Supporting Information

Carr *et al.* 10.1073/pnas.0801667105

SI Text

PCR and Sequencing Protocols. SSU rDNA products were sequenced directly by using the primers listed in Table S1 at the Genomics Lab, University of Birmingham. For all other genes, gel extracted PCR products were first cloned into the plasmid pGEM-T-Easy (Promega) and transformed into Subcloning Efficiency DH5α competent cells (Invitrogen). For sequencing, plasmid DNA was prepared from 5 ml overnight cultures by using the Qiagen Spin Miniprep kit. Sequencing was performed on a minimum of four clones per PCR product at Macrogen Inc. by using T7 and SP6 primers. Sequences were assembled into contigs by using the program BBEdit (version 6.1, Bare Bones Software).

SSU and LSU rDNA. A single PCR fragment was amplified by using the primers 1F and 1528R (1) with an annealing temperature of 52°C. Sequencing reactions were performed on both forward and reverse strands by using the primers listed in Table S1. Four overlapping regions of LSU rDNA were amplified by using the primer pairs NLF184/21 + NLR1126/22, NLF1105/22 + NLR2098/24, NLF1999/18 + NLR2781/19, and NLF2551/21 + NLR3535/22 (2). Annealing was performed at 53°C for the primer pair NLR108/18 + NLR1126/22 and at 58°C for all other pairings.

alpha-tubulin (tubA). PCR amplifications of *tubA* were performed in two rounds. A primary PCR was performed with the primer pair atF1 + atR3 (Table S1), and a 10°C touchdown step with annealing temperature reduced from 60°C to 50°C over 10 cycles. 1μl of product from the initial round of PCR was then used as the template for a second round of PCR by using the nested primers atF3 and one of either atR2 or atR1. All nested PCRs used an annealing temperature of 58°C.

90-kDa heat shock protein (hsp90). All primers used for *hsp90* amplification are listed in Table S1. All amplifications involved an initial round of touchdown PCR with the primer pair HSP2F + HSP2R and an annealing temperature of 60.7°C, which was reduced to 55.7°C over five cycles. The lower temperature was then used for annealing in a further 25 cycles. Product (1μl) from the initial round of PCR was then used as template for a second round of PCR by using HSP3F + HSP3R as nested primers and an annealing temperature of 50°C.

The remainder of the gene was obtained by using seminested PCR with primers designed after conceptual translation of the obtained *hsp90* sequences. For the upstream portion of the gene, amplification proceeded by an initial round of touchdown PCR with primers HSP1F and HSP3R followed by a second PCR round by using 1μl of product as template with the nested primers HSP2F and HSP4R. Amplification of the downstream region followed the same protocol but used primers HSP3F and HSP1R for the first round of amplification and primers HSP4F and HSP2R for the second. Longer *hsp90* sequences were obtained for *Diaphanoeca grandis* because of successful amplification with the primer pair HSP1F + HSP4R.

Diplotheca costata hsp90 was amplified as three overlapping fragments. The 5′ terminus was obtained by two rounds of amplification by using primers HSP1F and HSP3R in the first round and 1$\mu\mu$l of primary product with primers HSP2F and HSP4R in the second. The downstream region was obtained by a similar protocol but used the primers HSP3F and HSP1R, followed by the primers HSP-loricate-F and HSP2R. These sequences were then used to design the *Diplotheca costata* specific primers HSP-Dc-F and HSP-Dc-R to obtain the middle portion of the gene.

Downloaded Sequences. The concatenated four gene dataset included the following sequences downloaded from GenBank: Apusozoa (*Ancyromonas sigmoides* DQ207566, AY752988, AY615870, and EF455765; *Apusomonas proboscidea* DQ207568, DQ980467, DQ980460, and DQ98053), Fungi (*Batrachochytrium dendrobatidis* AF164301, AY079001, and AY546693; *Rhizophlyctis rosea* AY635829 and DQ273787; *Saccharomyces cerevisiae* EU011664, J01355, K01387, and M28428; *Schizosaccharomyces pombe* X54866, EU011663, NM 001019786, and NM 001023795), Nucleariidae (*Nuclearia simplex* AF349566, AY148095, and AY582858); Ichthyosporea (*Amoebidium parasiticum* Y19155; *Icthyophonus hoferi* U25637, and AY026370), Metazoa (*Beroe ovata* AF293694, and AY026369; *Halichondria* sp. AY226083, and AY226049; *Haliclona rubens* AY226050, and AY226084; *Leucosolenia* sp. AF100945, AY226053, and AY226087; *Leucosolenia* sp. MMM-2001 AY026372; *Mnemiopsis leidyi* AF293700, AY026373, AY226058, and AY226092; *Nematostella vectensis* AF254382, AY226056, and AY226090; *Suberites fuscus* AY226051 and AY226085; *Sycon* sp. AY226088 and AY226054; *Trichoplax adhaerens* AY652581 and AY303975), Choanoflagellida (*Helgoeca nana* L10823; Taxonomy ID258836 AF27200, *Desmarella moniliformis* AF084231; *Lagenoeca* sp. *antarctica* DQ995807; *Proterospongia* sp. ATCC 50818 AY226048 and AY226082; *Salpingoeca* sp. *abyssalis* DQ995808; *Salpingoeca infusionum* AF100941 and AY026380). Holozoa *incertae sedis* (*Capsaspora owczarzaki* AF349564 and AY724688; *Corallochytrium limacisporum* L42528, AY582859, and AY236712).

Additional sequences were obtained as follows. EST sequences were downloaded from the TBestDB database (3) for *Amoebidium parasiticum* (Cluster ABL00000331) and *Capsaspora owczarzaki* (Clusters NUL00000500, NUL0000031, and NUL00000511). The LSU sequence for *Nematostella vectensis* was obtained from StellaBase c436002894.Contig1 (4). The *hsp90* and *tubA* sequences for *Trichoplax adhaerens* were obtained from Scaffold 6 and 1001, respectively, of the Joint Genome Institute genome sequence assembly (5). The *hsp90* and *tubA* sequences (loci BDEG 05207.1 and BDEG 00078.1, respectively) for *Batrachochytrium dendrobatidis* were obtained from the Broad Institute genome database (6).

Measurement of Character States. Because the living cell of choanoflagellates is so consistent for all taxa (7, 8), morphological aspects of the protoplast are not included in the following matrix. The periplast structures, in particular the theca and lorica, provide the major sources of variation and therefore their morphology is included in considerable detail. Also included are data on ecology and habitat.

Description of Characters for Matrix. ***Organic covering.***

Glycocalyx. Term used to describe an evenly dispersed layer of fine fibrils on the external surface of a plasma membrane. The glycocalyx can be visualised in thin sections by transmission electron microscopy (TEM) after material is fixed in glutaraldehyde followed by osmium tetroxide. Epon sections are stained with lead citrate and uranyl acetate (9–11).

Glycocalyx only. Specifically refers to the outer covering of a cell consisting of a glycocalyx alone. This terminology is used to distinguish this condition from cells that possess a glycocalyx and another organic covering, such as a theca (10).

Theca. (Latin from Greek = sheath or case) A continuous investment that surrounds the mid- and posterior part of a

choanoflagellate cell. Usually consists of microfibrils, probably of carbohydrate composition, embedded in an amorphous matrix (Fig. 1 *B* and *C*).

Cup-shaped theca. A cup-shaped investment surrounding the posterior part of the cell comprising a compact arrangement of microfibrils embedded in an amorphous matrix (Fig. 1*C*). In section, the cup-shaped theca appears as a single dark layer. Sometimes there is evidence of an inner extension of material to attach the theca to the enclosed cell.

Flask-shaped theca. A theca in the shape of a round-bottomed flask with a number of highly distinctive characters (Fig. 1*B*). The neck of the flask usually reveals a pattern of longitudinal microfibrils; the anterior edge of the neck is curved outwards; on the inner surface of the base of the neck is an inwardly directed flange with regular longitudinal thickenings (Fig. 1*B*, arrow). In section the wall of the flask comprises three layers – a less dense staining layer sandwiched between two darker staining granular layers (10, 11).

Microfibrillar stalk (peduncle). Parallel-sided stalk usually comprising a compact arrangement of longitudinal microfibrils (Fig. 1*C*). Sometimes with a pattern of horizontal markings (9). In some species the microfibrils are obviously confluent with microfibrils of the cup or flask. Serves to attach the theca and cell to a substratum.

Protoplast sheath. Investment consisting of compacted microfibrils within an amorphous matrix, similar in appearance to a theca, that surrounds the posterior part of the protoplast of loricate species. Where the protoplast is held closely within the lorica, the sheath covers the inner surface of the lorica and serves to hold the protoplast in position (12, 13).

Regular microfibrils or ''veil''. Diaphanous investment of microfibrils that covers specific parts of the upper lorica chamber. In some species e.g., *Savillea* and *Acanthoeca*, the investment consists of a meshwork of horizontal and vertical microfibrils giving a regular arrangement of rectangular spaces (14–18). In other species e.g., *Diplotheca* and *Diaphanoeca*, the diaphanous arrangement of microfibrils appears not to be so regularly arranged and comprises the ''veil'' (19). The veil usually covers a specific portion of the inner surface of the lorica chamber allowing spaces for the passage of a water current created by the flagellum.

Lorica Features. *Lorica.* (Latin = armour, shell, case) Term used to describe the basket-like siliceous covering of exclusively marine choanoflagellates (Acanthoecidae) (Fig. 1 *E* and *F*). The lorica is composed of rod-shaped costal strips attached to each other end-to-end to form costae (ribs) (20, 21).

Costa. Elongated siliceous rib composed of rod-shaped costal strips attached end-to-end.

Costal strip. Siliceous rod-shaped structure – the unit of costal construction.

Longitudinal costae. Costae orientated parallel to the long axis of the cell. Longitudinal costae usually form the outer layer of the lorica (15). Longitudinal costae originate from a single bundle of costal strips.

Helical costae. Costae orientated to form a helix (15). Helical costae, in common with longitudinal costae, originate from a single bundle of costal strips. Helical costae often comprise the inner layer of the lorica.

Transverse (ring) costae. Costae arranged perpendicular (horizontal) to the long axis of the cell usually forming the inner layer of the lorica (Fig. 1*F*). Often the anterior costa (ring) is on the outer surface of the lorica (22). Transverse costae are assembled from several (usually 4) bundles of costal strips. They are not pulled out from a single bundle and must therefore be considered as analogous to longitudinal and helical costal strips.

Logistics of costal strip accumulation. Before lorica assembly, a full complement of costal strips is produced and stored externally on the cell surface. Lorica assembly is then a single continuous process that takes between 5–15 min (16, 23–25). The location order in which the strips are accumulated and the positioning of strips on the juvenile before lorica assembly differ between nudiform and tectiform taxa (14).

Location of costal strips-Nudiform taxa. Costal strips are accumulated in bundles on the surface of the juvenile (previously motile) cell (15–17).

1. *Tectiform taxa* - Costal strips are accumulated in bundles at the top of the inner surface of the collar (24).

Order in which costal strips are produced. 1. *Nudiform taxa* - The costal strips for longitudinal costae are deposited first followed by those for the helical costae.

2. *Tectiform taxa* - The costal strips for transverse costae are deposited first followed by those for the longitudinal costae.

Positioning of strips on the juvenile cell. 1. *Nudiform taxa* - All costal strips are located vertically on the surface of the juvenile cell (15–18).

2. *Tectiform taxa* - Costal strips destined for longitudinal and helical costae are located vertically on juvenile cell. Costal strips destined for transverse costae are located horizontally (26).

Lorica dimensions. Length of the lorica chamber is recorded; that is from the top of the stalk to the extreme anterior end. The length of the stalk is omitted.

Cell division. *Lateral cell division.* This is when a cell divides in a horizontal plane with respect to the longitudinal axis of the cell. This type of division is characteristic of a cell that is not restricted in any way (7).

Diagonal division. This is when a nudiform cell divides within a lorica (15, 17). One daughter protoplast remains within the parent lorica, whereas the other (the juvenile) appears diagonally above it (Fig. 1*E*).

Emergent division. This form of division is characteristic of thecate species where lateral division is impossible (10). The parent cell becomes amoeboid, emerges from the theca, the nucleus divides, one nucleus remains in the protoplast within the thecal neck whereas the other moves into the daughter cell that bears a flagellum and will swim away.

Inverted division. This form of division is exclusive to tectiform species (26). Nuclear division proceeds normally in a lateral fashion; the cell divides diagonally and is inverted as it is pushed into the accumulated costal strips at the top of the collar (Fig. 1*F*).

Motile juvenile. A motile juvenile cell is produced when a parent cell divides to give two daughter cells, one of which remains in the parent periplast whereas the other daughter (the juvenile) swims away (9, 10, 14, 17). This serves as an efficient means of dispersal for an otherwise sedentary or passively moving species.

Ecology and habitat. *Cyst.* A thick-walled completely enclosed cell that undergoes a period of perennation. The cyst of *Monosiga ovata* develops from a cell that produces a flask-shaped wall that subsequently thickens (27).

Marine. In the current context marine is taken to include brackish water. Whereas the usual salinity of undiluted seawater is 35‰, some marine species can survive salinities as low as 2‰.

Freshwater. Water with negligible salt content (<0.1‰ salinity).

Embedded in biofilm. Biofilms, comprising bacteria, fungi, and other microorganisms, occur on surfaces exposed to natural aquatic conditions. A few choanoflagellate species actually become incorporated into biofilms (15).

Sedentary. Attached to a surface, usually by a stalk.

Freely suspended. Refers to cells that are not attached to a substratum. They may be benthic but are still capable of being moved passively in the water currents. Other species may be suspended within the surface plankton.

Introduction to Morphological Systematics of Choanoflagellates. Choanoflagellates are small colorless protozoa comprising a spherical to ovoid cell with a single anterior flagellum surrounded by a funnel-shaped collar composed of 30 or more actin-based tentacles or microvilli (8). Cells are universally distributed in aquatic habitats where they are filter feeders; the flagellum creates a current of water from which particles, in particular bacteria, are trapped on the outer surface of the collar and subsequently ingested. Although this is an effective method of acquiring prey, a major limitation of the system is that, unless a cell is restrained in some way, movement of the flagellum creates a locomotory force (28, 29). Thus, moving cells are poor feeders and extracellular structures, in the form of an organic theca or basket like lorica, have evolved to render individual cells immobile. In contrast to the constancy of cell morphology, the outer coverings of choanoflagellates are much more variable and it is the capacity of these structures for variation that has been key to the group's evolutionary diversity and ecological success.

Choanoflagellate systematics, as presently construed, is based on observations made in the nineteenth century, and in particular the work of Saville-Kent (30). Because cell structure is so constant, Kent based his classification on the morphology of the extracellular coverings. Working with mixed samples collected from the field and laboratory aquaria Kent distinguished two families on the basis of the thickness and rigidity of their outer organic coverings. Members of the Family Codonosigidae Kent are distinguished by a thin almost invisible coat, now called a glycocalyx, which surrounds the cell and is usually associated with a stalk of variable length. The glycocalyx is sufficiently thin and expandable to permit lateral cell division. The equal products of division may remain as a multicellular head on a single stalk or, alternatively, one daughter cell may break loose, swim away and settle down on a suitable surface. Should this cell divide while motile, the beginnings of a 'proterospongia' colony may be found. Members of the second family, the Salpingoecidae Kent, possess a substantial microfibrillar theca that may be cup-, tube- or flask-shaped and is attached to a surface by a stalk similar to that found in the Codonosigidae. However, the theca is thicker and more rigid than the glycocalyx and does not allow lateral expansion of the cell during division. This limitation is overcome by the cell becoming amoeboid and emerging through the anterior aperture of the theca, so that cell division can occur outside the theca (10).

It was not until 1965 that a third group of choanoflagellates was recognized and given family status (Acanthoecidae Norris) (31). Members of this group are distinguished by the possession of a basket-like lorica composed of siliceous costae comprising rod-like costal strips attached to each other end-to-end. The lorica is not entirely inorganic; all loricate species possess an inner organic investment which, in some species, directs the flow of water across the collar tentacles.

All loricae are based on two layers of siliceous costae. The outer layer is usually longitudinal and is found in all species sampled. These longitudinal costae are held in place by an inner layer of helical or transverse (ring) costae. Longitudinal and helical costae are homologous in that they are assembled from single groups of costal strips that are located vertically on the juvenile cell (15). Ring costae, however, are analogous in that each ring is derived from several bundles of costal strips. The latter are located horizontally on the juvenile cell.

Two distinctive groups of loricate choanoflagellates can be distinguished on the basis of the timing of costal strip production in the cell cycle and the morphology of cell division. One group displays *nudiform* replication, whereby a cell with a lorica divides to produce a 'naked' motile cell that swims away from the parent lorica, settles on to a surface, produces a set of costal strips and then assembles its own lorica. Before lorica construction, the costal strips are stored on the surface of the cell as vertically orientated bundles (15, 17).

The other loricate group displays *tectiform* replication. In tectiform species the lorica comprises vertical and horizontal costae, and sometimes helical costae. Costal strips are produced by the parent cell before cell division. Thus, a complete set of strips is transferred to the daughter cell as it is inverted and pushed out of the parent lorica. This set of strips is then rapidly assembled into a mature lorica (25, 26).

Currently, only six nudiform species attributable to three genera have been recognized, in comparison with over one hundred tectiform species (15, 17). Whereas nudiform species are most often components of marine biofilms, tectiform species have broken with the necessity of being attached to a surface and diversified to occupy the planktonic environment on a global scale (20, 22, 32).

1. Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* 71:491–499.
2. European Ribosomal RNA Database. http://www.psb.ugent.be/rRNA/.
3. TBestDB: Taxonomically Broad EST Database http://amoebidia.bcm.umontreal.ca/pepdb/searches/welcome.php
4. StellaBase. *Nematostella vectensis* Genomics Database. http://evodevo.bu.edu/stellabase/.
5. JGI *Trichoplax adhaerens v1.0.* http://genome.jgi-psf.org/Triad1/Triad1.home.html.
6. Broad Institute *Batrachochytrium dendrobatidis* Database. http://www.broad.mit-.edu/annotation/genome/batrachochytrium dendrobatidis.3/Home.html *Batrachochytrium dendrobatidis* Database.
7. Karpov SA, Leadbeater BSC (1997) Cell and nuclear division in a freshwater choanoflagellate. *Monosiga ovata* Kent. *Europ J Protistol* 33:323–334.
8. Karpov SA, Leadbeater BSC (1998) Cytoskeleton structure and composition in choanoflagellates. *J Eukaryot Microbiol* 45:361–367.
9. Leadbeater BSC, Morton C (1974) A microscopical study of a marine species of *Codosiga* James-Clark (Choanoflagellata) with special reference to the ingestion of bacteria. *Biol J Linn Soc* 6:337–347.
10. Leadbeater BSC (1977) Observations on the life-history and ultrastructure of the marine choanoflagellate *Choanoeca perplexa* Ellis. *J Mar Biol Assoc UK* 57:285–301.
11. Leadbeater BSC (1983) Life-history and ultrastructure of a new marine species of *Proterospongia* (Choanoflagellida). *J Mar Biol Assoc UK* 63:135–160.
12. Leadbeater BSC (1972) Ultrastructural observations on some marine choanoflagellates from the coast of Denmark. *Br Phycol J* 7:195–211.
13. Leadbeater BSC (1979) Developmental studies on the loricate choanoflagellate *Stephanoeca diplocostata* Ellis. I. Ultrastructure of the non-dividing cell and costal strip production. *Protoplasma* 98:241–262.
14. Leadbeater BSC (1975) A microscopical study of the marine choanoflagellate *Savillea micropora* (Norris) comb. nov. and preliminary observations on lorica development in *S. micropora* and *Stephanoeca diplocostata* Ellis. *Protoplasma* 83:119–129.
15. Leadbeater BSC (2008) Choanoflagellate lorica construction and assembly: The nudiform condition. I. *Savillea* species *Protist* 159:259–268.
16. Leadbeater BSC, Morton C (1974) A light and electron microscope study of the choanoflagellates *Acanthoeca spectabilis* Ellis and *A. brevipoda* Ellis. *Biol J Linn Soc* 6:337–347.
17. Leadbeater BSC, Henouil M, Berovic N (2008) Choanoflagellate lorica construction and assembly: The nudiform condition. II. *Acanthoeca spectabilis. Protist* 159:495–505.
18. Buck KR, Marchant HJ, Thomsen HA, Garrison DL (1990) *Kakoeca antarctica* gen. et sp. n., a loricate choanoflagellate (Acanthoecidae, Choanoflagellida) from Antarctic sea ice with a unique protoplast suspensory membrane. *Zool Scripta* 19:389–394.
19. Manton I, Bremer G, Oates K (1981) Problems of structure and biology in a large collared flagellate (*Diaphanoeca grandis* Ellis) from Arctic Seas. *Proc Roy Soc Lond B* 213:15–26.
20. Thomsen HA, Buck KB, Coale SL, Garrison DL, Gowing MM (1990) Loricate choanoflagellates (Acanthoecidae Choanoflagellida) from the Weddell Sea, Antarctica. *Zool Scripta* 19:367–387.
21. Thomsen HA, Buck KR (1991) Choanoflagellate diversity with particular emphasis on the Acanthoecidae. In: *Free Living Heterotrophic Flagellates,* eds Patterson DJ, Larsen J (Clarendon Press, Oxford) Vol 45, pp 259–284.
22. Thomsen HA, Boonruang P (1984) A light and electron microscopical investigation of loricate choanoflagellates (Choanoflagellida, Acanthoecidae) from the Andaman Sea, S.W. Thailand and Denmark; species of *Cosmoeca* gen. n. *Zool Scripta* 13:165–181.
23. Leadbeater BSC (1979) Developmental studies on the loricate choanoflagellate *Stephanoeca diplocostata* Ellis. I. Ultrastructure of the non-dividing cell and costal strip production. *Protoplasma* 98:241–262.
24. Leadbeater BSC (1994) Developmental studies on the loricate choanoflagellate *Stephanoeca diplocostata* Ellis. VII. Dynamics of costal strip accumulation and lorica assembly. *Europ J Protistol* 30:111–124.
25. Leadbeater BSC (1994) Developmental studies on the loricate choanoflagellate *Stephanoeca diplocostata* Ellis. VIII. Nuclear division and cytokinesis. *Europ J Protistol* 30:171–183.

26. Leadbeater BSC (1979) Developmental studies on the loricate choanoflagellate *Stephanoeca diplocostata* Ellis. II. Cell division and lorica assembly. *Protoplasma* 98:311–328.
27. Leadbeater BSC, Karpov SA (2000) Cyst formation in a freshwater strain of the choanoflagellate *Desmarella moniliformis* Kent. *J Eukaryot Microbiol* 47:433–439.
28. Pettitt MA, Orme BAA, Blake JR, Leadbeater BSC (2002) The hydrodynamics of filter feeding in choanoflagellates. *Europ J Protistol* 38:313–332.
29. Lighthill J (1976) Flagellar hydrodynamics. *SIMA Reviews* 18:161–230.
30. Kent SW (1882) *A Manual of the Infusoria* (D. Bogue, London),Vols 1–3, p 913.
31. Norris RE (1965) Neustonic marine Craspedomonadales (Choanoflagellata) from Washington and California. *J Protozool* 12:589–612.
32. Leakey RJG, Leadbeater BSC, Mitchell E, McCready SMM, Murray AWA (2002) The abundance and biomass of choanoflagellates and other nanoflagellates in waters of contrasting temperature to the north-west of South Georgia in the Southern Ocean. *Europ J Protistol* 38:333–350.

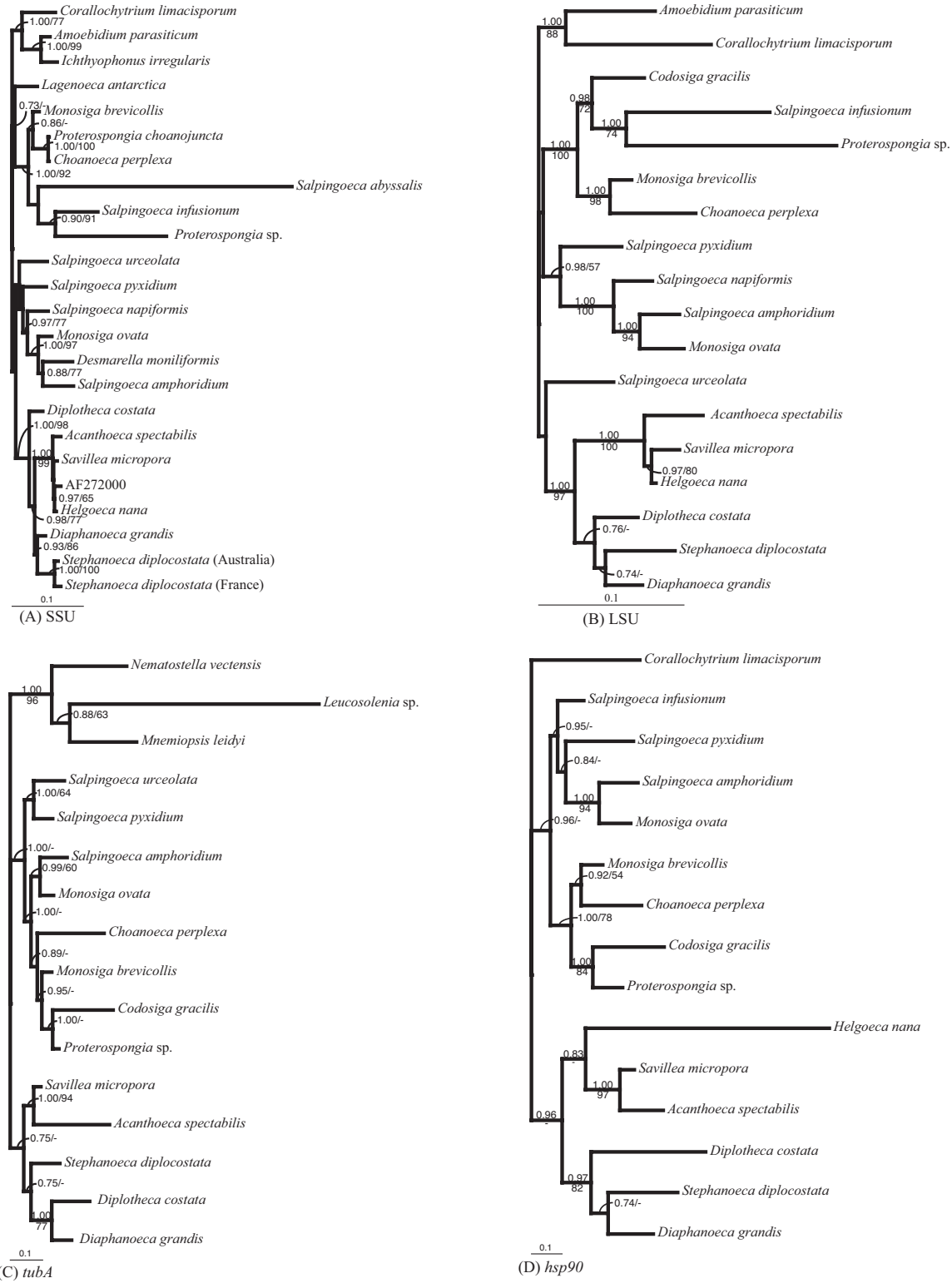

Fig. S1. Single gene phylogenies of Choanoflagellates. Trees are shown for individual analyses of each of the four genes comprising the concatenated dataset used for Fig. 2. Trees were derived by Bayesian inference on individual alignments of SSU rDNA (*A*), LSU rDNA (*B*), *tubA* (*C*) and *hsp90* (*D*). All analyses were conducted on nucleotide sequences including third codon positions for protein coding genes. Branches are drawn proportional to the number of nucleotide substitutions per site as indicated by the scale bar at the lower left. biPP and mlBP values are given above and below the branches respectively.

Table S1. PCR primers designed for this study

Gene	Primer	Nucleotide Sequence
SSU	300F	5′- AGGGTTCGATTCCGGAG -3′
	528F	5′- CGGTAATTCCAGCTCC -3′
	640F	5′- YAGAGGTGAAATTCT -3′
	1055F	5′- GGTGGTGCATGGCCG -3′
	1200F	5′- CAGGTCTGTGATGCTC -3′
	300R	5′- TCAGGCTCCCTCTCCGG -3′
	536R	5′- GWATTACCGCGGCKGCTG -3′
	690R	5′- AGAATTTCACCTCTG -3′
	1055R	5′- CGGCCATGCACCACC -3′
	1200R	5′- GGGCATCACAGACCTG -3′
tubA	atF1	5′-GGGCCCCAGGTCGGCAAYGCNTGYTGG-3′
	atF3	5′-CTCAGGGGNAARGCAGAYGC-3′
	atR1	5′-CACCAGGTTNGTYTGRAAYTC-3′
	atR2	5′-GCGCATAACCTCNCCNACRTACCA-3′
	atR3	5′-AAGGCCTTCNCCYTCYTCCAT-3′
hsp90	HSP1F	5′-CCAGCCTACGAC(TCN/AGY)AAYAARGAR-3′
	HSP2F	5′-GACATCAGCATGATCGGNCARTTYGGN-3′
	HSP3F	5′-ACGACCGCGTCAARCTNTAYGTN-3′
	HSP4F	5′-AAGCTGGGGATHCAYGARGAY-3′
	HSP1R	5′-CCATCTTCCRTCNACYTCYTCCAT-3′
	HSP2R	5′-CCGCGCYTTCATBATYCKYTCCAT-3′
	HSP3R	5′-AACAGCGTACTCRTCBATNGGNTC-3′
	HSP4R	5′-CTCGCAGTTRTCCATDATRAANAC-3′
	HSP-loricate-F	5′-CTCAAGGACTACGTCACCCGC-3′
	HSP-Dc-F	5′-TGGGAGGATCATCTTGCCGT-3
	HSP-Dc-R	5′-CTGCATCCTTAGTTTCACCAG-3

Table S2. Morphological and ecological characters and states for the sixteen choanoflagellate species shown in the matrix (Table S3)

Character	Character State
Organic Covering	
- Only Glycocalyx In Sedentary Stage	0 = absent, 1 = present
- Glycocalyx In Motile Stage Only	0 = absent, 1 = present
- Cup-Shaped Theca	0 = absent, 1 = present
- Flask-Shaped Theca	0 = absent, 1 = present
- Microfibrillar Stalk	0 = absent, 1 = present
- Protoplast Sheath	0 = absent, 1 = present
- Regular Microfibrils or Veil	0 = absent, 1 = present
Lorica Features	
- Longitudinal Costae In Lorica	0 = absent, 1 = present
- Helical Costae In Lorica	0 = absent, 1 = present
- Transverse Rings In Lorica	0 = absent, 1 = present
- Location of Costal Strip Bundles	0 = absent, 1 = around surface of juvenile cell; 1 = inside top of parent cell collar
- Order of Costal Strip Production	0 = absent, 1 = longitudinal costae first, helical costae second; 2 = transverse costae first, longitudinal costae second
- Vertical Bundles on Juvenile	0 = absent, 1 = present
- Horizontal Bundles on Juvenile	0 = absent, 1 = present
- Lorica Size (Length)	0 = absent, 1= <8 mm, 2 = 9–12 mm, 3 = 12–15 mm, 4=>15 mm
Cell Division	
- Lateral Division In Sedentary Stage	0 = absent, 1 = present
- Lateral Division In Motile Stage	0 = absent, 1 = present
- Diagonal Division	0 = absent, 1 = present
- Emergent Division	0 = absent, 1 = present
- Inverted Cell Division	0 = absent, 1 = present
- Motile Juvenile Dispersal Stage	0 = absent, 1 = present
Ecology and Habitat	
- Cyst Forming	0 = absent, 1 = present
- Marine	0 = absent, 1 = present
- Freshwater	0 = absent, 1 = present
- Embedded in Biofilm	0 = absent, 1 = present
- Sedentary	0 = absent, 1 = present
- Passively Suspended	0 = absent, 1 = present

Table S3. Matrix of morphological and ecological characters for the 16 choanoflagellate species in Fig. 3

	Organic Covering							Loricate Features							
	Only Glycocalyx In Sedentary Stage	Glycocalyx In Motile Stage	Cup-shaped Theca	Flask-shaped Theca	Microfibrillar Stalk	Protoplast Sheath	Regular Microfibrils/ Veil	Longitudinal Costae In Lorica	Helical Costae In Lorica	Transverse Rings In Lorica	Location of Costal Storage Bundles	Order of Costal Strip Production	Vertical Storage Bundles	Horizontal Storage Bundles	Lorica Size
Codosiga gracilis	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Salpingoeca infusionum	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0
Proterospongia sp.	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0
Monosiga brevicollis	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Choanoeca perplexa	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
Salpingoeca urceolata	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
Salpingoeca pyxidium	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
Salpingoeca napiformis	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
Salpingoeca amphoridium	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
Monosiga ovata	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Helgoeca nana	0	1	0	0	0	1	1	1	1	0	1	1	1	0	2
Savillea micropora	0	1	0	0	0	1	1	1	1	0	1	1	1	0	1
Acanthoeca spectabilis	0	1	0	0	0	1	1	1	1	0	1	1	1	0	2
Diplotheca costata	0	0	0	0	0	1	1	1	0	1	2	2	1	1	2
Diaphanoeca grandis	0	0	0	0	0	1	1	1	0	1	2	2	1	1	4
Stephanoeca diplocostata	0	0	0	0	0	1	0	1	1	1	2	2	1	1	3

	Cell Division						Ecology And Habitat					
	Lateral Division In Sedentary Stage	Lateral Division In Motile Cell	Diagonal Division	Emergent Division	Inverted Division	Motile Juvenile Dispersal Stage	Cyst Forming	Marine	Freshwater	Embedded In Biofilm	Sedentary	Passively Suspended
Codosiga gracilis	1	1	0	0	0	1	0	1	0	0	1	0
Salpingoeca infusionum	0	1	0	1	0	1	0	1	0	0	1	0
Proterospongia sp.	0	1	0	1	0	1	0	1	0	0	1	0
Monosiga brevicollis	1	1	0	0	0	1	0	1	0	0	1	0
Choanoeca perplexa	0	1	0	1	0	1	0	1	0	0	1	0
Salpingoeca urceolata	0	1	0	1	0	1	0	1	0	0	1	0
Salpingoeca pyxidium	0	1	0	1	0	1	0	1	0	0	1	0
Salpingoeca napiformis	0	1	0	1	0	1	0	0	1	0	1	0
Salpingoeca amphoridium	0	1	0	1	0	1	0	0	1	0	1	0
Monosiga ovata	1	1	0	0	0	1	1	0	1	0	1	0
Helgoeca nana	0	0	1	0	0	1	0	1	0	1	1	0
Savillea micropora	0	0	1	0	0	1	0	1	0	1	1	0
Acanthoeca spectabilis	0	0	1	0	0	1	0	1	0	0	1	0
Diplotheca costata	0	0	0	0	1	0	0	1	0	0	1	1
Diaphanoeca grandis	0	0	0	0	1	0	0	1	0	0	0	1
Stephanoeca diplocostata	0	0	0	0	1	0	0	1	0	0	1	1