J. Eukaryot. Microbiol., 53(5), 2006 pp. 379–384

DOI: 10.1111/j.1550-7408.2006.00118.x

Insights into the Evolutionary Origin and Genome Architecture of the Unicellular Opisthokonts *Capsaspora owczarzaki* and *Sphaeroforma arctica*

IÑAKI RUIZ-TRILLO, CHRISTOPHER E. LANE, JOHN M. ARCHIBALD and ANDREW J. ROGER

Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS, Canada B3H 1X5

ABSTRACT. Molecular phylogenetic analyses have recently shown that the unicellular amoeboid protist *Capsaspora owczarzaki* is unlikely to be a nucleariid or an ichthyosporean as previously described, but is more closely related to Metazoa, Choanoflagellata, and Ichthyosporea. However, the specific phylogenetic relationship of *Capsaspora* to other protist opisthokont lineages was poorly resolved. To test these earlier results we have expanded both the taxonomic sampling and the number of genes from opisthokont unicellular lineages. We have sequenced the protein-coding genes elongation factor 1-α (*EF1-α*) and heat shock protein 70 (*Hsp70*) from *C. owczarzaki* and the ichthyosporean *Sphaeroforma arctica*. Our maximum likelihood (ML) and Bayesian analyses of a concatenated alignment of *EF1-α*, *Hsp70*, and *actin* protein sequences with a better sampling of opisthokont-related protist lineages indicate that *C. owczarzaki* is not clearly allied with any of the unicellular opisthokonts, but represents an independent unicellular lineage closely related to animals and choanoflagellates. Moreover, we have found that the ichthyosporean *S. arctica* possesses an EF-like (*EFL*) gene copy instead of the canonical *EF1-α*, the first so far described in an ichthyosporean. A maximum likelihood phylogenetic analysis shows that the *EF-like* gene of *S. arctica* strongly groups with the EF-like genes from choanoflagellates. Finally, to begin characterizing the *Capsaspora* genome, we have performed pulsed-field gel electrophoresis (PFGE) analyses, which indicate that its genome has at least 12 chromosomes with a total genome size in the range of 22–25 Mb.

Key Words. *EFL* gene, genome size, molecular systematics, pulsed-field gel electrophoresis.

ONGOING molecular systematic studies have revealed that two major eukaryotic groups, Metazoa (or Animals) and Fungi, share a common ancestor and form a clade known as the opisthokonts (Baldauf and Palmer 1993; Cavalier-Smith and Chao 2003; Lang et al. 2002; Philippe et al. 2004; Ruiz-Trillo et al. 2004; Snell, Furlong, and Holland 2001; Steenkamp, Wright, and Baldauf 2006). The opisthokonts also include some unicellular lineages, such as choanoflagellates, ichthyosporeans, nucleariids, *Corallochytrium*, *Ministeria*, and *Capsaspora owczarzaki* (for a review, see Steenkamp and Baldauf 2004). The monophyly of the opisthokonts is also supported on morphological grounds as they share several ultrastructural characteristics, such as the presence of a unicellular motile stage bearing a single posterior flagellum and flattened mitochondrial cristae (Baldauf and Palmer 1993; Cavalier-Smith and Chao 2003; Steenkamp and Baldauf 2004). Opisthokonts have also been shown to possess a unique ~ 12 amino acid insertion in the protein *EF1-α* (Baldauf and Palmer 1993; Steenkamp et al. 2006). It is among the unicellular opisthokont lineages that we should expect to find the closest extant relatives of both animals and fungi. Identification of these relatives is, therefore, a first crucial step in trying to understand the evolutionary transitions to multicellularity.

Among the unicellular opisthokont lineages, *Capsaspora owczarzaki* is one of the most phylogenetically controversial. *Capsaspora owczarzaki* was first named as *Nuclearia* sp. and, based on its morphology, described as a member of the nucleariids, a unicellular opisthokont lineage of amoeboids (Owczarzak, Stibbs, and Bayne 1980). Molecular phylogenetic analyses of small subunit ribosomal RNA (SSU rRNA) showed *Capsaspora* (then still known as *Nuclearia* sp.) occupies an unclear phylogenetic position near the fungal-metazoan boundary (Amaral Zettler et al. 2001). Later, *Nuclearia* sp. was renamed as *C. owczarzaki* and based on the SSU rRNA tree and its parasitic lifestyle, it was proposed to be a member of the ichthyosporeans (Hertel, Bayne, and Loker 2002), a clade of parasitic and saprophytic opisthokont microbes (also known as DRIPs or mesomycetozoans, for a review, see Mendoza, Taylor, and Ajello (2002)). However, the statistical support for the phylogenetic position of *C. owczarzaki* in that single-gene analysis was weak. A recent multi-gene analysis (*SSU+LSU+Actin*), however, showed that *Capsaspora* is probably not a nucleariid or an ichthyosporean, but an independent lineage closely related to animals, choanoflagellates, and ichthyosporeans (Ruiz-Trillo et al. 2004). Consistent with previous single-gene phylogenies (Medina et al. 2003; Nikolaev et al. 2004) this analysis also showed that nucleariids are the sister-group to Fungi; a position that has been further corroborated by a recent multi-gene analysis (Steenkamp et al. 2006).

The specific phylogenetic position of the ichthyosporeans also remains elusive. Molecular phylogenies based on either ribosomal RNA genes (Amaral Zettler et al. 2001; Hertel et al. 2002; Medina et al. 2003; Mendoza et al. 2002) or mitochondrial proteins (Lang et al. 2002) have consistently shown ichthyosporeans to be the sister-group to choanoflagellates. In the two most recent multigene analyses ichthyosporeans appear either as sister-group to *Capsaspora* and choanoflagellates (Ruiz-Trillo et al. 2004), or grouping with *Corallochytrium* as a sister-group to a clade comprising metazoans, choanoflagellates, and *Ministeria* (Steenkamp et al. 2006). However, despite being the two taxonomically most comprehensive phylogenies of the opisthokonts to date, both those analyses were taxonomically incomplete. They included information from just one member of the ichthyosporeans and lacked molecular data from some key opisthokont taxa, such as *Capsaspora* in Steenkamp et al. (2006) or *Corallochytrium* and *Ministeria* in Ruiz-Trillo et al. (2004). In any case, the recent data have clearly shown that ichthyosporeans, choanoflagellates, *Ministeria*, *Corallochytrium,* and *Capsaspora* are all more closely related to animals than to Fungi (Ruiz-Trillo et al. 2004; Steenkamp et al. 2006).

Thus, under the current phylogenetic scenario of the opisthokonts, the unicellular lineages *Capsaspora*, *Ministeria*, *Corallochytrium*, and the ichthyosporeans appear as new targets (in addition to choanoflagellates) for comparative genomic analyses aimed at understanding the transition from unicellular protozoans to multicellular animals. Furthermore, *C. owczarzaki* is a parasite of the pulmonate snail *Biomphalaria glabrata*, which is also the intermediate host of *Schistosoma*, a human parasite platyhelminth that is the causative agent of schistosomiasis. Besides being a parasite of the same host, *C. owczarzaki* has been also found to attack and kill the sporocysts of *Schistosoma* (Owczarzak et al. 1980). Knowledge of the *C. owczarzaki* genome and biology may provide insights into the co-evolution of this host–parasite system

Corresponding Author: A. J. Roger, Department of Biochemistry and Molecular Biology; Dalhousie University, Halifax, Nova Scotia, B3H 1X5, Canada—Telephone number: 1-902-494-2620; FAX number: 1-902-494-1355; e-mail: Andrew.Roger@Dal.Ca

and, eventually, provide biological strategies to control this human parasite. For these reasons, we have conducted a more in-depth investigation into the phylogenetic position of *C. owczarzaki* and the ichthyosporeans, as well as into *C. owczarzaki* genomic features.

Here we present an improved and updated phylogenetic analysis of opisthokonts, which includes new sequences for the protein-coding genes *EF1-α*, and *Hsp70* from both *C. owczarzaki* and the ichthyosporean *Sphaeroforma arctica*. It represents the first multigene phylogenetic analysis to date that includes all known opisthokont lineages and two members from the Ichthyosporea. The Bayesian and maximum likelihood analyses of a combined dataset of *actin*, *EF1-α*, and *Hsp70* here reported are aimed at testing the phylogenetic position of *Capsaspora* and the ichthyosporeans within the opisthokonts. We also analyze the origin of the *EF-like* gene of the ichthyosporean *S. arctica*. Finally, we present pulsed-field gel electrophoresis (PFGE) data of *C. owczarzaki* that yield insight into its genome size and structure.

MATERIALS AND METHODS

Material sources. A culture of *Capsaspora owczarzaki* (ATCC 30864) was grown at 25 °C on 803 M7 medium as specified by the American Type Culture Collection (ATCC) and cultures of *Sphaeroforma arctica* were grown in MAP-medium (18.6 g/l Difco marine broth 2216, 20 g/l Bacto peptone, 10 g/l NaCl) at 12 °C.

RNA extraction, PCR, cloning, and sequencing. Total RNA was obtained following TRI-reagent (Molecular Research Center, Cincinnati, OH) protocols. Purified polyA+mRNA was obtained using Poly(A) Purist mRNA purification kit from Ambion (Austin, TX). cDNA was obtained via a reverse transcription-PCR by using SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA) following the manufacture protocols. *EF1-α* and *Hsp70* sequences were amplified using different combinations of primers as in Simpson, Inagaki, and Roger (2005). Polymerase chain reaction products were purified and cloned using TOPO-TA cloning kit for sequencing (Invitrogen). Both coding and non-coding strands were sequenced using an ABI PRISM BigDye Termination Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) from three different clones. Both *EF1-α* and *Hsp70* sequences were also obtained by an EST project of *C. owczarzaki* and *S. arctica*. New sequences have been deposited in GenBank under accession numbers: DQ403163–166.

Alignment and phylogenetic analyses. A concatenated alignment, which is available upon request, of the three genes (*EF1-α*, *Hsp70*, and *actin*), including a broad sampling of opisthokonts, was constructed by eye using BioEdit (Tom Hall, Ibis Therapeutics, Carlsbad, CA). The amoebozoans *Dictyostelium* and *Acanthamoeba* were used as outgroups. Some sequences were obtained from publicly accessible genome or EST projects. In some cases, sequences from closely related species were combined into one ''taxon'' (see Table 1). In three instances (*Salpingoeca amphoridium*, *Mucor racemosus*, and *Trichoplax adhaerens*) one protein gene was missing. These organisms were included but we did not consider taxa with more than one missing gene (see Table 1 for a complete list of the taxa used and GenBank accession numbers). Positions that could not be unambiguously aligned were excluded from the analyses, resulting in a total of 1,151 positions.

A maximum likelihood (ML) analysis of the concatenated (*EF1-α*, *Hsp70*, and *actin*) dataset was performed using Iqpnni (Vinh le and Von Haeseler 2004) using a JTT+Γ model of evolution with rates across sites (eight categories) and a minimum of a 100 iterations ($I = 100$). Statistical support was assessed by Bayesian posterior probability (BPP) and by bootstrapping (100 replicates) in phyml (Guindon and Gascuel 2003) under a JTT+Γ model of evolution (four rates across sites categories). The Bayesian analysis was done using Mr. Bayes 3.0 (Ronquist and Huelsenbeck 2003) by three independent runs of all three partitions (*EF1-α*, *Hsp70*, and *actin*), each one with 1,000,000 Markov Chain Monte Carlo generations with three simultaneous heated (0.2) chains and a covarion model implemented. The WAG+Γ+cov model of evolution was used with rates across sites (four categories) and branch lengths unlinked (i.e. each dataset has its own

Table 1. List of species used in this study with GenBank accession numbers.

Species	*Actin*	*Hsp70*	*EF1-α*
Amoebozoa			
Acanthamoeba culbertsoni	CAA23399[a]	AAU94664	AAU94656
Dictyostelium discoideum	P02577	AAL92992	X55972
Choanoflagellata			
Monosiga ovata	CO435450/ CO435566	AAG45150	AAU94651
Salpingoeca amphoridium	AAY99759	—	AAY99757
Ichthyosporea			
Sphaeroforma arctica	AJ780965	**DQ403166**	AAL87078[b] **DQ403164***
Amoebidium parasiticum	AAU94670	AAU94658	AAU94655
Other unicellular opisthokonts			
Ministeria vibrans	AAU94673	AAU94663	AAU94652
Corallochytrium limacisporum	AAU94671	AAU94661	AAU94653
Nuclearia simplex	AAQ75373	AAU94662	AAU94654
Capsaspora owczarzaki	AY724689	**DQ403165**	**DQ403163**
Fungi			
Candida albicans	CAA34413	CAA82929	XP_717655
Saccharomyces cerevisiae	P60010	NP_011029	AAT92946
Ajellomyces capsulatus	AAA57122	Q00043	P40911
Filobasidiella neoformans	AAW41026	BAD72840	O42671
Mucor racemosus	CAC28274	—	P06805
Rhizopus stolonifer	CAC28295	AAN52150	AAG29040
Chytriomyces confervae	AAU94668	AAU94660	AAU94650
Rhizophlyctis rosea	—	AAU94659	ABB90961
Metazoa			
Geodia cydonium	AAU94666[c]	CAA64441	CAA70221
Trichoplax adhaerens	CAD70272	—	CAD70273
Hydra magnipapillata	AAA29205[d]	Q05944	BAA11471
Nematostella vectensis	AAG61116	AAF12746[e]	BAD02195
Ciona intestinalis	CAC82547	NP001029006	BAB63212
Strongylocentrotus purpuratus	P12431	XP_780020	AAT06191
Gallus gallus	CAA25004	CAA06233	Q90835
Bombyx mori	P04829	BAB92074	P29520
Crassostrea gigas	AAB81845	CAC83009	BAD15289

Sequences generated in this study are in bold. Asterisk (*) indicates that *Sphaeroforma arctica* possess the EFL gene instead of the canonical EF1-α (see text for details).

Sequences obtained from other, closely related species, of those listed are indicated as follows:

[a] *Acanthamoeba castellanii*.
[b] *Ichthyophonus irregularis*.
[c] *Discodermia* sp.
[d] *Hydra vulgaris*.
[e] *Stylophora pistillata*.

α shape parameter and set of branch lengths). The first 50,000 trees were discarded as burn-in.

A preliminary analysis of the *EF1-α* alignment showed that *Sphaeroforma arctica* had an *EF*-like gene instead of the *EF1-α*. Therefore, the *S. arctica* EF1-α gene was not used in the concatenated analysis; instead, as a ''surrogate'' for this gene we substituted the EF1-α homolog of the ichthyosporean *Ichthyophonus irregularis* (see Table 1), a species that is known to be a close relative of *S. arctica* to the exclusion of *Amoebidium parasiticum* (Mendoza et al. 2002). Moreover, an ML analysis of the *EF*-like gene was performed including other published *EF*-like sequences. Taxa and GenBank Accession Number were as follows: *Pleodorina* sp. (BAC67663); *Helicosporidium* sp. (AAV34148); *Heterocapsa triquetra* (AAV34145); *Monosiga brevicollis* (AAK27413); *Proterospongia* (EST project); *Isochrysis galbana* (AAV34146); *Pavlova lutheri* (AAV34147). A phyml tree was performed under a JTT+Γ model of evolution (four categories). Branch support was assessed by bootstrapping (100 replicates).

Pulsed-Field Gel Electrophoresis. Two different concentrations of agarose plugs for PFGE were prepared using the protocol of Eschbach et al. (1991) using 2 L of log-phase culture for a final density of 1.5–2.0 × 10^8 cells. Plugs were run in a CHEF-DR® III Pulsed-Field Electrophoresis System (Bio-Rad Laboratories, Hercules, CA), on 0.9% or 1% agarose (1 × TBE or 1 × TAE) gels in 0.5% TBE buffer, at 14.0 °C. In order to resolve all chromosome sizes, we performed three independent PFGE runs, each one with its own particular conditions. To resolve the smaller chromosomes (i.e. < 1.5 Mb) plugs were run on 1% agarose (1 × TAE), using a 70 h run time, a voltage of 2.5 V/cm, a 500 s switch time, and an angle of 120°. To resolve the medium-sized chromosomes (i.e. > 1.5 to < 3 Mb), plugs were run on 1% agarose (1 × TBE), using a 24-h run time, a voltage of 6 V/cm, a 60–120-s switch time, and an angle of 120°. For the largest chromosomes (i.e. > 3 and < 6 Mb), plugs were run on 0.9% agarose (1 × TAE), using a 65-h run time, a voltage of 2 V/cm, a 20–30 min switch time, and an angle of 106°. Commercial yeast plugs were used as markers (Bio-Rad Laboratories, *Saccharomyces cerevisiae* strain YNN295, *Hansenula wingei* strain YB-4662-VIA, and *Schizosaccharomyces pombe* strain 972 h).

RESULTS

Phylogenetic analysis. The ML analysis of the concatenated dataset performed with all main opisthokonts lineages available (Table 1) yielded a topology identical to that derived by the Bayesian analysis. Our results strongly show that *Capsaspora* is not a nucleariid or an ichthyosporean (Fig. 1). Instead, *Capsaspora* is most likely an independent lineage closely related to metazoans, choanoflagellates, *Corallochytrium*, and *Ministeria*. The ichthyosporeans appear as sister-group to the Metazoa, Choanoflagellata, *Capsaspora*, *Ministeria*, and *Corallochytrium* clade (Fig. 1). The group that includes *Capsaspora*, *Ministeria*, *Corallochytrium*, and choanoflagellates has strong (100% BPP) support in Bayesian analyses but weak (< 50%) in the ML analysis. Finally, our results confirm (and with a strong ML bootstrap and BPP support) that the Nucleariidae are the protists most closely related to Fungi. We further analyzed our tree topology by testing its overall statistical support if more data (namely the SSU rRNA gene, *SSU rRNA*) were added. Unexpectedly, the addition of the *SSU RNA* nucleotide sequences resulted in the failure to recover some of the groups of Fig. 1, such as the one comprising *Capsaspora*, *Ministeria*, *Corallochytrium*, and choanoflagellates, with *Corallochytrium* instead appearing as a sister-group to the fungi+*Nuclearia* clade (data not shown).

***Sphaeroforma arctica* EF-like gene.** Our ML phylogenetic tree of *EFL* genes shows very strongly (100% bootstrap support) that the ichthyosporean *S. arctica EFL* gene groups with the two choanoflagellates homologs (Fig. 2).

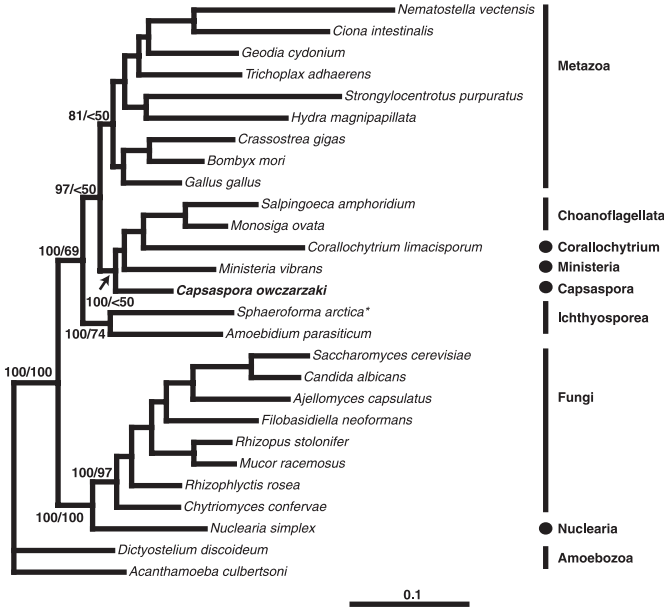

Fig. **1.** Maximum likelihood (ML) tree of the combined (*EF-1α*, *HSP70*, and *actin*) dataset under JTT+Γ model of evolution with rates across sites (eight categories) as inferred by Ipqnni (Vinh le and Von Haeseler 2004). Values above key nodes refer to the Bayesian posterior probabilities and a ML bootstrap support (shown as percentage). * indicates that in these analyses the EF-1α sequence from *Ichthyophonus irregularis* was used as a ''surrogate'' for *Sphaeroforma arctica* (see text for details).

Genome size estimate of *Capsaspora owczarzaki*. We used PFGE analysis to elucidate the number and size of *Capsaspora* chromosomes. Our preliminary analyses revealed that the size diversity of *Capsaspora* chromosomes was too wide to be resolved in one gel. We, therefore, performed three different runs with two

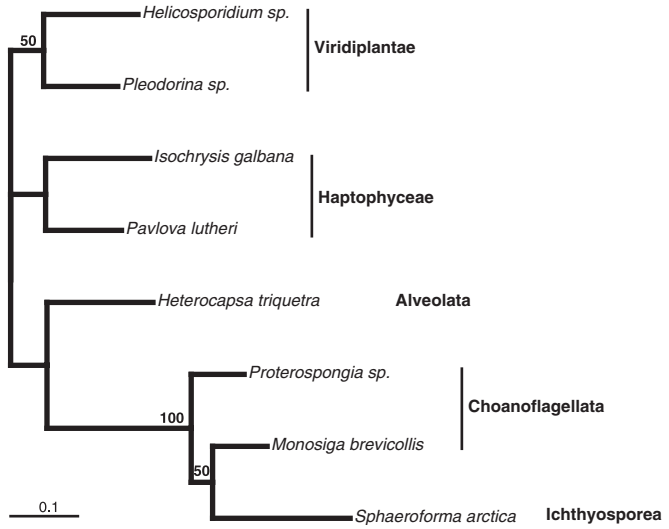

Fig. **2.** Maximum likelihood (ML) tree of the *EF*-like gene under JTT+Γ model of evolution and rates across sites (four categories) performed in phyml (Guindon and Gascuel 2003). Values above key nodes refer to a 100-replicate ML bootstrap (shown as percentage).

different *Capsaspora* DNA concentrations (Fig. 3A–C). The combination of the three runs resolved 12 chromosomes, with sizes estimated by comparison with the size markers as follows: 1: ~ 0.60 Mb; 2: ~ 0.68 Mb; 3: ~ 1.02 Mb; 4: ~ 1.06 Mb; 5: ~ 1.12 Mb; 6: ~ 1.5 Mb; 7: ~ 1.6 Mb; 8: ~ 1.9 Mb; 9: ~ 2.3 Mb; 10: ~ 2.7 Mb; 11: ~ 2.9 Mb; 12: ~ 4 Mb. The conditions used in the first gel, allowed us to exclude the possibility of chromosomes smaller than 0.60 Mb. (Fig. 3A), while the last run confirmed that there appear to be no chromosomes larger than 4 Mb. (Fig. 3C).

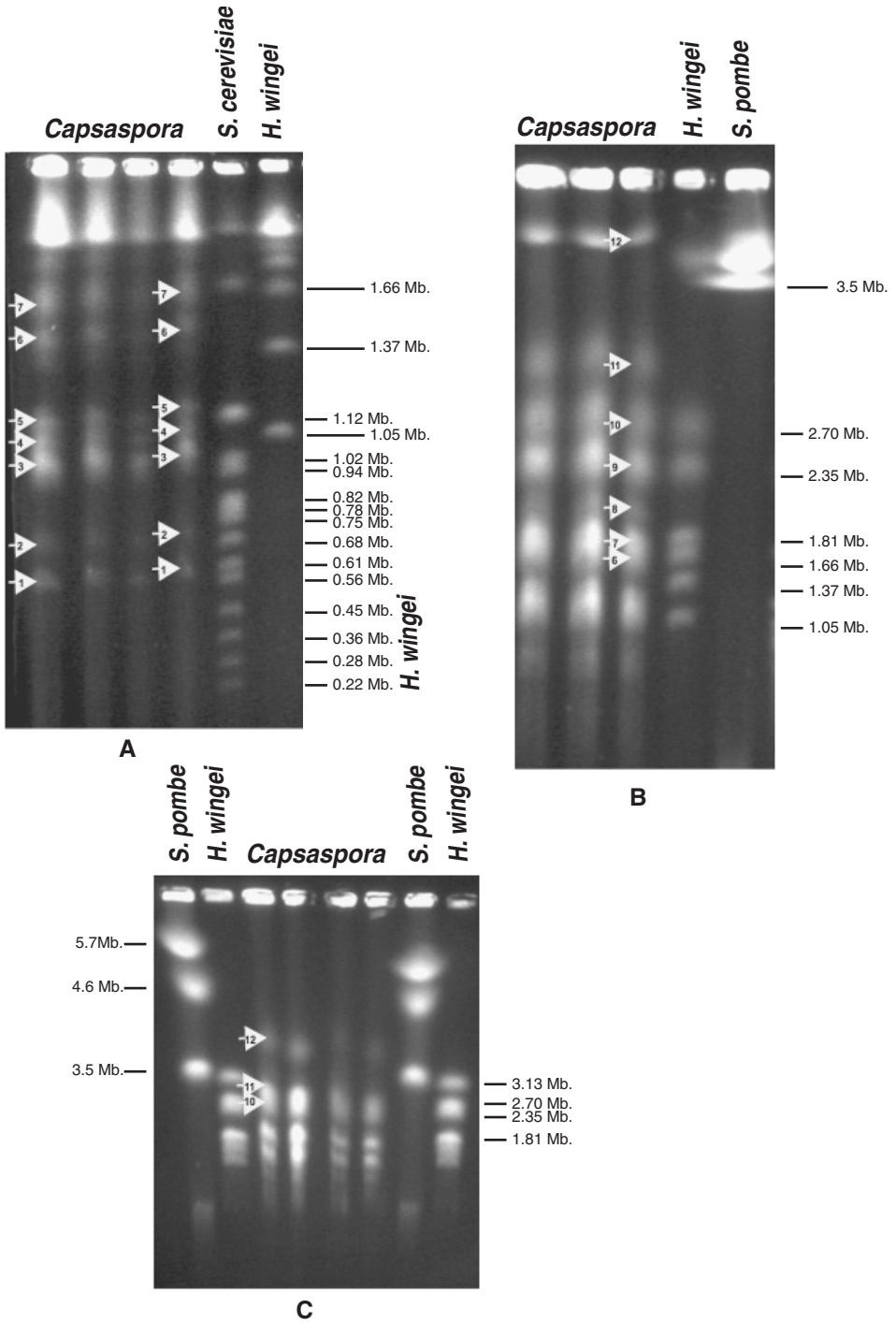

Fig. **3.** Ethidium bromide-stained chromosomes of *Capsaspora owczarzaki* obtained using three different runs of pulsed-field gel electrophoresis (PFGE) with conditions optimized to resolve (**A**) smaller, (**B**) medium-sized, and (**C**) largest chromosomes. **A.** The size standards shown are *Saccharomyces cerevisiae* and *Hansenula wingei*. Putative *Capsaspora* chromosomes 1–7 are indicated by white arrowheads. **B.** The size standards shown are *H. wingei* and *Schizosaccharomyces pombe*. Putative *Capsaspora* chromosomes 6–12 are indicated by white arrowheads. **C.** The size standards shown are *H. wingei* and *S. pombe*. Putative *Capsaspora* chromosomes 10–12 are indicated by white arrowheads. Size estimates (Mb) of *Capsaspora* chromosomes are as follow; 1: ~ 0.60; 2: ~ 0.68; 3: ~ 1.02; 4: ~ 1.06; 5: ~ 1.12; 6: ~ 1.5; 7: ~ 1.6; 8: ~ 1.9; 9: ~ 2.3; 10: ~ 2.7; 11: ~ 2.9; 12: ~ 4.

The number and relative sizes of *Capsaspora* chromosomes were consistent among the three runs (Fig. 3), among the two DNA concentrations used (Fig. 3), and between independent runs and plugs (data not shown). Nevertheless, both the number and sizes should be considered rough estimates. We cannot exclude the possibility of two chromosomes having the same size, and thus appearing as a single band. In fact, some bands are more intense than others. However, our data allow us to provide a rough size estimate of the *C. owczarzaki* genome size probably in the range of 22–25 Mb, with a best guess of 21.38 Mb.

DISCUSSION

***Capsaspora owczarzaki* is an independent unicellular opisthokont lineage.** Our updated phylogenetic analysis of the opisthokonts represents, to our knowledge, the first multi-gene analysis to include several choanoflagellates, several ichthyosporeans, a wide representation of basal fungi and metazoans, plus *Corallochytrium*, *Ministeria*, and *Capsaspora*. The results confirm that *Capsaspora* is indeed not a nucleariid or an ichthyosporean, but an independent lineage closely related to metazoans, choanoflagellates, *Corallochytrium*, and *Ministeria*. Thus, *C. owczarzaki* and also *Ministeria* or *Corallochytrium* may be important lineages to target for comparative genomics aimed at understanding the origin of multicellular animals. Our results show some differences compared with previous studies, especially in the position of ichthyosporeans as sister-group (although not strongly supported in the ML analysis) to *Capsaspora*, metazoans, choanoflagellates, *Corallochytrium*, and *Ministeria*. This may be due to the differences in the taxa used in the different studies; since two ichthyosporeans are included in our analysis versus one in previous multi-gene analyses (Ruiz-Trillo et al. 2004; Steenkamp et al. 2006). Moreover, the nucleariids are again shown to be the sister-group of Fungi as previously shown in single gene (Medina et al. 2003; Nikolaev et al. 2004) and multi-gene phylogenies (Ruiz-Trillo et al. 2004; Steenkamp et al. 2006). However, the specific phylogenetic relationship among ichthyosporeans, metazoans, choanoflagellates, *Corallochytrium*, *Ministeria*, and *Capsaspora* remains to be fully resolved, since statistical support is still weak on some branches (such as those within the clade containing *Ministeria*, *Corallochytrium* and choanoflagellates). The failure to obtain strong support for some of those nodes is probably a sign that the $\sim$ 1,000 sites used do not contain enough phylogenetic information and additional data (more genes) are needed in order to fully resolve the deep branching order within opisthokonts. Finally, the reason for the topological differences obtained upon addition of the *SSU rRNA* nucleotide sequences is unclear, but may result from lack of true convergence of the MCMC procedure in the Bayesian analysis and/or could be related to long-branch effects in the *SSU rRNA* dataset. In any case, this observation is not consistent with other published *rRNA* analyses (Cavalier-Smith and Allsopp 1996) and thus we are inclined to treat it with skepticism.

The ichthyosporean *Sphaeroforma arctica* lacks a canonical *EF1-α*. It has recently been shown that some eukaryotes lack a canonical *EF1-α* gene, but instead posses a similar, but clearly different, gene known as *EF-like* (*EFL*) (Keeling and Inagaki 2004). All our attempts to amplify and sequence the *EF1-α* gene from *Sphaeroforma arctica* cDNA produced only the *EFL* homolog. Moreover, a preliminary analysis of an on-going EST project on this organism reveals the presence of *EFL* sequences but no canonical *EF1-α* homologs. Our failure to obtain a canonical *EF1-α* gene from *S. arctica* cDNA by either PCR or EST sequencing implies that *S. arctica* lacks this gene, and that it may have been replaced by the *EFL* gene. However, we cannot rule out that *S. arctica* may have a canonical *EF1-α* gene and we simply failed to isolate it. In any case, these data show that within the ichthyosporeans, as in fungi and choanoflagellates, there are species with the canonical *EF1-α* gene, such as *Amoebidium parasiticum* and *Ichthyophonus irregularis*, and species with the *EFL* gene, such as the *S. arctica*. How this *EFL* evolved in those organisms and how it replaces the canonical *EF1-α* is still a matter of debate. Keeling and Inagaki (2004) suggested that eukaryote-to-eukaryote lateral gene transfer (LGT) is the most plausible mechanism. Our results show that the *EFL* gene of *S. arctica* is more closely related to the *EFL* gene from choanoflagellates than to the other eukaryotes *EFL* genes. This grouping implies that if LGT is the mechanism, both *S. arctica* and the two choanoflagellates got *EFL* from the same donor or one of those organisms was the donor for the others. To fully understand this process and how *S. arctica* acquired its *EFL* gene, much better taxonomic sampling within the unicellular opisthokonts is needed.

***Capsaspora owczarzaki* has a relatively small genome size.** Our PFGE data has provided the first estimate of a genome size from an opisthokont protist. Our data show that *C. owczarzaki* has a relatively small genome size, in the range of 22–25 Mb, which is roughly double the genome size of yeast. Although the *Capsaspora* nuclear genome is relatively small compared with some other eukaryotes, especially to metazoans, it is not as reduced as might be expected for a parasitic organism. This may be an indication that its metabolism and genomic composition have not yet been dramatically reduced. Animals, by comparison, have much larger genome sizes. For example, the basal cnidarians have genome sizes above 200 Mb (see http://www.genomesize.com for a review). However, poriferans, which are morphologically very simple animals, have been reported to have genome sizes as small as 58 Mb (see http://www.genomesize.com). Thus, the genome size of *C. owczarzaki* may be roughly half that of the simplest metazoans, a promising size if we want to use comparative genomics to investigate the origin of the multicellular animals.

ACKNOWLEDGMENTS

We thank Jacquie de Mestral, Jessica Leigh, Maria J. Barberà, Yuji Inagaki, Karen Li, and Julia Tufts for their technical support. We thank Bjarne Landfald for providing a *Sphaeroforma arctica* culture. IR-T was supported by a long-term postdoctoral fellowship from EMBO (ALTF230-2003) and CIHR (200502MFE-143033-135733). This study was supported by the Genome Atlantic/Genome Canada Large Scale project ''The Protist EST Program'' in which AJR is a co-investigator and an NSERC Discovery grant awarded to AJR.

LITERATURE CITED

Amaral, Z. L. A., Nerad, T. A., O'Kelly, C. J. & Sogin, M. L. 2001. The nucleariid amoebae: more protists at the animal-fungal boundary. *J. Eukaryot. Microbiol.*, **48**:293–297.

Baldauf, S. L. & Palmer, J. D. 1993. Animals and fungi are each other's closest relatives: congruent evidence from multiple proteins. *Proc. Natl. Acad. Sci. USA*, **90**:11558–11562.

Cavalier-Smith, T. & Allsopp, M. T. E. P. 1996. *Corallochytrium*, an enigmatic non-flagellate protozoan related to choanoflagellates. *Eur. J. Protistol.*, **32**:306–310.

Cavalier-Smith, T. & Chao, E. E. 2003. Phylogeny of Choanozoa, Apusozoa, and other Protozoa and early eukaryote megaevolution. *J. Mol. Evol.*, **56**:540–563.

Eschbach, S., Hofmann, C. J., Maier, U. G., Sitte, P. & Hansmann, P. 1991. A eukaryotic genome of 660 kb: electrophoretic karyotype of nucleomorph and cell nucleus of the cryptomonad alga, *Pyrenomonas salina*. *Nucleic Acids Res.*, **19**:1779–1781.

Guindon, S. & Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.*, **52**:696–704.

Hertel, L. A., Bayne, C. J. & Loker, E. S. 2002. The symbiont *Capsaspora owczarzaki*, nov. gen. nov. sp., isolated from three strains of the pulmonate snail *Biomphalaria glabrata* is related to members of the Mesomycetozoea. *Int. J. Parasitol.*, **32**:1183–1191.

Keeling, P. J. & Inagaki, Y. 2004. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1alpha. *Proc. Natl. Acad. Sci. USA*, **101**:15380–15385.

Lang, B. F., O'Kelly, C., Nerad, T., Gray, M. W. & Burger, G. 2002. The closest unicellular relatives of animals. *Curr. Biol.*, **12**:1773–1778.

Medina, M., Collins, A. G., Taylor, J. W., Valentine, J. W., Lipps, J. H., Amaral Zettler, L. & Sogin, M. L. 2003. Phylogeny of Opisthokonta and the evolution of multicellularity and complexity in fungi and metazoa. *Int. J. Astrobiol.*, **2**:203–211.

Mendoza, L., Taylor, J. W. & Ajello, L. 2002. The class Mesomycetozoea: a heterogeneous group of microorganisms at the animal-fungal boundary. *Annu. Rev. Microbiol.*, **56**:315–344.

Nikolaev, S. I., Berney, C., Fahrni, J. F., Bolivar, I., Polet, S., Mylnikov, A. P., Aleshin, V. V., Petrov, N. B. & Pawlowski, J. 2004. The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. *Proc. Natl. Acad. Sci. USA*, **101**:8066–8071.

Owczarzak, A., Stibbs, H. H. & Bayne, C. J. 1980. The destruction of *Schistosoma mansoni* mother sporocysts in vitro by amoebae isolated from *Biomphalaria glabrata*: an ultrastructural study. *J. Invertebr. Pathol.*, **35**:26–33.

Philippe, H., Snell, E. A., Bapteste, E., Lopez, P., Holland, P. W. & Casane, D. 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. *Mol. Biol. Evol.*, **21**:1740–1752.

Ronquist, F. & Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**:1572–1574.

Ruiz-Trillo, I., Inagaki, Y., Davis, L. A., Sperstad, S., Landfald, B. & Roger, A. J. 2004. *Capsaspora owczarzaki* is an independent opisthokont lineage. *Curr. Biol.*, **14**:R946–R947.

Simpson, A. G., Inagaki, Y. & Roger, A. J. 2005. Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of 'primitive' eukaryotes. *Mol. Biol. Evol.*, **23**:615–625.

Snell, E. A., Furlong, R. F. & Holland, P. W. 2001. Hsp70 sequences indicate that choanoflagellates are closely related to animals. *Curr. Biol.*, **11**:967–970.

Steenkamp, E. T. & Baldauf, S. L. 2004. Origin and evolution of animals, fungi and their unicellular allies (Opisthokonta). *In*: Hirt, R. P. & Horner, D. S. (ed.), Organelles, Genomes and Eukaryote Phylogeny: an Evolutionary Synthesis in the Age of Genomics. CRC Press, Boca Raton. p. 109–129.

Steenkamp, E. T., Wright, J. & Baldauf, S. L. 2006. The protistan origins of animals and fungi. *Mol. Biol. Evol.*, **23**:93–106.

Vinh le, S. & Von Haeseler, A. 2004. IQPNNI: moving fast through tree space and stopping in time. *Mol. Biol. Evol.*, **21**:1565–1571.

Received: 03/13/06, 05/09/06; accepted: 06/13/06