster together with relatively high BS (88) and PP (0.98) values as a sister group of clade M. Although relationships between clades J, K and L remain still elusive due to the short phylogenetic distance and the different topology when using a Maximum Likelihood (Fig. 1) or a Bayesian (Fig. 3) approach (BS: 49 between clade K and L and PP: 0.38 between clade J and K). All members of this lineage share a very small spherical inner most shell with four or more radial beams connecting to the second shell, where 12 or more radial beams come out (Fig. 2J-M). lineage IV differs from lineage I in the absence of discrete radial beams from the second or later shells. Clade J is characterized by five sequences obtained in this study and three previously available sequences. Members of this clade share a flattened skeleton with two or more concentric shells and radial spines from the innermost or second innermost shell (Fig. 2J1, J1′ and J2), agreeing with the definition of the Superfamily Stylodictyoidea (Haeckel 1882; Matsuzaki et al. 2015). Morphological differences between sub-clade J1 (Fig. 2J1) and J2 (Fig. 2J2) were not possible to determine, yet they have been separated into two different subclades due to high BS (96 in J1 and 100 in J2) and PP (0.98 in J1 and 1 in J2) values. Clade K contains one novel and four previously available sequences. All specimens of this clade have three spherical shells with more than eight main spines that extend from the inner shell (Fig. 2K), representing the Superfamily Actinommoidea (Haeckel 1862). Clade L is represented by five novel and 14 previously available sequences. Members of this clade are morphologically distant. Sub-clade L1 has a tightly concentric or coiled center with one thick protoplasmic pseudopodium (axopodium) (AB61787) or a rhomboidally inflated center (Fig. 2L1 and L1′), which is not observed in any other Spumellaria. In the first case, this morphology is attributed to the family Sponguridae, and in the second to Litheliidae. Sub-clade L2 is represented by either: a flatten circular test with a tunnel-like pylome and a test comprised by a very densely concentric structure (Fig. 2L2a) attributed to Spongodiscoidea; or by a spherical translucent protoplasm with an encrypted flat consolidated skeletal shell in distal position and star-like solid soluble materials (Fig. 2L2b and Supplementary Material Fig. S2) assigned to Collodaria. In clade M there are 13 novel sequences and seven previously available sequences. A morphological characteristic of this clade is the broken, or fenestrated, outermost shell with the presence of pyloniid central structure (Fig. 2M), corresponding to the Superfamily Pylonioidea (Haeckel 1882), thereafter, the symmetry and the different opening of the outermost shell distinguishes the sub-clades. Yet, phylogenetic relationships within this clade remains elusive. The first sub-clade (M1) consists of three basal sequences poorly supported as a clade yet

Figure 2. Scanning Electron Microscope (SEM) images of Spumellaria specimens used in this study for phylogenetic analysis or morphologically related to one of the morpho-molecular clades of Fig. 1. Letters correspond to its phylogenetic clade in Fig. 1. Scale bars = 50 μm. (**A**) Osh194: *Hexacontium sp.* (specimen not included in phylogeny). (**Ba**) Vil484: *Stigmostylus ferrusi*. (**Bb**) Mge17-15: *Hexacontella*? sp. (**C**) Vil174: *Hollandosphaera hexagonia*. (**D**) Mge17-82: *Spongosphaera streptacantha*. (**E1a**) Vil186: *Didymocyrtis* sp. (**E1b**) Vil458: *Cypassis irregularis*. (**E2a**) Osh48: Spongodiscidae gen et sp. Indet. A (**E2b**) Mge17-17: *Spongobracium ellipticum*. (**E3**) Vil246: *Dictyocoryne koellikeri*. (**F1**) Osh34: *Heliosphaera bifurcus*. (**F2**) Vil499: *Cladococcus* sp. A. (**Ga**) Vil240: *Haliommilla capillacea* (broken specimen). (**Gb**) Mge17-81: *Rhizosphaera trigonacantha*. (**I**) Vil210: *Plegmosphaeromma*? sp. (**J1**) Vil441: *Stylodictya arachnea*. (**J1′**) detail of Vil441. (**J2**) Osh90: *Stylodictya stellata*. (**K**) Vil438: *Actinomma trinacrium*. (**L1**) Mge17-118: *Trilobatum* aff. *acuferum*. (**L1′**) detail of Mge17-118. (**L2a**) Osh191: *Schizodiscus* spp. (specimen not included in phylogeny). (**L2b**) Vil296: *Calcaromma morum*. (**M1**) Vil200: *Tholomura transversara*. (**M2**) Osh16: *Larcopyle* aff. *buetschlii*. (**M4a**) Mge17-20: *Tetrapyle octacantha*. (**M4b**) Vil452: *Tetrapyle octacantha* (specimen not included in phylogeny).

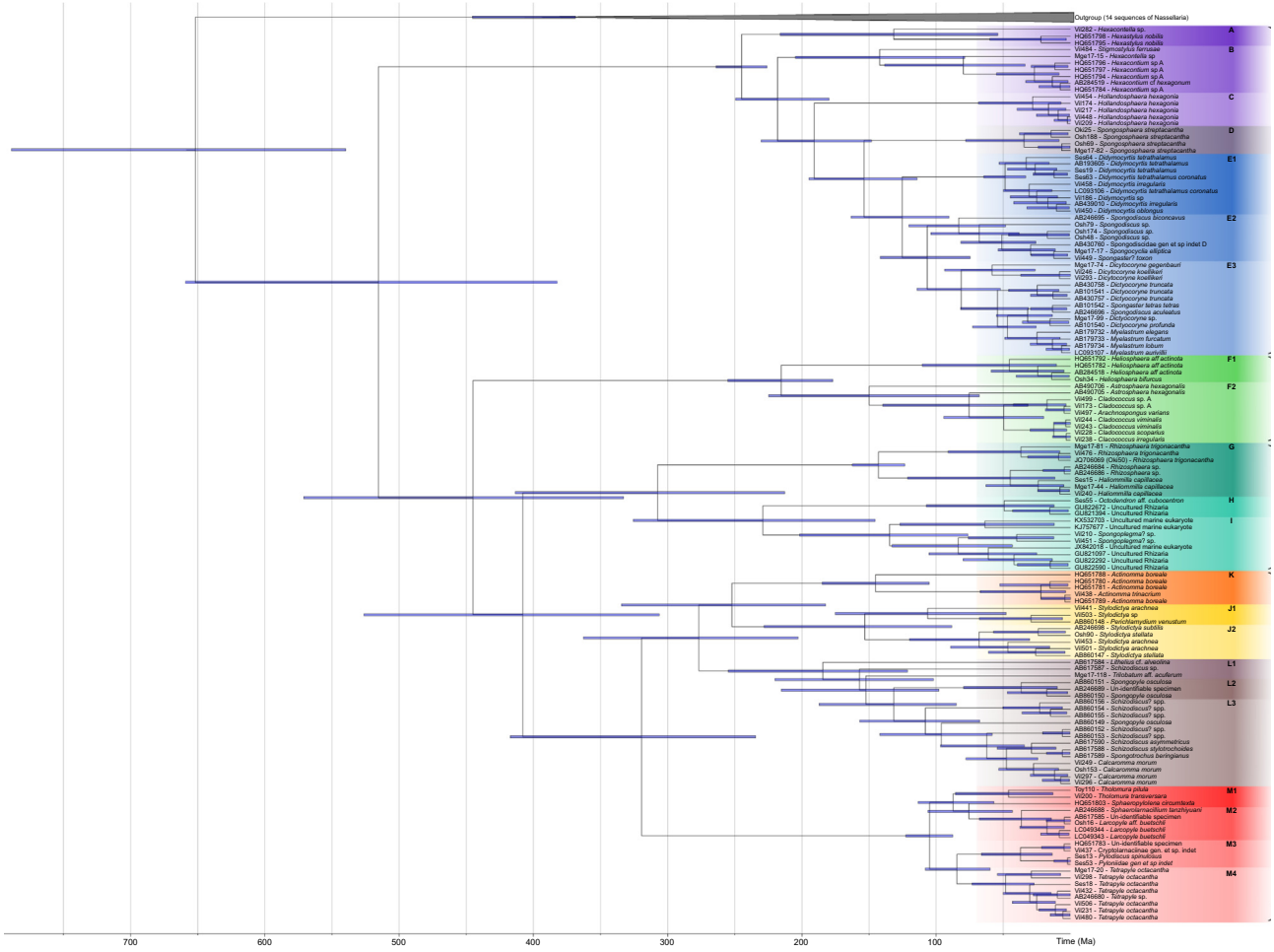


Figure 3. Time-calibrated tree (Molecular clock) of Spumellaria, based on alignment matrix used for phylogenetic analyses. Node divergences were estimated with a Bayesian relaxed clock model and the GTR + γ + I evolutionary model, implemented in the software package BEAST. Nine different nodes were selected for the calibration (blue dots). Blue bars indicate the 95% highest posterior density (HPD) intervals of the posterior probability distribution of node ages.

sharing a cubic morphology and the two opposite and closed gates (Fig. 2M1), which agrees with the family Tholoniidae (Haeckel 1887). The three other sub-clades have different morphologies (Fig. 2M2; Ses13, Ses53 and Vil437, Supplementary Material Fig. S2) agreeing with the definition of different families within Pylonioidea. Sub-clade M2 (Fig. 2M2) has a spherical to ellipsoidal skeleton without distinctive openings. On the other side, sub-clades M3 and M4 are highly supported as a group with a flattened box-shaped skeleton for the former sub-clade (Ses13, Ses53 and Vil437, Supplementary Material Fig. S2) and a cubic skeleton characterized by big openings of the second shell for the later sub-clade (Fig. 2M4a and 4b).

Molecular Dating

The molecular clock dated the diversification between Nassellaria and Spumellaria (the root of

the tree) at a median value of 651 Ma (95% Highest Posterior Density -HPD-: between 789 and 540 Ma) (Fig. 3). From here on, all dates are expressed as median values followed by the 95% HPD interval. The first diversification of Spumellaria happened around 515 (659–382) Ma, followed by a rapid branching of the lineages. Despite their dubious phylogenetic relationships, lineage II splits apart 445 (571–332) Ma and lineage III and IV diversified from each other 407 (526–306) Ma. The next diver-

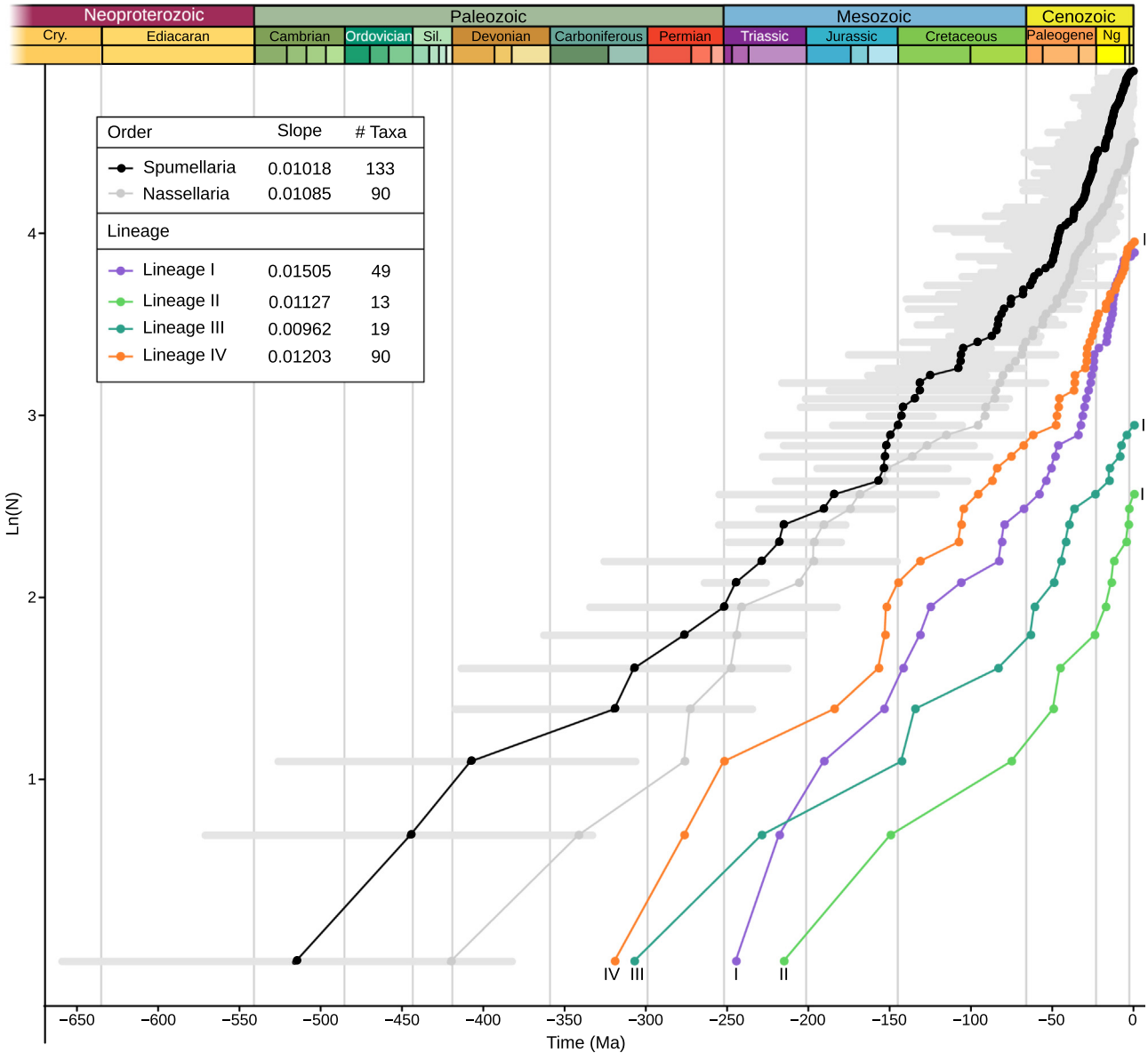


Figure 4. Lineages Through Time (LTT) analysis based on the molecular clock results for Spumellaria, removing the outgroup (in black) and of each lineage independently: lineage I (purple), lineage II (light green), lineage III (dark green) and lineage IV (orange) and of Nassellaria (data taken from Sandin et al. 2019). The y-axis represents the number of lineages (N) expressed in logarithmic (base e) scale (Ln(N)) and in the x-axis it is represented the time in millions of years ago (Ma). Horizontal gray bars in black slope represent the 95% Highest Posterior Density (HPD) of molecular clock estimates.

sification events correspond to the first radiation of lineage IV 319 (417–234) Ma and that of lineage III 307 (414–213) Ma. Within lineage IV, the phylogenetic relationships are doubtful, yet two radiation events at 277 (363–203) Ma and 252 (334–182) Ma separate clades J, K and L. clade L is the first of these clades diversifying at 184 (255–121) Ma, followed by clade J 153 (228–88) Ma, clade K at 145 (185–105) Ma and clade M at 105 (123–88) being the youngest clade of lineage IV. Lineage III diverged soon after lineage IV, yet the next ramification was between clade G and clade I 229 (326–145) Ma. It was not until 143 (162–123) and 135 (202–76) Ma when clade G and I respectively radiate, and the latest diversification within this lineage corresponds to clade H 49 (107–12) Ma. Lineage I appeared 245 (264–226) Ma, and thereafter is characterized by a series of relatively continuous diversification events until the radiation of clades B, A, E, D and C at 142 (205–79), 131 (216–54), 153 (163–90), 34 (78–8) and 28 (68–7) Ma respectively. Regarding lineage II, it is the youngest of all the lineages appearing at 215 (255–177) Ma despite its early branching from any other lineage.

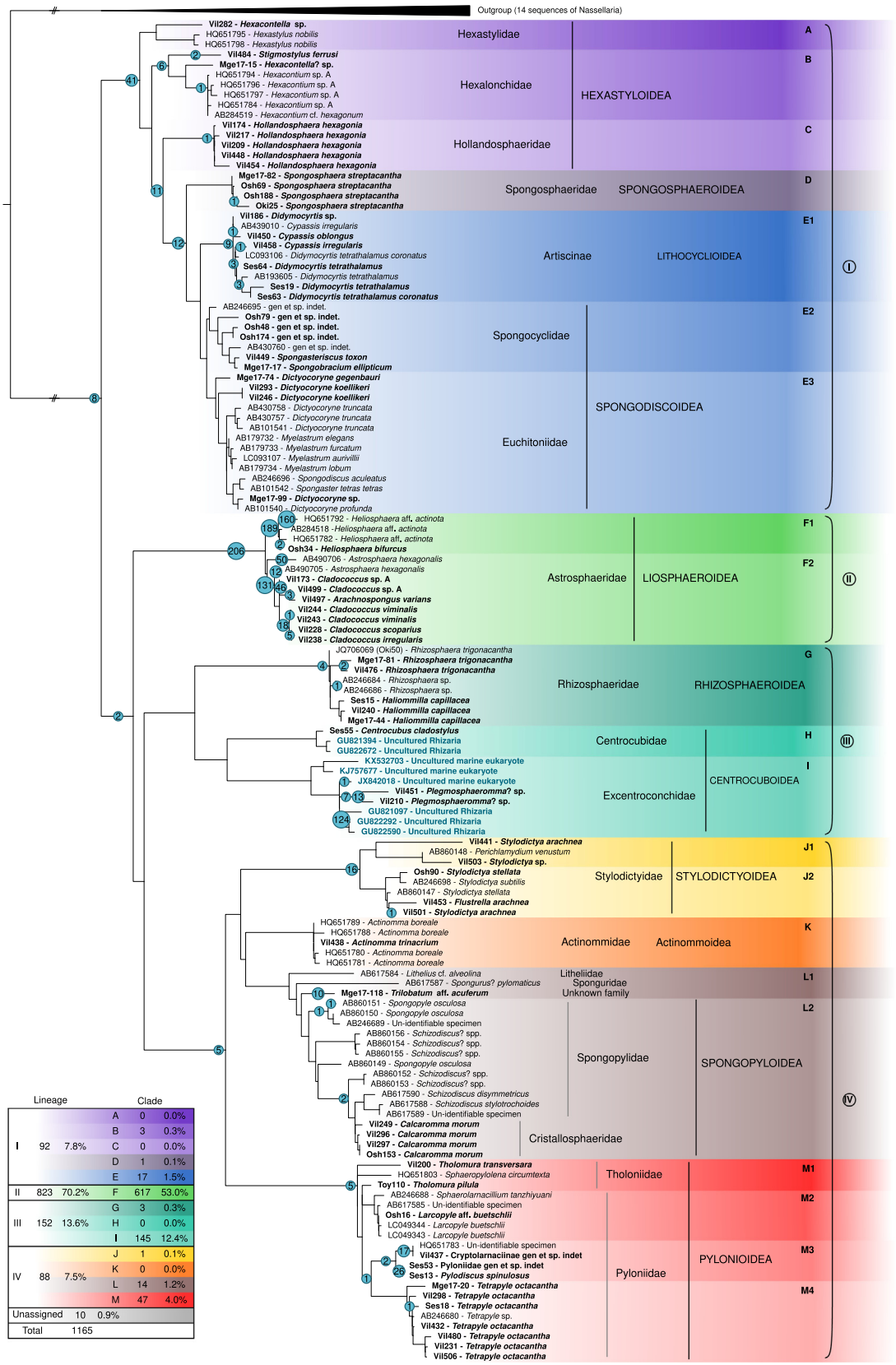
Post-hoc Analyses

The lineages through time analysis (Fig. 4) showed a classic exponential diversification slope, with a 0.0102 rate of speciation (Ln(lineages)·million years^{-1}). The first diversification of extant Spumellaria corresponds to an early and fast divergence of the different lineages from ~515 to ~407 Ma. After that, the diversification of living groups remained at standstill until ~313 Ma when lineage IV and III radiated. From ~276 to ~215 Ma all different groups diversified: clades K, J and L (within lineage IV) split apart; lineage I appeared and started a continued diversification; within lineage III the three clades split apart; and lineage II emerged. Another important diversification event happened from ~157 to ~125 Ma when lineage IV greatly diversified, followed by lineage I and II. During the following years there was a tiered diversification, where clade M appeared in isolated events as well as the ramification of clade L or the branching between subclades E2 and E3. Finally, from ~54 Ma onwards there was a continuous diversification where the rest of the clades appeared or keep diversifying. During this time lineage III diversified notably and so did lineage II.

The ancestral state reconstruction analysis (Supplementary Material Fig. S3) established a spherical skeleton shape with a line of symmetry for the common ancestor to all Spumellaria, with a central structure that can be either empty (state 0) or with an entactinarian-type microsphere or fused fibers or soft spongy system (state 1), but with a wide space in between the central structure and the cortical shell, without spines arising from it and one distinctive cortical shell. The common ancestor of lineage I and that of lineages II, III and IV share a spherical skeleton shape with a line of symmetry, a wide internal cavity and one cortical shell. The difference between these two ancestors are that the ancestor of lineage I is characterized by an entactinarian-type central structure or fused fibers or soft spongy system with more or less six spines, whereas that of lineages II, III, IV could also have an empty central structure and have no spines arising from the central structure. The common ancestor for lineage II has the most ancestral state in every reconstructed trait (state 0). The common ancestor of lineages III and IV also have a spherical skeleton shape with a line of symmetry, like the common ancestor of these two lineages independently. Although the common ancestor of lineages III and IV differs from that of lineage II in a higher complexity of the central structure, the appearance of structures in the internal cavity, the presence of significantly more than six spines and that it could have three cortical shells. Lastly, the common ancestor of the lineage III and that of lineage IV could also have one or three cortical shells with significantly more than six radial spines. Yet in lineage III the internal cavity gets denser than that of lineage IV, whereas the central structure tends to be entactinarian-type, fused fibers or soft spongy in lineage III and that of the common ancestor of lineage IV is undefined since it has equal probabilities every character state.

Environmental Genetic Diversity of Spumellaria

A total of 1165 environmental sequences affiliated to Spumellaria, and retrieved from NCBI public data-



Lineage			Clade		
Ⅰ	92	7.8%	A	0	0.0%
			B	3	0.3%
			C	0	0.0%
			D	1	0.1%
			E	17	1.5%
Ⅱ	823	70.2%	F	617	53.0%
Ⅲ	152	13.6%	G	3	0.3%
			H	0	0.0%
			I	145	12.4%
Ⅳ	88	7.5%	J	1	0.1%
			K	0	0.0%
			L	14	1.2%
			M	47	4.0%
Unassigned		10	0.9%		
Total		1165			

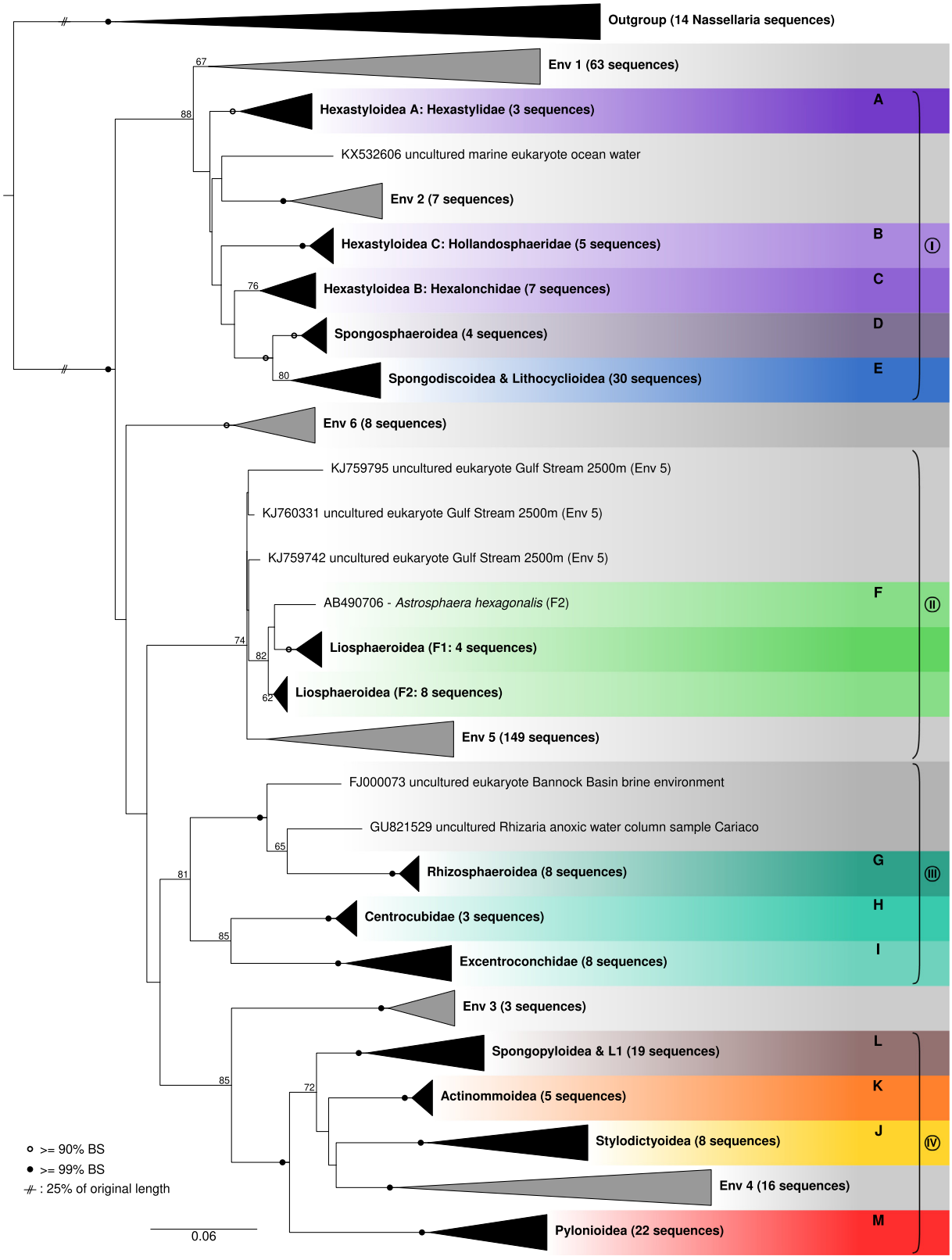
base, were placed in our reference phylogenetic tree (Fig. 5; Supplementary Material Table S3). From these, 617 were closely related to clade F. Two other clades with a high number of environmental sequences related to are clade I and clade M, with 145 and 47 sequences respectively. While up to 40 sequences were scattered between clades E, L, B, G, D and J (17, 14, 3, 3, 1 and 1 respectively), no environmental sequences were clustered within clades A, C, H and K. The rest of the sequences (317) where distributed across the tree, mainly at basal nodes, with no possible assignation: 206 basal to clade F in lineage II, 70 in lineage I (of which 41 basal to the lineage, 11 basal to clades C, D and E, 12 basal to clades D and E, 6 basal to clade B and 1 basal to clade C), 26 in lineage IV (of which 16 basal to clade J, 5 basal to clade M and 5 basal to the lineage), 4 in lineage III basal to clade G, 2 basal to lineages II, III and IV and 8 final sequences basal to all Spumellaria.

To go beyond rapid phylogenetic placement using the pplacer tool, these 317 environmental sequences dubiously assigned, were later included in a phylogenetic tree of the 18S rDNA gene (RAxML v8.2.10, GTR + G + I, with 1000 rapid bootstraps, Stamatakis 2014) after removing chimeras (52 sequences were detected as chimeric) and suspicious sequences (11 sequences with multiple phylogenetic placements and/or long branches). Up to six different new clades were found (Env1 – 6; Fig. 6), six sequences scattered over the phylogenetic tree (one related to Env2, three sequences related to clade F and Env5 and two sequences closely related to clade G) and two last sequences were clustered with high support basal to clade M (Fig. 6; further details of sequence assignations can be found in Supplementary Material Table S3). From these clades, Env5 gathers 149 sequences related to clade F in lineage II, yet low bootstrap values (51) avoid a formal assignation. Two clades (Env1 with 63 sequences and Env2 with seven) are found at basal positions of lineage I. Another environmental clade appears in lineage IV (Env4 with 16 sequences), and two final clades are found basal to lineages I, III and IV (Env6 with eight sequences) and basal to lineage IV (Env3 with three sequences).

In order to improve the understanding of the diversity of the environmental clades (Fig. 6) and try to explain the high chimeric ratio (∼13–25%) observed, we performed a second phylogenetic analysis (RAxML v8.2.10, GTR + G + I, with 1000 rapid bootstraps, Stamatakis 2014). This analysis includes the dataset used to build Fig. 6 and nine other sequences obtained from single-cell isolates but previously removed due to the contrasting morphological assignation in the clustered phylogenetic clade. It turns out that these nine sequences clustered with moderate to high support among one of the environmental clades, clade Env1 (Fig. 7; Supplementary Material Table S1). These nine specimens show a morphology that matches clades E (Osh87 and Osh174), F (Osh156), G (Osh82, Osh97 and Osh114), I (Osh187), L (Osh116) and a last specimen with a morphology not covered in our morpho-molecular framework (Osh108). Our morpho-molecular classification show that each molecular clade has a clear and endorsed (two or more specimens) morphological patterns distinct from one another. The combination of different morphologies within the same clade Env1 contrasts with this global pattern. In addition, one of the specimens from clade Env1 (Osh174) is also found within clade E. Although in clade E it is composed by the concatenation of the first part of the 18S rDNA and the 28S rDNA gene (1489 bp in total) and that of clade Env1 by the second part of the 18S rDNA gene (871 bp).

To assess possible pseudogenes for clade Env1, we performed two analyses for the 18S rDNA: (i) an entropy analysis of the alignment and (ii) the GC content estimation of every sequence. No differences were found between the entropy of Env1 and that of the different lineages (Supplementary Material Fig. S4), showing an average Shannon entropy of 0.07 (±0.16 standard deviation; sd) in Env1, 0.08 (±0.19 sd) in lineage I, 0.03 (±0.08 sd) in lineage II, 0.11 (±0.24 sd) in lineage III and 0.14 (±0.26 sd) in lineage IV. The greatest entropy values correspond to the V4 hyper-variable region of the 18S rDNA gene (positions ∼ 450-∼850). Other important regions correspond to the V2 (around the position ∼ 200), V7 (∼1400) and V9 (∼1720) hyper-variable regions of the 18S rDNA gene. GC

Figure 5. Pplacer phylogenetic placement of 1165 environmental sequences into a concatenated phylogenetic tree of Spumellaria (complete 18S + partial 28S rDNA genes). Numbers at nodes represent the amount of environmental sequences assigned to a branch or a node.


Outgroup (14 Nassellaria sequences)
Env 1 (63 sequences)
Hexastyloidea A: Hexastylidae (3 sequences)
KX532606 uncultured marine eukaryote ocean water
Env 2 (7 sequences)
Hexastyloidea C: Hollandosphaeridae (5 sequences)
Hexastyloidea B: Hexalonchidae (7 sequences)
Spongosphaeroidea (4 sequences)
Spongodiscoidea & Lithocyclioidea (30 sequences)
Env 6 (8 sequences)
KJ759795 uncultured eukaryote Gulf Stream 2500m (Env 5)
KJ760331 uncultured eukaryote Gulf Stream 2500m (Env 5)
KJ759742 uncultured eukaryote Gulf Stream 2500m (Env 5)
AB490706 - Astrosphaera hexagonalis (F2)
Liosphaeroidea (F1: 4 sequences)
Liosphaeroidea (F2: 8 sequences)
Env 5 (149 sequences)
FJ000073 uncultured eukaryote Bannock Basin brine environment
GU821529 uncultured Rhizaria anoxic water column sample Cariaco
Rhizosphaeroidea (8 sequences)
Centrocubidae (3 sequences)
Excentroconchidae (8 sequences)
Env 3 (3 sequences)
Spongopyloidea & L1 (19 sequences)
Actinommoidea (5 sequences)
Stylodictyoidea (8 sequences)
Env 4 (16 sequences)
Pylonioidea (22 sequences)
A
B
C
D
E
F
G
H
I
L
K
J
M
I
II
III
IV
67
88
76
80
74
82
62
65
81
85
85
72
>= 90% BS
>= 99% BS
: 25% of original length
0.06

content of Env1 also reveals comparable results to the other clades, where Env1 has an average GC content of 50.1% (±1.36% sd) and the average of all the sequences analyzed is of 48.9% (±1.61% sd) (Supplementary Material Fig. S5). Yet Env1 has the highest average GC content among its relatives in lineage I and decreases towards more distal groups, finding a similar pattern in lineage IV.

Discussion

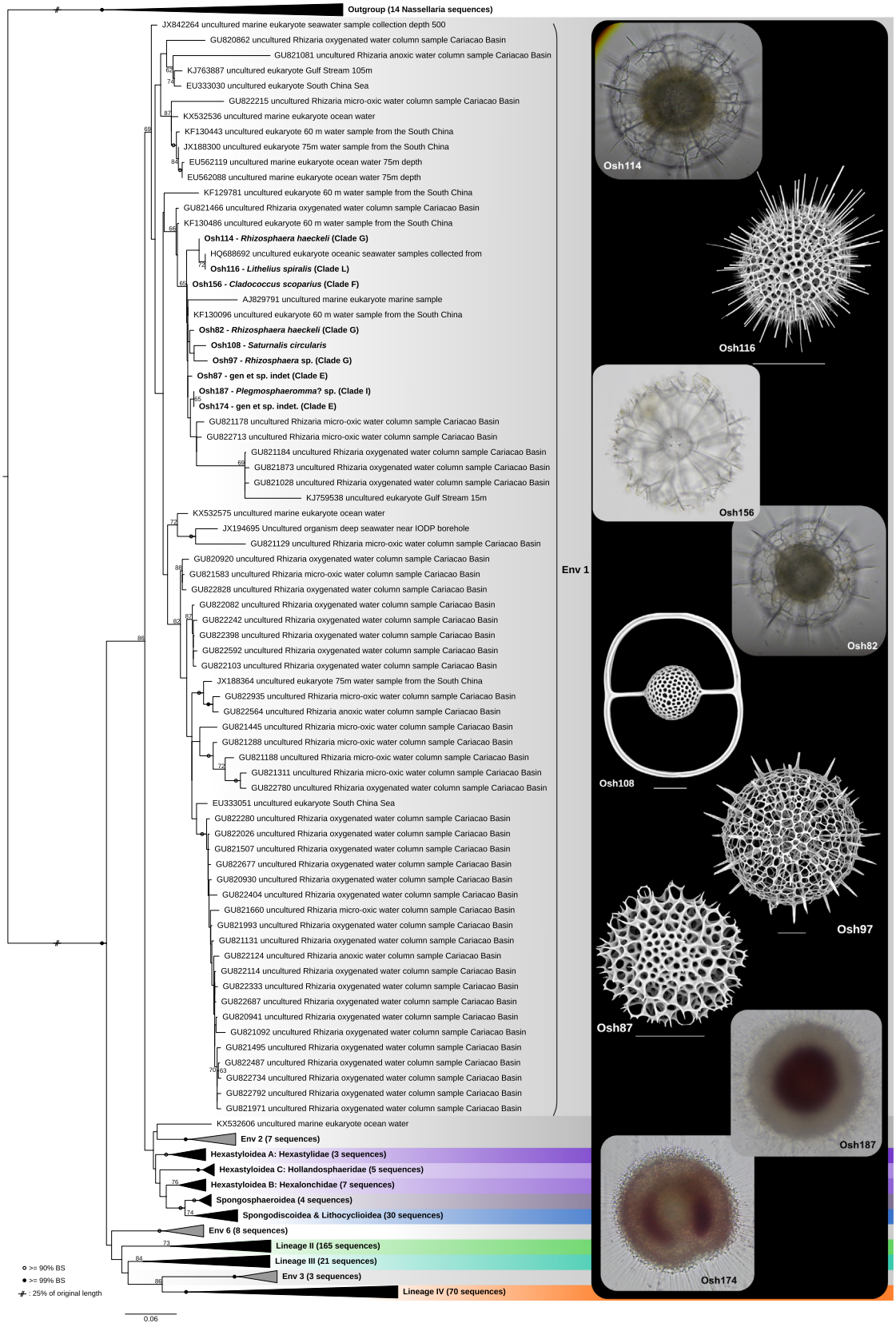
Morpho-molecular Classification of Spumellaria

Our study shows that the symmetry of the skeleton and the overall morphology are key features describing spumellarian phylogeny at higher rank classifications, as for other radiolarian groups such as Acantharia (Decelle et al. 2012b) or Nassellaria (Sandin et al. 2019). Central structure patterns are also well defined between clades helping to discern similar symmetric trends in distantly related clades (e.g. flat and radial symmetry in the super-families Spongodiscoidea, Spongopyloidea and Stylodictyoidea). These results contrast the recognized importance of internal structures in previous morphology-based radiolarian classifications (Afanasieva et al. 2005; De Wever et al. 2001) in which the initial spicular system was the key element for taxonomic delimitation and characterization. Yet, overall shape has always played a significant role in the further description of different taxonomic groups. We therefore establish the symmetry and the overall morphology of the skeleton of Spumellaria as main morphological markers, followed by central structure patterns at finer taxonomic ranks.

The morpho-molecular framework established herein understands the Spumellaria as polycystines radiolarians with concentric structure and a spherical or radial symmetry. This concept includes some living groups classified under the order Entactinaria in De Wever et al. (2001) such as Hexastyloidea that were already suggested to be included in Spumellaria (Ishitani et al. 2012; Matsuzaki et al. 2015; Yuasa et al. 2009). Our results also confirm the inclusion of the Entactinaria super-families Centrocuboidea and Rhizosphaeroidea within Spumellaria. The Radiolaria orders Spumellaria and Entactinaria were traditionally separated based on the absence or presence of internal structures respectively (De Wever et al. 2001). Here we show that the Entactinaria order in the sense of De Wever et al. (2001) is scattered into clades A, B, C (from lineage I) G, H and I (from lineage III). Recent advances in imaging techniques have suggested the Spumellaria as a sub-order within the Entactinaria based on findings of homologous structures with the initial spicular system in Spumellaria among primitive radiolarian forms (Kachovich et al. 2019). Despite our results agreeing with Kachovich et al. (2019) in the taxonomic inclusion of both orders, we consider Entactinaria as a polyphyletic group within the Rhizaria as shown in Nakamura et al. (2020). The prevalence of the overall shape at higher taxonomic levels would therefore place the Spumellaria as the name of polycystines radiolarians with concentric structure and a spherical or radial symmetry, in contrast to Kachovich et al. (2019). The formal inclusion of most entactinarian families with extant Spumellaria provides a direct and documented link from the Cambrian to current ecosystems, improving the connection for paleoenvironmental reconstruction analyses. In addition to our conclusions, there are no genetic evidences of a purely Entactinaria clade so far and at this point we cannot exclude that other groups sharing these symmetric patterns might also be emended within Spumellaria, as for the Centrocubidae and Rhizosphaeridae families. Despite the demonstrated polyphyletic nature of Entactinaria, different appearance and extinction ranges along the fossil record makes it a valuable taxon for biostratigraphic studies.

Extant Spumellaria included in our study can be divided in 13 morpho-molecular clades, grouped in four main evolutionary lineages based on the combi-

Figure 6. Molecular phylogeny of environmental sequences associated to Spumellaria. The tree was obtained by using 1000 rapid bootstrap RAxML GTR + γ model of sequence evolution. Black circles at nodes indicate BS $\geq$ 99% and hollow circles indicates BS $\geq$ 90%. BS lower than 60 were removed for clarity. Branches with a double barred symbol are fourfold reduced for clarity. Black triangles represent clades which morphology is known (Figs. 1,2). Gray triangles represent environmental clades. Sequence composition of each environmental clade is described in Supplementary Material Table S3.


Outgroup (14 Nassellaria sequences)
JX842264 uncultured marine eukaryote seawater sample collection depth 500
GU820862 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821081 uncultured Rhizaria anoxic water column sample Cariacao Basin
KJ763887 uncultured eukaryote Gulf Stream 105m
EU333030 uncultured eukaryote South China Sea
GU822215 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
KX532536 uncultured marine eukaryote ocean water
KF130443 uncultured eukaryote 60 m water sample from the South China
JX188300 uncultured eukaryote 75m water sample from the South China
EU562119 uncultured marine eukaryote ocean water 75m depth
EU562088 uncultured marine eukaryote ocean water 75m depth
KF129781 uncultured eukaryote 60 m water sample from the South China
GU821466 uncultured Rhizaria oxygenated water column sample Cariacao Basin
KF130486 uncultured eukaryote 60 m water sample from the South China
Osh114 - Rhizosphaera haeckeli (Clade G)
HQ688692 uncultured eukaryote oceanic seawater samples collected from
Osh116 - Lithelius spiralis (Clade L)
Osh156 - Cladococcus scoparius (Clade F)
AJ829791 uncultured marine eukaryote marine sample
KF130096 uncultured eukaryote 60 m water sample from the South China
Osh82 - Rhizosphaera haeckeli (Clade G)
Osh108 - Saturnalis circularis
Osh97 - Rhizosphaera sp. (Clade G)
Osh87 - gen et sp. indet (Clade E)
Osh187 - Plegmosphaeromma? sp. (Clade I)
Osh174 - gen et sp. indet. (Clade E)
GU821178 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU822713 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU821184 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821873 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821028 uncultured Rhizaria oxygenated water column sample Cariacao Basin
KJ759538 uncultured eukaryote Gulf Stream 15m
KX532575 uncultured marine eukaryote ocean water
JX194695 Uncultured organism deep seawater near IODP borehole
GU821129 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU820920 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821583 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU822828 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822082 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822242 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822398 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822592 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822103 uncultured Rhizaria oxygenated water column sample Cariacao Basin
JX188364 uncultured eukaryote 75m water sample from the South China
GU822935 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU822564 uncultured Rhizaria anoxic water column sample Cariacao Basin
GU821445 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU821288 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU821188 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU821311 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU822780 uncultured Rhizaria oxygenated water column sample Cariacao Basin
EU333051 uncultured eukaryote South China Sea
GU822280 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822026 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821507 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822677 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU820930 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822404 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821660 uncultured Rhizaria micro-oxic water column sample Cariacao Basin
GU821993 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821131 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822124 uncultured Rhizaria anoxic water column sample Cariacao Basin
GU822114 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822333 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822687 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU820941 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821092 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821495 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822487 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822734 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU822792 uncultured Rhizaria oxygenated water column sample Cariacao Basin
GU821971 uncultured Rhizaria oxygenated water column sample Cariacao Basin
KX532606 uncultured marine eukaryote ocean water
Env 2 (7 sequences)
Hexastyloidea A: Hexastylidae (3 sequences)
Hexastyloidea C: Hollandosphaeridae (5 sequences)
Hexastyloidea B: Hexalonchidae (7 sequences)
Spongosphaeroidea (4 sequences)
Spongodiscoidea & Lithocyclioidea (30 sequences)
Env 6 (8 sequences)
Lineage II (165 sequences)
Lineage III (21 sequences)
Env 3 (3 sequences)
Lineage IV (70 sequences)
Env 1
>= 90% BS
>= 99% BS
: 25% of original length
0.06
Osh114
Osh116
Osh156
Osh82
Osh108
Osh97
Osh87
Osh187
Osh174

nation of phylogenetic clustering support, morphological features and molecular dating. Our results confirm the monophyly of some families (e.g., Liosphaeroidea, Rhizosphaeridae, Actinommoidea or Pylonioidea) while others such as Hexastyloidea found in clades A, B and C or Spongodiscoidea previously gathering morphologies from clades E and L are polyphyletic (Matsuzaki et al. 2015). Different families included within Hexastyloidea have been classified as independent super-families in previous morphology-based taxonomic studies (Afanasieva et al. 2005; De Wever et al. 2001), as seen in the paraphyletic relationships of our phylogenetic results. For instance, flat spumellarians have been classified all together on the basis of their flat symmetry, yet the previously stated polyphyletic nature of this group is confirmed by our results, due to differences in the test and the overall morphology of the skeleton (De Wever et al. 2001). Despite our efforts, some uncertainties regarding specific families such as Heliodiscidae, Ethmosphaeridae or Saturnalidae remain yet to be genetically described. This could be addressed by performing further single celled morpho-molecular characterization analyses on targeted specimens from a broad variety of environments (e.g. deep water). Although these families are represented by very few species and have similar symmetric and central structure patterns to other families presented herein; they could be therefore emended within one of the described clades.

Diversity of Spumellaria

The great environmental diversity found at early diverging positions (i.e.; Env1, Env3, Env6 from Fig. 6) prevents a comprehensive exploration of the extant spumellarian morpho-molecular diversity and challenges the reliability of such molecular clades. However, our analyses allowed us to hypothesizing different morphologies for such environmental clades, precisely Env1. Various lines of evidences led us to conclude the existence of two different and phylogenetically distant 18S rDNA gene copies within the same specimen:

(i) The notably high ratio of chimeric sequences obtained from each single cell analyzed (~13–25%) that would reflect the presence of different rDNA genes amplified with the same Spumellaria-specific primer sets from a single cell/specimen.

(ii) The common relatively long branches within clade Env1 (Fig. 7) traducing the decreased accuracy and quality of sequences obtained by the sequencing of a mix of rDNA gene copies.

(iii) Distantly related morpho-species unexpectedly clustered in the same phylogenetic clade Env1 (Fig. 7), contrasting with the globally well supported morpho-molecular clades in Spumellaria.

(iv) The presence of the same single cell specimen in 2 different and highly supported clades (Osh174 in Env1 and E2), when different amplification and sequencing runs have been carried out.

Further analyses based on the entropy and GC content of sequences clustering in Env1 showed similar trends compare to the rest of spumellarian sequences, disregarding the possibility of a pseudogene (e.g. Balakirev et al. 2014). Intra-genomic or intra-specific polymorphism may also be excluded as possible causes since phylogenetic distances between families and superfamilies are of grater magnitude. Given the regular occurrence of symbiotic relationships within the SAR supergroup (Bjorbækmo et al. 2020; Stoecker et al. 2009, 2016), the most likely explanation left for clade Env1 is the existence of a naked symbiotic group of Spumellaria developing within another, skeleton-bearing Spumellaria. The exact symbiotic nature (parasitic, mutualistic or commensalistic) of such group remains beyond the scope of this study and cannot be inferred from our data. It is worth mentioning that the possibility of environmental DNA contamination cannot be excluded. All 9 specimens come from different samples within the water column (from surface down to 2000 m deep; Supplementary Material Table S1) yet belonging to

Figure 7. Detailed molecular phylogeny of the environmental clade Env1, showing the different morphotypes within the same phylogenetic clade. In brackets after the species name is indicated the expected clade based in the morpho-molecular framework established herein. RAxML bootstrap values (1000 replicates, BS) are shown at the nodes. Black circles indicate BS $\geq$ 99% and hollow circles indicates BS $\geq$ 90%. BS lower than 60 were removed for clarity. Branches with a double barred symbol are fourfold reduced for clarity.

the same water mass. Further imaging analysis on live or fixed specimens (e.g. fluorescence in situ hybridization) could help settle this hypothesis by visualizing two cells/nucleus within the same skeleton or detecting skeleton-less Spumellaria. Similar associations have been already identified between benthic bathyal Foraminifera and dead tests of Xenophyophorea (Foraminifera) (Hughes and Gooday 2004; Pawlowski et al. 2003). To our knowledge the few molecular data available for marine Heliozoa have appeared to be polyphyletic (Burki et al. 2016; Cavalier-Smith et al. 2015), as for freshwater heliozoans (Nikolaev et al. 2004). Indeed, some groups of Heliozoa, such as Gymnosphaeridae, have been moved to the Retaria lineage (precisely within Radiolaria) based on ultrastructural features of the axopodial complex (Yabuki et al. 2012). Therefore, one cannot rule out that some of these basal environmental clades are represented by heliozoan-like organisms and some of them (clade Env1) developed symbiosis within other Spumellaria.

The morpho-molecular framework of Spumellaria established in this study allows an accurate phylogenetic placement of environmental sequences for which it would not be possible to infer their taxonomic position otherwise. Most of the environmental (i.e. not morphologically described) sequences placed in our reference morpho-molecular framework are highly related to Liosphaeroidea, coming primarily from deep environments (1500–2500 m) (Edgcomb et al. 2011; Lie et al. 2014). Another group displaying a large environmental diversity is the family Excentroconchidae, mainly represented by sequences coming from anoxic environments (Edgcomb et al. 2011; Kim et al. 2012). These results agree with the data obtained by morphological observations from the fossil record (De Wever et al. 2001) or from plankton and sediments traps and surface sediment materials (Boltovskoy et al. 2010; Boltovskoy and Correa 2016), in which both the Liosphaeroidea and the Excentroconchidae are among the most abundant groups. Representatives of these families are characterized by a thin and fragile skeleton with a dense protoplasm. As previously suggested (Not et al. 2007), cell breakage may also contribute to the outstanding amount of sequences found among these taxonomic groups compared to other clades.

Evolutionary History of Spumellaria

Our molecular clock results dated the first diversification of Spumellaria at ~ 515 (HPD: 659–382) Ma, in agreement with recent findings of the oldest radiolarian fossils with spherical forms, taxonomically classified in the extinct spumellarian genus *Paraantygopora* (early Cambrian, Series 2, ca. 521–509 Ma; Zhang and Feng 2019). This early diversification separated the two contrasting lineages I and IV, followed soon afterwards by the branching of linages II and III, during the so-called Great Ordovician Biodiversification Event (Noble and Danelian 2004; Servais et al. 2016). At this moment many different symmetrically spherical radiolarians appeared in the fossil record (Aitchison et al. 2017; Noble et al. 2017). Yet, poor radiolarian-bearing rocks (Aitchison et al. 2017), different topologies of the different marker genes and the short phylogenetic distance at these nodes in the concatenated phylogenetic analysis make relationships between lineages obscure. From this period until the end of the Paleozoic the diversification of extant lineages is low and basically restricted to the end of the Carboniferous (358.9 – 298.9 Ma) when lineages III and IV start diversifying.

All over the Paleozoic, extinct specimens previously attributed to Entactinaria show a great diversity that mostly became extinct towards the beginning of the Mesozoic (De Wever et al. 2003, 2006). Some representatives of this group have been tentatively considered as Nassellaria in latest classifications and later suggested as ancestors of this order due to overall morphology and molecular clock results (Noble et al. 2017; Sandin et al. 2019). The morpho-molecular framework here proposed allows the attribution of some radiolarian extinct groups with spherical or radial symmetry as possible ancestors of the extant Spumellaria, that can be reflected in the long branches of clades belonging to lineages II, III and IV (Fig. 1). These phylogenetic patterns may be interpreted as a bottleneck effect, in which many groups went extinct before they had a chance to radiate.

At the beginning of the Mesozoic the first extant representatives appear with the diversification of lineage II (Liosphaeroidea) and Hexastyloidea. Although, the polyphyletic nature of Hexastyloidea shows a later diversification within the clades (A, B

and C), explaining why the Triassic (251.9–201.3 Ma) genera does not resemble that of the Cenozoic (66–0 Ma) (O'Dogherty et al. 2011). During the Triassic, lineage III also diversifies and all clades from lineage IV are already separated from each other, yet their diversifications are not happening until the end of the Mesozoic. Similar diversification patterns at the beginning of the Mesozoic have been reported in living Nassellaria inferred from the combination of the molecular clock and the direct observation of the fossil record (De Wever et al. 2006; Sandin et al. 2019); in which ancient forms, presumably extinct during the end-Permian extinction (Aitchison et al. 2017; De Wever et al. 2006; O'Dogherty et al. 2011; Sepkoski 1981; Takahashi et al. 2009) led to the diversification of extant groups.

The second part of the Mesozoic is characterized by a stepped diversification of Spumellaria, as already described for Nassellaria (Sandin et al. 2019) and Foraminifera (Hart et al. 2003; Leckie et al. 2002). This phenomenon is classically explained by the onset of a global oceanic anoxia during the Jurassic (Jenkyns 1998) and a series of Oceanic Anoxic Events during the Cretaceous (Erbacher et al. 1996; Jenkyns 2010; Kemp and Izumi 2014; Schalanger and Jenkyns 1976; Yilmaz et al. 2012). In contrast to Nassellaria diversification, towards the end of the Jurassic (201.4–145 Ma) and beginning of Cretaceous (145–66 Ma), there is a big increase in Spumellaria diversification where most of the clades suddenly diversified. This trend is also reported in the fossil record (De Wever et al. 2003; Kiessling 2002). Such differences between diversification patterns of Spumellaria regarding other radiolarian groups is a question already raised in recent studies (Kachovich et al. 2019). Both extant Nassellaria and Spumellaria share similar environmental preferences (Suzuki and Not 2015). Yet Spumellaria tend to possess a larger protoplasmic volume (Des Combes and Abelmann 2009; Takahashi 1982). There is also a drastic diversification of flat spumellarians, a feature that increases the surface/volume ratio of a given organisms. Probably such morphological adaptations provided an advantage to thrive during low nutrient availability periods, as known to be case during the Jurassic (Cárdenas and Harries 2010). Once the conditions became more favorable heterotrophic Spumellaria started a sustained diversification that followed throughout the Cenozoic (66–0 Ma). Similar diversification patterns linked to climatic oscillations over

Table 1. Primer sequences and temperatures used for DNA amplification and sequencing.

Targeted gene	Primer	Specificity	Sequence 5′-3′	Direction	Tm °C	Reference
18S (1st part)	SA	Eukaryotes	AAC CTG GTT GAT CCT GCC AGT	Forward	56	Medlin et al., 1988
	S879	Radiolaria	CCA ACT GTC CCT ATC AAT CAT	Reverse		Decelle et al., 2012b
18S (2nd part)	S32_TASN	Radiolaria	CCA GCT CCA ATA GCG TAT RC	Forward	57	Ishitani et al. 2012
	V9R	Eukaryotes	CCT TCY GCA GGT TCA CCT AC	Reverse		Romac (unpub.)
28S (D1 + D2)	28S Rad2	Radiolaria	TAA GCG GAG GAA AAG AAA	Forward	52	Ando et al., 2009
	ITSa4R	Radiolaria	TCA CCA TCT TTC GGG TCC CGG CAT	Reverse		This study

the Cenozoic were found in Nassellaria (Sandin et al. 2019), Foraminifera (Hallock et al. 1991) and coccolithophores (Bown et al. 2004).

Methods

Sampling and single cell isolation: Plankton samples were collected: 1- off Sendai (11 samples: 38° 0' 28.8'' N, 142° 0' 28.8'' E), 2- Sesoko (8 samples: 26° 48' 43.2'' N, 73° 58' 58.8'' E), 3- the Southwest Islands, South of Japan (2 samples: 28° 14' 49.2'' N, 129° 5' 27.6'' E) 4- in the bay of Villefranche-sur-Mer (38 samples: 43° 40' 51.6'' N, 7° 19' 40.8'' E) and 5- in the western Mediterranean Sea (9 samples: MOOSE-GE cruise) by net tows (20, 64 or 200 μm), Vertical Multiple-opening Plankton Sampler (VMPS) or Bongo net (24-200 μm). Samples related metadata can be found in the RENKAN database (http://abims.sb-roscoff.fr/renkan). Spumellarian specimens were individually handpicked using Pasteur pipettes from the plankton tow sample and transferred 3 to 4 times into 0.2 μm filtered seawater to allow self-cleaning from debris, particles attached to the cell or prey digestion. Images of live specimens were taken under an inverted microscope and thereafter transfer into 1.5 ml Eppendorf tubes containing 50 μl of molecular grade absolute ethanol and stored at -20 °C until DNA extraction.

DNA extraction, amplification and sequencing: DNA was extracted using the MasterPure Complete DNA and RNA Purification Kit (Epicentre) following manufacturer's instructions. Both 18S rDNA and partial 28S rDNA genes (D1 and D2 regions) were amplified by Polymerase Chain Reaction (PCR) using Radiolaria and Spumellaria specific and general primers (Table 1). For further details about rDNA amplification see: dx.doi.org/10.17504/protocols.io.xwvfpe6. PCR amplicons were visualized on 1% agarose gel stained with ethidium bromide. Positive reactions were purified using the Nucleospin Gel and PCR Clean up kit (Macherey Nagel), following manufacturer's instructions and sent to Macrogen Europe for sequencing.

Scanning Electron Microscopy (SEM): After DNA extraction, spumellarian skeletons were recovered from the eluted pellet and handpicked under binoculars or inverted microscope. After cleaning and preparing the skeletons (detailed protocol in dx.doi.org/10.17504/protocols.io.ug9etz6) images were taken with a FEI Phenom tabletop Scanning Electron Microscope (FEI technologies).

Single cell morphological identification: Spumellaria specimens were identified at the species level, referring to pictures of holotypes, through observation of live images and analysis of the skeleton by scanning electron microscopy when available (see Supplementary Material Table S1 for taxonomic authority of specimens included in our study). Further details in taxonomic assignment of the specimens can be found in the Methods of Sandin et. al. 2019.

Phylogenetic analyses: After sequencing, forward and reverse sequences were checked and assembled using ChromasPro software version 2.1.4 (2017). Sequences were compared to the GenBank reference database (GenBank) using the *BLAST search* tool integrated in ChromasPro to discriminate radiolarian sequences from possible contamination. Presence of chimeras, at a frequency as high as 13-25%, was detected by mothur v.1.39.3 (Schloss et al. 2009) against previously available reference sequences of Spumellaria. Sequences not detected as chimeras, were in turn included in our reference database for future phylogenetic reconstruction and chimeric analysis.

Two different datasets for each genetic marker (18S rDNA gene and partial 28S rDNA gene) were aligned separately using MAFFT v7.395 (Katoh and Standley 2013) with a L-INS-i algorithm ('--localpair') and 1000 iterative refinement cycles for high accuracy. Each alignment was manually checked in SeaView version 4.6.1 (Gouy et al. 2010) and trimmed automatically using trimal (Capella-Gutiérrez et al. 2009) with a 30% gap threshold. For both genes, the 18S rDNA (133 taxa, 1789 positions) and the 28S rDNA (55 taxa, 746 positions), phylogenetic analyses were performed independently. The best nucleotide substitution model was chosen following the corrected Akaike Information Criterion (AICc) using the *modelTest* function implemented in the R version 3.6.0 (R Core Team 2014) package *phangorn* version 2.5.5 (Schliep 2011). The obtained model (General Time Reversible with Gama distribution and proportion of Invariable sites, GTR + G + I) was applied to each data set in R upon the packages *APE* version 5.3 (Paradis et al. 2004) and *phangorn* version 2.5.5 (Schliep 2011). A Maximum Likelihood (ML) method (Felsenstein 1981) with 10 000 replicates of bootstrap (Felsenstein 1985) was performed to infer phylogenies.

Despite specific discrepancies in the topology of the two analyses (Supplementary Material Fig. S1), the two genes were concatenated in order to increase taxonomic coverage and improve phylogenetic resolution. A final data set containing 133 taxa and 2535 positions was used to infer phylogenies following the previous methodology. Fourteen sequences of Nassellaria were assembled to form the outgroup as seen in previous classifications to be the sister clade of Spumellaria (e.g., Cavalier-Smith et al. 2018; Krabberød et al. 2011). The best model obtained was GTR + G + I, with four intervals of the discrete gamma distribution, and a ML method with 100 000 bootstraps. In parallel, a Bayesian analysis was performed using BEAST version 1.8.4 (Drummond et al. 2012) with the same model parameters over 100 million generations sampled every 1000 states, to check the consistency of the topology and to calculate posterior probabilities. The final tree was visualized and edited with FigTree version 1.4.3 (Rambaut 2016). All sequences obtained in this study and used for the phylogeny were submitted in GenBank under the accession numbers: MT670437 - MT670549.

Morphological observations performed with light and scanning electron microscopy (Supplementary Material Fig. S2) allowed assigning these sequences to nine super-families (Actinommoidea, Hexastyloidea, Liosphaeroidea, Lithocyclioidea, Pylonioidea, Spongodiscoidea, Spongopyloidea, Spongosphaeroidea and Stylodictyoidea), two super-families considered to belong to Entactinaria (Centrocuboidea and Rhizosphaeridae) and two unresolved families (Litheliidae and Sponguridae).

Molecular clock analyses: Molecular clock estimates were performed according to Sandin et al. (2019). Nine nodes were chosen to carry out the calibration. The selection of these nodes is explained below from the oldest to the newest calibration age, and given the name of the node for the taxa they cover:

1. Root: The calibration for the root of the tree corresponds to the hypothesized last common ancestor between Nassellaria and Spumellaria (De Wever et al. 2001). A uniform distribution (U) with a minimum bound of 300 million years ago (Ma) and a maximum of 800 Ma (U[300, 800]) was set to allow uncertainty in the diversification of both groups and to establish a threshold restricting the range of solutions for the entire tree.

2. Nassellaria N(410, 20): The outgroup of the phylogeny is calibrated based in a consensus between the first appearance of nassellarian-like fossils in the Upper Devonian (ca. 372.2 Ma) in the fossil record (Cheng 1986) and the first diversification of Nassellaria dated with previous analysis of the molecular clock (ca. 423 Ma; 95% HPD: 500–342 Ma; Sandin et al. 2019). Therefore, the node was nor-

mally distributed (N) with a mean of 410 and a standard deviation of 20: N(410, 20).

3. Spumellaria U[700, 200]: Recent studies have found the oldest spumellarian representatives in the early middle Cambrian (Zhang and Feng 2019). Yet, De Wever et al. (2001) have argued that the initial spicular system may be subjected to preservation bias. Since many spumellarians and entactinarians (presence of the initial spicular system or not, respectively) are superficially similar and it is not possible to distinguish from one another (Suzuki and Oba 2015), a uniform distribution was set to allow uncertainty in the diversification of Spumellaria.

4. Hexastyloidea N(242, 10): The family Hexastylidae is the first representative of the superfamily Hexastyloidea and has its first appearance in the fossil record in the Middle Triassic (Late Anisian: ca. 246.8–241.5 Ma; O'Dogherty et al. 2011).

5. Liosphaeroidea N(233, 20): The Liosphaeroidea seems to have appeared in the Triassic (De Wever et al. 2001; Pessagno and Blome 1980). Although it is not sure whether there is a continuity between members from the Mesozoic and Cenozoic (Matsuzaki et al. 2015).

6. Actinommidae N(170, 20): The family Actinommidae appears for the first time in the fossil record in the Middle Jurassic (Aalenian: ca. 174.2–170.3 Ma; O'Dogherty et al. 2011) yet some morphologies from the Triassic resemblance this Superfamily (De Wever et al. 2001).

7. Rhizosphaeroidea N(148, 10): The Rhizosphaeridae appeared during the late Jurassic (Tithonian: ca. 145–152.1 Ma) in the fossil record (Afanasieva and Amon 2006; De Wever et al. 2001; Dumitrica 2017; Petrushevskaya 1975).

8. Pylonioidea N(97, 10): The family Larnacillidae are the first representatives of the superfamily Pylonioidea appearing at the beginning of the Late Cretaceous (Cenomanian: ca. 100.5–93.9 Ma; Afanasieva and Amon 2006; De Wever et al. 2001, 2003).

9. Coccodiscoidea N(45, 10): The family Coccodiscoidea appears for the first time in the fossil record in the Early Eocene (Afanasieva and Amon 2006; De Wever et al. 2001). This family corresponds to clade E1 in Fig. 1.

Post hoc analyses: Two different analyses were performed a posteriori: a diversification of taxa over time (Lineages Through Time: LTT) and an ancestral state reconstruction. The former analysis was carried out with the *ltt.plot* function implemented in the package *APE* (Paradis et al. 2004) upon the tree obtained by the molecular dating analyses after removing the outgroup. The second analysis used the resulting phylogenetic tree to infer the evolution of morphological characters. A numerical value was assigned to each state of a character trait, being 0 for the considered ancestral state, and one to three for the presumed divergence state. In total six traits were considered (Supplementary Material Table S2): the skeleton shape, the symmetry, the central structure, the internal cavity (outside of the central structure), the number of spines arisen from the central structure and the number of distinctive cortical shells. Once the character matrix was established a parsimony ancestral state reconstruction was performed to every character independently in Mesquite version 3.2 (Maddison and Maddison 2017).

Environmental sequences: Each of the reference 18S rDNA and partial 28S rDNA spumellarian sequences available in our study was compared with publicly available environmental sequences in GenBank (NCBI) using BLAST (as of May 2019). It allowed estimating the environmental genetic diversity of Spumellaria and to assess the genetic coverage of our phylogenetic tree. A total of 1171 and six environmental sequences affiliated to Spumellaria were retrieved for the genes 18S rDNA and 28S rDNA respectively. Out of the six 28S rDNA sequences, four were also found in the 18S rDNA survey covering the full rDNA operon. These four sequences along with the eight sequences already included in the phylogeny (two in clade H and eight in clade I) were therefore removed to avoid duplicates. Final dataset containing 1165 environmental sequences were aligned against the reference alignment used for the phylogenetic analysis using MAFFT v7.395 (Katoh and Standley 2013) with the "--add" option and placed in our reference phylogenetic tree using the pplacer software (Matsen et al. 2010). A RAxML (GTR + G + I) tree was built for the placement of the sequences with a rapid bootstrap analysis and search for best-scoring ML tree and 1000 bootstraps.

In order to discriminate pseudo-genes among environmental clades due to a high frequency of chimeras encountered and long phylogenetic branches, two preliminary analyses were performed: a Shannon entropy analysis and an estimation of the GC content. Shannon entropy was calculated for every position of the resulting alignment matrix, without considering gaps due to differences in length of environmental sequences (script available on https://github.com/MiguelMSandin/DNA-alignment-entropy). The GC content was estimated for every sequence independently as follows: (G + C)/(G + C + A + T), where G, C, A and T are the number of guanines, cytosines, adenines and thymines respectively. Analysis were performed in R version 3.6.0 (R Core Team 2014) and plotted with the package *ggplot2* version 3.2.1 (Wickham 2016).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the IMPEKAB ANR 15-CE02-0011 grant and the Brittany Region ARED C16 1520A01, the Japan Society for Promotion of Science KAKENHI Grant No. K16K0-74750 for N. Suzuki and "the Cooperative Research Project with the Japan Science and Technology Agency (JST) and Centre National de la Recherche Scientifique (CNRS, France) "Morpho-molecular Diversity Assessment of Ecologically, Evolutionary, and Geologically Relevant Marine Plankton (Radiolaria)". We are grateful to the CNRS-Sorbonne University ABiMS bioinformatics platform (http://abims.sb-roscoff.fr) for providing computational resources. The authors are grateful to the MOOSE observation national network (funded by CNRS-INSU and Research Infrastructure ILICO) which sustain the annual ship-based hydrographic sections in the northwestern Mediterranean Sea (MOOSE-GE), as well as John Dolan for hosting us multiple times at

the Laboratoire d'Océanographie of Villefranche sur Mer. We are greatly thankful to Cedric Berney for the phylogenetic advice and the valuable help on the interpretation of the "symbiotic" clade, as well as Vasily Zlatogursky for his contributions and feedback on the heliozoan-like organism.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.protis.2021.125806.

References

Abelmann A, Nimmergut A (2005) Radiolarians in the Sea of Okhotsk and their ecological implication for paleoenvironmental reconstructions. Deep Sea Res. Part II Top Stud Oceanogr **52**:2302–e2331

Adl SM, Bass D, Lane CE, Lukeš J, Schoch CL, Smirnov A, Agatha S, Berney C, Brown MW, Burki F, Cárdenas P, Čepička I, Chistyakova L, del Campo J, Dunthorn M, Edvardsen B, Eglit Y, Guillou L, Hampl V, Heiss AA, Hoppenrath M, James TY, Karnkowska A, Karpov S, Kim E, Kolisko M, Kudryavtsev A, Lahr DJ, Lara E, Le Gall L, Lynn DH, Mann DG, Massana R, Mitchell EA, Morrow C, Park JS, Pawlowski JW, Powell MJ, Richter DJ, Rueckert S, Shadwick L, Shimano S, Spiegel FW, Torruella G, Youssef N, Zlatogursky V, Zhang Q (2019) Revisions to the classification, nomenclature, and diversity of eukaryotes. J Eukaryot Microbiol **66**:4–119

Afanasieva MS, Amon EO (2006) Biotic crises and stages of radiolarian evolution in the Phanerozoic. Paleontol J **40**: S453–S467

Afanasieva MS, Amon EO, Agarkov YU, Boltovskoy BS (2005) Radiolarians in the geological record. Paleontol J **39 (Suppl. 3)**:135–392

Aita Y, Suzuki N, Ogane K, Sakai T, Lazarus D, Young J, Tanimura Y (2009) Haeckel Radiolaria Collection and the H. M.S. Challenger plankton collections. Nat Mus Natl Sci Monogr **40**:35–45

Aitchison J, **Suzuki N**, **Caridroit M**, **Danelian T**, **Noble PJ** (2017) Paleozoic Radiolarian Biostratigraphy. In Danelian T, Caridroit M, Noble P, Aitchison JC (eds) Catalogue of Paleozoic Radiolarian Genera. Geodiversitas **31**:503–531

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

Balakirev ES, Chechetkin VR, Lobzin VV, Ayala FJ (2014) Computational methods of identification of pseudogenes based on functionality: entropy and GC content. Methods Mol Biol **1167**:41–62

Biard T, Bigeard E, Audic S, Poulain J, Stemmann L, Not F (2017) Biogeography and diversity of Collodaria (Radiolaria) in the global ocean. ISME J **11**:1331–1344

Bjorbækmo MFM, Evenstad A, Røsæg LL, Krabberød AK, Logares R (2020) The planktonic protist interactome: where do we stand after a century of research? ISME J **14**:544–559

Boltovskoy D, Correa N (2016) Biogeography of Radiolaria Polycystina (Protista) in the World Ocean. Progr Oceanogr **149**:82–105

Boltovskoy D, Kling SA, Takahashi K, Bjørklund K (2010) World atlas of distribution of recent Polycystina (Radiolaria). Palaeontol Electron **13**:230

Bown PR, Lees JA, Young JR (2004) Calcareous Nannoplankton Evolution and Diversity through Time. In Thierstein HR, Young JR (eds) Coccolithophores. Springer; Berlin, Heidelberg, pp 481–508

Burki F, **Kaplan M**, **Tikhonenkov DV**, **Zlatogursky V**, **Minh BQ**, **Radaykina LV**, **Smirnov A**, **Mylnikov AP**, **Keeling PJ** (2016) Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. Proc R Soc B Biol Sci:20152802

Burki F, Keeling PJ (2014) Rhizaria. Curr Biol **24**:103–107

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics **25**:1972–1973

Cárdenas AL, Harries PJ (2010) Effect of nutrient availability on marine origination rates throughout the Phanerozoic eon. Nat Geosci **3**:430–434

Cavalier-Smith T, Chao EE, Lewis R (2015) Multiple origins of Heliozoa from flagellate ancestors: New cryptist subphylum Corbihelia, superclass Corbistoma, and monophyly of Haptista, Cryptista, Hacrobia and Chromista. Mol Phylogenet Evol **93**:331–362

Cavalier-Smith T, Chao EE, Lewis R (2018) Multigene phylogeny and cell evolution of chromist infrakingdom Rhizaria: contrasting cell organisation of sister phyla Cercozoa and Retaria. Protoplasma **255**:1517–1574

Cheng YN (1986) Taxonomic Studies on Upper Paleozoic Radiolaria. Spec Publ National Museum of Natural Science. Taiwan **1**:311 p

de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney C, Le Bescot N, Probert I, Carmichael M, Poulain J, Romac S, Colin S, Aury JM, Bittner L, Chaffron S, Dunthorn M, Engelen S, Flegontova O, Guidi L, Horák A, Jaillon O, Lima-Mendez G, Lukeš J, Malviya S, Morard R, Mulot M, Scalco E, Siano R, Vincent F, Zingone A, Dimier C,

Picheral Tara Oceans, Searson S, Kandels-Lewis S, Coordinators Tara Oceans, Acinas SG, Bork P, Bowler C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Not F, Ogata H, Pesant S, Raes J, Sieracki ME, Speich S, Stemmann L, Sunagawa S, Weissenbach J, Wincker P, Karsenti E (2015) Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. Science **348**: 1261605

De Wever P, O'Dogherty L, Goričan Š (2006) The plankton turnover at the Permo-Triassic boundary, emphasis on radiolarians. Eclogae Geol Helv **99**:S49–S62

De Wever P, Dumitrica P, Caulet JP, Nigrini C, Caridroit M (2001) Radiolarians in the Sedimentary Record. Gordon & Breach Science Publishers, Amsterdam, 533 p

De Wever P, O'Dogherty L, Caridroit M, Dumitrica P, Guex J, Nigrini C, Caulet JP (2003) Diversity of radiolarian families through time. Bull la Soc Geol Fr **174**:453–469

Decelle J, Martin P, Paborstava K, Pond DW, Tarling G (2013) Diversity, ecology and biogeochemistry of cyst-forming Acantharia (Radiolaria) in the Oceans. PLoS ONE **8(1)**: e53598

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

Decelle J, Probert I, Bittner L, Desdevises Y, Colin S, Cde Vargas, Galí M, Simó R, Not F An original mode of symbiosis in open ocean plankton. Proc Natl Acad Sci USA **109**:18000–18005

Deflandre G (1953) Radiolaires fossiles. In Grassé PP (ed), Traité de Zoologie, Anatomie, Systématique, Biologie. vol. 1, part 2, Masson, Paris, pp 389–436

Deflandre G (1973) Observations et remarques sur les Radiolaires Sphaerellaires du Paléozoïque, à propos d'une nouvelle espèce viséenne, du genre Foremaniella Defl., parfait intermédiare entre les Périaxoplastidiés et les Pylentonémidés. Comptes rendus hebdomadaires des séances de l'Académie des sciences Ser D, Sciences naturelles **276**:1147–1151

Des Combes HJ, Abelmann A (2009) From species abundance to opal input: Simple geometrical models of radiolarian skeletons from the Atlantic sector of the Southern Ocean. Deep Res Part I Oceanogr Res Pap **56**:757–771

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol **29**:1969–1973

Dumitrica P Internal skeletal structures of the superfamily Pyloniacea (Radiolaria), a basis of a new systematics. Rev Española Micropaleontol **XXI**:207–264.

Dumitrica P (1994) *Pyloctostylys* n. gen., a Cretaceous spumellarian radiolarian genus with initial spicule. Rev Micropaleontol **37**:235–244

Dumitrica P (2017) Contribution to the knowledge of the Entactinaria radiolarian family Rhizosphaeridae Haeckel and description of some new genera and species. Rev Micropaleontol **60**:469–491

Dumitrica P, Tekin UK, Bedi Y (2010) Eptingiacea and Saturnaliacea (Radiolaria) from the middle Carnian of Turkey and some late Ladinian to early Norian samples from Oman and Alaska. Paläontol Zeitschr **84**:259–292

Edgcomb V, Orsi W, Bunge J, Jeon S, Christen R, Leslin C, Holder M, Taylor G, Suarez P, Varela R, Epstein S (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. ISME J **5**:1344–1356

Erbacher J, Thurow J, Littke R (1996) Evolution patterns of radiolaria and organic matter variations: A new approach to identify sea-level changes in mid-Cretaceous pelagic environments. Geology **24**:499–502

Felsenstein J (1981) Evolutionary trees from DNA sequences: A maximum likelihood approach. J Mol Evol **17**:368–376

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat **125**:1–15

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**:221–224

Haeckel E (1862) Die Radiolarien (Rhizopoda radiaria): eine Monographie. Reimer G, Berlin, pp 1862–1888

Haeckel E (1882) Entwurf eines Radiolarien-Systems auf Grund von Studien der Challenger-Radiolarien. Jenaische Z Naturwiss **15**:418–472

Haeckel E (1887) Report on the Radiolaria collected by H. M. S Challenger during the years 1873–1876. Zoology **18**:1–1803

Hallock P, Premoli Silva I, Boersma A (1991) Similarities between planktonic and larger foraminiferal evolutionary trends through Paleogene paleoceanographic changes. Palaeogeogr Palaeoclimatol Palaeoecol **83**:49–64

Hart MB, Hylton MD, Oxford MJ, Price GD, Hudson W, Smart CW (2003) The search for the origin of the planktic Foraminifera. J Geol Soc London **160**:341–343

Hollande A, Enjumet M (1960) Cytologie, évolution et systématique des Sphaeroidés (Radioiaires). Arch Mus Nain Hist Nat **7**:1–134

Hughes JA, Gooday AJ (2004) Associations between living benthic foraminifera and dead tests of *Syringammina fragilissima* (Xenophyophorea) in the Darwin Mounds region (NE Atlantic). Deep-Sea Res. Part I - Oceanogr Res Pap **51**:1741–1758

Ishitani Y, Ujiié Y, Cde Vargas, Not F, Takahashi K (2012) Two distinct lineages in the radiolarian Order Spumellaria having different ecological preferences. Deep Res Part II Top Stud Oceanogr **61–64**:172–178

Jenkyns HC (1998) The Early Toarcian Anoxic event: Stratigraphic, sedimentary, and geochrmical evidence. Am J Sci **288**:101–151

Jenkyns HC (2010) Geochemistry of oceanic anoxic events. Geochemistry Geophys Geosystems **11**:1–30

Kachovich S, Sheng J, Aitchison JC (2019) Adding a new dimension to investigations of early radiolarian evolution. Sci Rep **9**:6450

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol **30**:772–780

Kemp DB, Izumi K (2014) Multiproxy geochemical analysis of a Panthalassic margin record of the early Toarcian oceanic anoxic event (Toyora area, Japan). Palaeogeogr Palaeoclimatol Palaeoecol **414**:332–341

Kiessling W (2002) Radiolarian diversity patterns in the latest Jurassic-earliest Cretaceous. Palaeogeogr Palaeoclimatol Palaeoecol **187**:179–206

Kim DY, Countway PD, Yamashita W, Caron DA (2012) A combined sequence-based and fragment-based characterization of microbial eukaryote assemblages provides taxonomic context for the Terminal Restriction Fragment Length Polymorphism (T-RFLP) method. J Microbiol Methods **91**:527–536

Krabberød AK, Brate J, Dolven JK, Ose RF, Klaveness D, Kristensen T, Bjørklund KR, Shalchian-Tabrizi K (2011) Radiolaria divided into Polycystina and Spasmaria in combined 18S and 28S rDNA phylogeny. PLoS ONE **6(8)**: 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. Europ J Protistol **42**:143–153

Leckie RM, Bralower TJ, Cashman R (2002) Oceanic anoxic events and plankton evolution: Biotic response to tectonic forcing during the mid-Cretaceous. Paleoceanography **17:**13–1–13-29

Lewitus E, Bittner L, Malviya S, Bowler C, Morlon H (2018) Clade-specific diversification dynamics of marine diatoms since the Jurassic. Nat Ecol Evol **2**:1715–1723

Lie AAY, Liu Z, Hu SK, Jones AC, Kim DY, Countway PD, Amaral-Zettler LA, Cary SC, Sherr EB, Sherr BF, Gast RJ, Caron DA (2014) Investigating microbial eukaryotic diversity from a global census: Insights from a comparison of pyrotag and full-length sequences of 18S rRNA genes. Appl Environ Microbiol **80**:4363–4373

Llopis-Monferrer N, Boltovskoy D, Tréguer P, Sandin MM, Not F, Leynaert A (2020) Estimating biogenic silica production of Rhizaria in the global ocean. Global Biogeochem. Cycl **34**: e2019GB006286

Maddison WP, Maddison DR (2017) Mesquite: a modular system for evolutionary analysis. Version **3**:2. http://www.mesquiteproject.org

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

Matsuzaki KM, **Suzuki N**, **Nishi H** (2015) Middle to Upper Pleistocene polycystine radiolarians from Hole 902-C9001C, northwestern Pacific. Paleontol Res **19**(Supp):1–77

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

Nakamura Y, Sandin MM, Suzuki N, Somiya R, Tuji A, Not F (2020) Phylogenetic revision of the order Entactinaria - Paleozoic relict Radiolaria (Rhizaria, SAR). Protist **171**: 125712

Nikolaev SI, Berney C, Fahrn JF, Bolivar I, Polet S, Mylnikov AP, Aleshin VV, Petrov NB, Pawlowski J (2004) The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. Proc Natl Acad Sci USA **101**:8066–8071

Nitsche F, Thomsen HA, Richter DJ (2016) Bridging the gap between morphological species and molecular barcodes – exemplified by loricate choanoflagellates. Europ J Protistol **57**:26–37

Noble P, **Aitchison JC**, **Danelian T**, **Dumitrica P**, **Maletz J**, **Suzuki N**, **Cuvelier J**, **Caridroit M**, **O'Dogherty L** (2017) Taxonomy of Paleozoic Radiolarian Genera. In Danelian, T, Caridroit M, Noble P, Aitchison JC (eds), Catalogue of Paleozoic Radiolarian Genera. Geodiversitas **31**: 419–502

Noble P, Danelian T (2004) Radiolarians. In Webby BD, Droser ML, Paris F, Percival IG (eds) The Great Ordovician Biodiversification Event. Columbia University Press; New York, pp. 97–101

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

O'Dogherty L (1994) Biochronology and paleontology of mid-Cretaceous radiolarians from Northern Apennines (Italy) and Betic Cordillera (Spain). Mémoires de Géologie (Lausanne) **21**:1–415

O'Dogherty L, De Wever P, Goričan Š, Carter ES, Dumitrica P (2011) Stratigraphic ranges of Mesozoic radiolarian families. Palaeoworld **20**:102–115

O'Dogherty L, Carter ES, Dumitrica P, Gori Š, de Wever P, Hungerbühler A, Bandini AN, Takemura A (2009) Catalogue of Mesozoic radiolarian genera. Part 1: Triassic. Geodiversitas **31**:213–270

Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. Bioinformatics **20**:289–290

Pawlowski JAN, Holzmann M, Fahrni J, Richardson SL (2003) Small subunit ribosomal DNA suggests that the Xenophyophorean *Syringammina corbicula* is a Foraminiferan. J Eukaryot Microbiol **50**:483–487

Pernice MC, Giner CR, Logares R, Perera-Bel J, Acinas SG, Duarte CM, Gasol JM, Massana R (2016) Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. ISME J **10**:945–958

Pessagno EA, Blome CD (1980) Upper Triassic and Jurassic. Pantanelliinae from California, Oregon and British Columbia, Micropaleontology **26**:225–273

Petrushevskaya MG (1975) Cenozoic Radiolarians of the Antarctic, Leg 29, DSDP. Texas A & M University, Ocean Drilling Program, College Station, TX, United States. In Kennett JP, Houtz RE, et al. (eds) Initial Reports of the Deep Sea Drilling Project, **29**. U.S. Government Printing Office, Washington, DC, pp 541-576

Probert I, Siano R, Poirier C, Decelle J, Biard T, Tuji A, Suzuki N, Not F (2014) B*randtodinium* gen. nov. and *B. nutricula* comb. nov. (Dinophyceae), A dinoflagellate commonly found In symbiosis with polycystine radiolarians. J Phycol **50**:388–399

R Core Team (2014) R: A Language and Environment for Statistical Computing. doi:10.1007/978-3-540-74686-7

Rambaut A (2016) FigTree version 1.4.3. http://tree.bio.ed.ac.uk/software/figtree/

Sandin MM, Pillet L, Biard T, Poirier C, Bigeard E, Romac S, Suzuki N, Not F (2019) Time calibrated morpho-molecular classification of Nassellaria (Radiolaria). Protist **170**:187–208

Schalanger SO, Jenkyns HC (1976) Cretaceous oceanic anoxic events: causes and consequences. Geol Mijnb **55**:179–184

Schliep KP (2011) phangorn: Phylogenetic analysis in R. Bioinformatics **27**:592–593

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol **75**:7537–7541

Sepkoski JJ (1981) A factor analytic description of the phanerozoic marine fossil record. Paleobiology **7**:36–53

Servais T, Vincent P, Danelian T, Klug C, Martin R, Munnecke A, Nowak H, Nützel A, Vandenbroucke TRA, Williams M, Rasmussen CMØ (2016) The onset of the 'Ordovician Plankton Revolution' in the late Cambrian. Palaeogeogr Palaeoclimatol Palaeoecol **458**:12–28

Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics **30**:1312–1313

Stoecker DK, Hansen PJ, Caron DA, Mitra A (2016) Mixotrophy in the Marine Plankton. Annu Rev Mar Sci **9**:311–335

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

Sugiyama K, Anderson OR (1997) Experimental and observational studies of radiolarian physiological ecology, 6. Effects of silicate-supplemented seawater on the longevity and weight gain of spongiose radiolarians *Spongaster tetras* and *Dictyocoryne truncatum*. Mar Micropaleontol **29**:159–172

Suzuki N, Aita Y (2011) Radiolaria: achievements and unresolved issues: taxonomy and cytology. Plankt Benthos Res **6**:69–91

Suzuki N, Not F (2015) Biology and Ecology of Radiolaria. In Ohtsuka S, Suzaki T, Horiguchi T, Suzuki N, Not F (eds) Marine Protists. Springer; Tokyo, pp 179–222

Suzuki N, Oba M (2015) Chapter 15. Oldest Fossil Records of Marine Protists and the Geologic History Toward the Establishment of the Modern-Type Marine Protist World. In Ohtsuka S, Suzaki T, Horiguchi T, Suzuki N, Not F (eds) Marine Protists. Springer; Tokyo, pp 359–394

Takahashi K (1982) Dimensions and sinking speeds of tropical radiolarian skeletons from the PARFLUX sediment traps: technical report. Woods Hole Oceanog Inst Tech Rep **WHOI-82-25**:111

Takahashi S, Yamakita S, Suzuki N, Kaiho K, Ehiro M (2009) High organic carbon content and a decrease in radiolarians at the end of the Permian in a newly discovered continuous pelagic section: A coincidence? Palaeogeogr Palaeoclimatol Palaeoecol **271**:1–12

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York, 260p

Yabuki A, Chao EE, Ishida KI, Cavalier-Smith T (2012) *Microheliella maris* (Microhelida ord. n.), an ultrastructurally highly distinctive new axopodial protist species and genus, and the unity of phylum Heliozoa. Protist **163**:356–388

Yilmaz IO, Altiner D, Tekin UK, Ocakoglu F (2012) The first record of the " Mid-Barremian" Oceanic Anoxic Event and the Late Hauterivian platform drowning of the Bilecik platform, Sakarya Zone, western Turkey. Cretac Res **38**:16–39

Yuasa T, Dolven JK, Bjørklund KR, Mayama S, Takahashi O (2009) Molecular phylogenetic position of *Hexacontium pachydermum* Jørgensen (Radiolaria). Mar Micropaleontol **73**:129–134

Yuasa T, Horiguchi T, Mayama S, Takahashi O (2016) *Gymnoxanthella radiolariae* gen. et sp. nov. (Dinophyceae), a dinoflagellate symbiont from solitary polycystine radiolarians. J Phycol **52**:89–104

Zhang K, Feng QL (2019) Early Cambrian radiolarians and sponge spicules from the Niujiaohe Formation in South China. Palaeoworld **28**:234–242

Zhang L, Suzuki N, Yasuhide N, Tuji A (2018) Modern shallow water radiolarians with photosynthetic microbiota in the western North Pacific. Mar Micropaleontol **139**:1–27