Phylogenomics of a new fungal phylum reveals multiple waves of reductive evolution across Holomycota

Luis Javier Galindo1✉, Purificación López-García1, Guifré Torruella1, Sergey Karpov2,3 & David Moreira1✉

Compared to multicellular fungi and unicellular yeasts, unicellular fungi with free-living flagellated stages (zoospores) remain poorly known and their phylogenetic position is often unresolved. Recently, rRNA gene phylogenetic analyses of two atypical parasitic fungi with amoeboid zoospores and long kinetosomes, the sanchytrids Amoeboradix gromovi and Sanchytrium tribonematis, showed that they formed a monophyletic group without close affinity with known fungal clades. Here, we sequence single-cell genomes for both species to assess their phylogenetic position and evolution. Phylogenomic analyses using different protein datasets and a comprehensive taxon sampling result in an almost fully-resolved fungal tree, with Chytridiomycota as sister to all other fungi, and sanchytrids forming a well-supported, fast-evolving clade sister to Blastocladiomycota. Comparative genomic analyses across fungi and their allies (Holomycota) reveal an atypically reduced metabolic repertoire for sanchytrids. We infer three main independent flagellum losses from the distribution of over 60 flagellum-specific proteins across Holomycota. Based on sanchytrids’ phylogenetic position and unique traits, we propose the designation of a novel phylum, Sanchytriomycota. In addition, our results indicate that most of the hyphal morphogenesis gene repertoire of multicellular fungi had already evolved in early holomycotan lineages.

1Ecologie Systématique Evolution, CNRS, Université Paris-Saclay, AgroParisTech, Orsay, France. 2 Zoological Institute, Russian Academy of Sciences, St. Petersburg, Russia. 3 St. Petersburg State University, St. Petersburg, Russia. ✉email: luisjagg92@gmail.com; david.moreira@u-psud.fr
Results and discussion

The new zoosporic fungal phylum Sanchytriomycota. We isolated individual sporangia of *A. gromovi* and *S. tribonematis* (Fig. 1) by micromanipulation and sequenced their genomes after whole-genome amplification (WGA). After thorough data curation (see ‘Methods’), we assembled two high-coverage genome sequences (123.9 x and 45.9 x, respectively) of 10.5 and 11.2 Mbp, encoding 7220 and 9638 proteins, respectively (Table 1). Comparison with a fungal dataset of 290 near-universal single-copy orthologs indicated very high completeness for the two genomes (92.41% for *A. gromovi*; 91.72% for *S. tribonematis*). Only half of the predicted sanchytrid proteins could be functionally annotated using the ‘evolutionary genealogy of genes: non-supervised orthologous groups’ (EggNOG) database (3838 for *A. gromovi*; 4772 for *S. tribonematis*) (Supplementary Data 1). This could be partly due to the fact that, being fast-evolving parasites, many genes have evolved beyond recognition by annotation programmes. However, low annotation proportions are common in Holomycota, including fast-evolving parasites (e.g. only 20 and 52% of the genes of the microsporidia *Nosema parisii* and *Encephalitozoon cuniculi* could be assigned to protein family ( Pfam) domains and Gene Ontology (GO) terms). And also the less fast-evolving metchnikovellids (*Amphianibys sp.*, 45.6%), rozellids (*Rozella allomycis*, 64.9%); *Paramicrosporidium saccamoeba*, 66.7%) and blastocladiomycetes (*Catenaria anguillulae*, 47.5%). Many of the non-annotated genes were unique to sanchytrids as deduced from orthologous gene comparison with 57 other species, including representatives of the other major fungal lineages, and several outgroups (Supplementary Data 2). After clustering of orthologous proteins with OrthoFinder, we identified 1217 that were only present in both *A. gromovi* and *S. tribonematis*. Their analysis using eggNOG resulted in only 93 proteins annotated. The remaining (93.4%) sanchytrid-specific proteins lack functional annotation (Supplementary Data 2).

The two sanchytrid genomes yielded similar sequence statistics (Table 1), but showed important differences with genomes from other well-known zoosporic fungi. They are 4–5 times smaller than those of blastocladiomycetes (40–50 Mb) and average chytrids (~20–101 Mb), an observation that extends to the number of protein-coding genes. Their genome G+C content (~35%) is much lower than that of blastocladiomycetes and most chytrids (40–57%), though some chytrids, like *Anaeromyces robustus*, may have values down to 16.3% . Low G+C content correlates with parasitic lifestyle in many eukaryotes. In Holomycota, low G+C is observed in microsporidian parasites and Neocallimastigomycota, both anaerobic and exhibiting reduced mitochondrion-derived organelles, and in the aerobic parasite *R. allomycis*. Although sanchytrids are aerobic parasites that have similar life cycles compared to those of blastocladiomycetes and chytrids, their smaller genome size and G+C content suggest that they are more derived parasites. This pattern is accompanied by a global acceleration of evolutionary rate (see below), a trend also observed, albeit at a lower extent, in *R. allomycis*. The mitochondrial genomes of *S. tribonematis* and *A. gromovi* showed similar trends. Gene order was highly variable (Supplementary Fig. 1), as commonly observed in Fungi, and their size (24,749 and 27,055 bp, respectively) and G+C content (25.86% and 30.69%, respectively) were substantially smaller than those of most other Fungi. However, despite these signs of reductive evolution, most of the typical core mitochondrial genes were present, indicating that they have functional mitochondria endowed with complete electron transport chains.
To resolve the previously reported unstable phylogenetic position of sanchytrids based on ribosomal RNA (rRNA) genes \(^{33}\), we carried out phylogenomic analyses on a manually curated dataset of 264 conserved proteins (93,743 amino acid positions) \(^{13,49,52}\) using Bayesian inference (BI) and maximum likelihood (ML) with the CAT\(^{53}\) and PMSF\(^{54}\) models of sequence evolution, respectively. Both mixture models are known to alleviate homoplasy and long-branch attraction (LBA) artefacts\(^{22,53}\). We used the same 59 species as above (dataset GBE59), including a wide representation of Holomycota plus two holozoan, two amoebae and one apusomonad as outgroups. BI and ML phylogenomic analyses yielded the same tree topology for major fungal groups with only minor changes in the position of terminal branches (Fig. 2a and Supplementary Fig. 2c), suggesting an LBA artefact on the tree topology. We removed the fast-evolving sanchytrids in a new dataset of 57 species and 93,421 amino acid positions (dataset GBE57; 72 species and 84,949 amino acid positions), which led to chytrids recovering their position as the sister group of all other fungi with full maximum support (BS 100%). However, this position of the root of fungi was as strongly supported as a well-accepted relationship, the monophyly of Dikarya. Similarly, the root of the fungal tree always received very weak support (<26% bootstrap). To confirm the possible impact of LBA on this topology, we removed the fast-evolving sanchytrids (dataset GBE72; 72 species and 84,949 amino acid positions), which led to chytrids recovering their position as the sister group of all other fungi with higher support (BS 91%; Supplementary Fig. 2e), corroborating the LBA induced by the long sanchytrid branch.

Second, we tested the influence of fast-evolving sites by applying a slow-fast approach\(^{58}\) that progressively removed the fastest-evolving sites (in 5% steps) in the 59-species alignment. The monophyly of sanchytrids and Blastocladiomycota obtained maximum support (BS > 99%) in all steps until only 20% of the sites remained, that is, when the phylogenetic signal was too low to resolve any deep-level relationship (Fig. 2b). This relationship was as strongly supported as a well-accepted relationship, the monophyly of Dikarya. Similarly, the root of the fungal tree between chytrids and the rest of the fungi was supported (>90% bootstrap) until only 40% of the sites remained. By contrast, a root between sanchytrids + Blastocladiomycota and the rest of fungi always received very weak support (<26% bootstrap).

We then further tested the robustness of the position of chytrids using alternative topology (AU) tests. For the dataset

---

**Fig. 1 Light microscopy observations of sanchytrid cells.**

- **a-i** Life cycle stages of *Sanchytrium tribonematis*. **f-j** Life cycle stages of *Amoeboradix gromovi*. **a-d**, **f-i** Amoeboid crawling zoospores with filopodia (f) and posterior pseudocilium (pc). **g-i** Zoospores with retracted pseudocilium. **e** Sporangium (sp) with one (e) or two (j) papillae (p) on the host (h) surface. Scale bars: **a-d, f-i**, 5 \(\mu\)m; **e, j**, 10 \(\mu\)m.
### Table 1: Statistics of sanchytrid genome assemblies before and after decontamination and comparison with related zoosporic lineages.

|                      | Sanchytrium tribotrematis | Allomyces robustus | Allomyces macrogynus | Spizellomyces punctatus | Anaeromyces tribonematis | Catenaria anguillulae | Rhizoclosmatium globosum | Allomyces macrogynus | Anaeromyces robustus | Allomyces macrogynus |
|----------------------|---------------------------|--------------------|----------------------|-------------------------|--------------------------|-----------------------|-------------------------|-----------------------|----------------------|-----------------------|
| Genome size (Mb)     | 57.02                     | 51.06              | 11.2                 | 11.2                    | 56                        | 3.1                   | 292.246                 | 12,673                | 12,673               | 12,673               |
| GC%                  | 44.9                      | 51.06              | 34.64                | 34.64                   | 3.1                       | 2.8                   | 929.86                  | 19,447                | 19,447               | 19,447               |
| Number of contigs    | 437                       | 437                | 35,497               | 35,497                  | 35.497                    | 35.497                | 35.497                  | 35.497                | 35.497               | 35.497               |
| N50                  | 292.246                   | 292.246            | 292.246              | 292.246                 | 292.246                   | 292.246              | 292.246                 | 292.246              | 292.246              | 292.246              |
| Predicted proteins   | 155,888                   | 155,888            | 155,888              | 155,888                 | 155,888                   | 155,888             | 155,888                 | 155,888             | 155,888             | 155,888             |
| Mt genome size (bp)  | —                         | 57,473             | —                    | 57,473                  | —                         | —                    | —                      | —                    | —                    | —                    |
| Mt genome GC%        | 29.88                     | 29.88              | 29.88                | 29.88                   | 29.88                     | 29.88                | 29.88                   | 29.88                | 29.88                | 29.88                |

GBE59, these tests did not reject alternative positions for the divergence of Chytridiomycota and Blastocladiomycota + Sanchytridiomycota (p values > 0.05; Supplementary Data 3), which likely reflected the LBA due to the long sanchytrid branch. In fact, the position of Blastocladiomycota at the base of fungi was significantly rejected (p values < 0.05; Supplementary Data 3) after removing the fast-evolving sanchytrid clade (dataset GBE57).

Finally, we compared the results based on the GBE dataset with those based on a different dataset. We decided to use the BMC dataset29 (which includes 53 highly conserved proteins and 14,965 amino acid positions) because it was originally designed to study fast-evolving Holomycota (e.g. Microsporidia), which made it appropriate to deal with the fast-evolving sanchytrids. Using the same taxon sampling of 59 species (BMC59), we recovered the same topology, particularly for the deeper nodes, with full ML ultrafast and conventional bootstrap (BS = 100) and Bayesian PP (PP = 1) supports for both the monophyly of sanchytrids + blastocladiomycetes and chytrids as the sister lineage to all other fungi (Fig. 2c and Supplementary Fig. 2f–h).

The origin of the conspicuous long branch exhibited by sanchytrids is unclear. It has been shown that fast-evolving organisms, including those within Holomycota, tend to lack part of the machinery involved in genome maintenance and DNA repair30,60. To verify if it was also the case in sanchytrids, we searched in the two sanchytrid genomes 47 proteins involved in genome maintenance and DNA repair that have been observed to be missing in several fast-evolving budding yeasts60. Our results confirm that most of these genes are also absent in both sanchytrids (Supplementary Fig. 3). However, they are also missing in the closely related short-branching blastocladiomycete A. macrogynus, suggesting that the long sanchytrid branch is not only due to the absence of these genes.

The relative position of Chytridiomycota or Blastocladiomycota as the first branch to diverge within Fungi has remained a major unresolved question61. If the earliest fungal split occurred ~1 billion years ago30, the phylogenetic signal to infer it may have been largely eroded over time. Likewise, if evolutionary radiations characterized early fungal evolution22, the accumulation of sequence substitutions during early diversification would have been limited. Both factors would explain the difficulty to resolve the deepest branches of the fungal tree so far. Our results, based on an improved gene and taxon sampling, provided strong support for the placement of chytrids as a sister clade to all other Fungi. This solid position of the root on the chytrid branch is additionally consistent with the distribution of so-considered derived characters in Blastocladiomycota, including sporic meiosis, relatively small numbers of carbohydrate metabolism genes and, in some species, hyphal-like apical growing structures (Allomyces) and narrow sporangia exit tubes (e.g. Catenaria spp.)56–64. Despite the use of a large dataset, some branches remained unresolved, in particular the position of Glomeromy-}

---

**Macceevolutionary trends in primary metabolism.** To assess if sanchytrid metabolic capabilities are as reduced as suggested by their small genome sizes, we inferred their metabolic potential in comparison with other major fungal clades as well as other opisthokonts and amoebozoa as outgroup (43 species). We compared the EggNOG38-annotated metabolic repertoires of holomycotan phyla by focusing on 1158 orthologous groups (Supplementary Data 4) distributed in eight primary metabolism categories. Unexpectedly, cluster analyses based on the presence/absence of these genes did not group sanchytrids with canonical
fungi and their closest aphelid relatives (i.e. *Paraphelidium*), but with non-fungal parasites (*R. allomycis*, *Mitosporidium daphniæ* and *P. saccamoebae*) that show evidence of reductive genome evolution[^11,^67] (Fig. 3a), suggesting gene loss-related convergence. An even further metabolic reduction was observed in Neocallimastigomycota, gut-inhabiting symbiotic anaerobic chytrids[^68–^70]. A principal coordinate analysis of the same gene matrix confirmed this result (Fig. 3b).

At a more detailed level, the main differences in the metabolic complement of sanchytrids and of canonical fungi (+*Paraphelidium*) (Supplementary Fig. 4). We further pairwise compared Kyoto Encyclopedia of Genes and Genomes (KEGG)^[^71] orthlogs of sanchytrids against *R. allomycis* and the blastocladiomycete *A. macrognys* (as representative of the sanchytrid closest canonical fungal relatives). The KEGG metabolic maps of *A. gromovi* and *S. tribonematis* contained 1222 and 1418 orthologous groups, respectively, whereas those of *R. allomycis* and *A. macrognys* contained 845 and 4860, respectively (Supplementary Fig. 5a–c). Blastocladiomycetes and sanchytrids shared more similarities, including the maintenance of amino acid and nucleotide metabolism and energy production with a complete electron transport chain, which were largely lost in *Rozella*[^8,^13]. Nonetheless, a reductive trend in energy production pathways could be observed in sanchytrid mitochondria, including the loss of ATP8, one F-type ATP synthase subunit that is also absent or highly modified in several metazoans, including chaetognaths, rotifers, most bivalve molluscs and flatworms[^72,^73]. *Sanchytrium tribonematis* also lacked the NADH dehydrogenase subunit NAD4L (Supplementary Fig. 1), although this loss is unlikely to impact its capacity to produce ATP since *R. allomycis*, which lacks not only ATP8 but also the complete NADH dehydrogenase complex, still seems to be able to synthesize ATP[^9].

Most carbohydrate-related metabolic pathways were retained in sanchytrids and canonical fungi except for the galactose and inositol phosphate pathways, absent in both sanchytrids and *Rozella*. Nonetheless, sanchytrids displayed a rich repertoire of carbohydrate-degrading enzymes (Supplementary Figs. 6–10), most of them being likely involved in the degradation of algal cell walls required for penetration into the host cells[^13,^74,^75]. The most important difference with canonical fungi concerned lipid metabolism, with the steroids and fatty acid metabolism missing in sanchytrids and also in *Rozella*[^9] (Supplementary Fig. 5d–i).

---

[^1]: Fusarium oxysporum
[^2]: Zygomyces
[^3]: Coniophora puteana
[^4]: Cryptococcus neoformans
[^5]: Holomycota
[^6]: Blastocladiomycota
[^7]: Sanchytriomycota
[^8]: Rozella
[^9]: Dikarya
[^10]: Dikarya
[^11]: Dikarya
[^12]: Dikarya
[^13]: Dikarya
[^14]: Dikarya
[^15]: Dikarya
[^16]: Dikarya
[^17]: Dikarya
[^18]: Dikarya
[^19]: Dikarya
[^20]: Dikarya
[^21]: Dikarya
[^22]: Dikarya
[^23]: Dikarya
[^24]: Dikarya
[^25]: Dikarya
[^26]: Dikarya
[^27]: Dikarya
[^28]: Dikarya
[^29]: Dikarya
[^30]: Dikarya
[^31]: Dikarya
[^32]: Dikarya
[^33]: Dikarya
[^34]: Dikarya
[^35]: Dikarya
[^36]: Dikarya
[^37]: Dikarya
[^38]: Dikarya
[^39]: Dikarya
[^40]: Dikarya
[^41]: Dikarya
[^42]: Dikarya
[^43]: Dikarya
[^44]: Dikarya
[^45]: Dikarya
[^46]: Dikarya
[^47]: Dikarya
[^48]: Dikarya
[^49]: Dikarya
[^50]: Dikarya
[^51]: Dikarya
[^52]: Dikarya
[^53]: Dikarya
[^54]: Dikarya
[^55]: Dikarya
[^56]: Dikarya
[^57]: Dikarya
[^58]: Dikarya
[^59]: Dikarya
[^60]: Dikarya
[^61]: Dikarya
[^62]: Dikarya
[^63]: Dikarya
[^64]: Dikarya
[^65]: Dikarya
[^66]: Dikarya
[^67]: Dikarya
[^68]: Dikarya
[^69]: Dikarya
[^70]: Dikarya
[^71]: Dikarya
[^72]: Dikarya
[^73]: Dikarya
[^74]: Dikarya
[^75]: Dikarya

---

**Fig. 2 Phylogenomic analysis of Holomycota.** a) Bayesian inference (BI) phylogenomic tree based on the GBE dataset of 264 conserved proteins (93,743 amino acid positions). The tree was reconstructed using 59 species and the CAT-GTR model and the PMSF approximation of the LG + R8 + C20 model for maximum likelihood (ML). b) Evolution of IQ-TREE ML bootstrap support for Chytridiomycota sister to all other fungi (C + F), Blastocladiomycota + Sanchytriomycota sister to all other fungi (B + F). Sanchytriomycota within Blastocladiomycota (S + B), and the monophyly of Dikarya (Dikarya) as a function of the proportion of fast-evolving sites removed from the GBE59 dataset. Holomycota are highlighted in violet, sanchytrids in pink and outgroup taxa in other colours (green, orange and yellow). All phylogenomic trees can be seen in Supplementary Fig. 2a–e. c) Schematic BI phylogeny showing the results obtained with the BMC dataset of 53 conserved proteins (14,965 amino acid positions) using 59 species (Fungi: violet; Sanchytriomycota: red; outgroup taxa: black) and the CAT-GTR model (BI) and the PMSF approximation of the LG + R7 + C60 model for maximum likelihood (ML). In both trees, branches with support values ≥0.99 BI posterior probability and 99% ML bootstrap are indicated by black dots.
Collectively, our data suggest that, compared to Blastocladiomycota and other fungal relatives, sanchytrids have undergone a metabolic reduction that seems convergent with that observed in the phylogenetically distinct rozellid parasites.

Convergent reductive flagellum evolution in Holomycota. The loss of the ancestral opisthokont single posterior flagellum in terrestrial fungi76,77 is thought to have been involved in their adaptation to land environments78. The number and timing of flagellum losses along the holomycotan branch remain to be solidly established. The flagellum is completely absent in nucleariids12,52, but is found in representatives of all other major holomycotan clades to the exclusion of the highly derived Microsporidia. They include rozellids79, a phelid10 and various canonical fungal groups, namely chytrids70, blastocladiomycetes80, Olpidium21,26 and sanchytrids33,34,39, although the latter are atypical. Sanchytrid amoeboid zoospores have never been observed swimming but were found gliding on solid surfaces via two types of pseudopods: thin filopodia growing in all directions and a broad hyaline pseudopodium at the anterior end (Fig. 1). The posterior flagellum, often described as a pseudocilium, drags behind the cell without being involved in active locomotion33,34. Its basal body (kinetosome) and axoneme ultrastructure differ from that of most flagellated eukaryotes. Instead of the canonical kinetosome with nine microtubule triplets and axonemes with nine peripheral doublets + two central microtubules81, sanchytrids exhibit reduced kinetosomes (nine singlets in S. tribonematis; nine singlets or doublets in A. gromovi) and axonemes (without the central doublet).

Fig. 3 Distribution patterns of primary metabolism genes in Holomycota. a Binary heatmap and b principal coordinate analysis (PCoA) species clustering based on the presence/absence of 1158 orthologous genes belonging to eight primary metabolism Gene Ontology categories across 43 eukaryotic genomes and transcriptomes. Species are colour- and shaped-coded according to their taxonomic affiliation as indicated in the legend. COG presence is depicted in blue and absence is depicted in white.
and only 4 microtubular singlets)\textsuperscript{33,34}. Despite this substantial structural simplification, sanchytrid kinetosomes are among the longest known in eukaryotes, up to 2.2 µm\textsuperscript{33,34}. Such long but extremely simplified kinetosomes have not been reported in any other zoosporic fungi, including Blastocladiomycota\textsuperscript{33,34}. Some of them, including \textit{P. sedebokerense}\textsuperscript{32}, display amoeboid zoospores during the vegetative cycle, with flagellated cells most likely to be gametes\textsuperscript{35–37}.

To better understand flagellar reduction and loss across Holomycota, we analysed 61 flagellum-specific proteins on a well-distributed representation of 43 flagellated and non-flagellated species. Sanchytrids lacked several functional and maintenance flagellar components (Fig. 4a), namely axonemal dyneins, single- and double-headed inner arm dyneins, all intraflagellar transport proteins (IFT) of the group IFT-A and several of the group IFT-B. Sanchytrid kinetosomes have also lost several components of the centriolar structure and tubulins, including Centrin2, which is involved in basal body anchoring\textsuperscript{38}, and Delta and Epsilon tubulins, which are essential for centriolar microtubule assembly and anchoring\textsuperscript{37}. These losses (Fig. 4b) explain why sanchytrids lack motile flagella. Cluster analyses based on the presence/absence of flagellar components (Supplementary Fig. 11) showed sanchytrids at an intermediate position between flagellated and non-flagellated lineages. Therefore, sanchytrids are engaged in an unfinished process of flagellum loss, thereby providing an interesting model to study intermediate steps of this reductive process.

In addition to sanchytrid reduction, between four and six independent flagellar losses have been inferred in Holomycota\textsuperscript{32}. Our new, more robust phylogenetic framework (Fig. 2) allowed us to infer three large independent flagellum losses, plus the ongoing one in sanchytrids (Fig. 4c). These losses occurred at the base of high-rank taxa: nuclearids, Microsporidia and the Zoopagomy- cota + Mucoromycota + Dikarya clade. A possible fourth loss event occurred in \textit{Hyaloraphidium curvatum}, an atypical non-flagellated fungus originally classified as a colourless green alga\textsuperscript{14} and later reclassified within the Monoblepharidomycota\textsuperscript{80,88}. Further analysis will be needed to confirm the loss of flagellar components in this species. In addition, a putative fifth independent loss might have occurred in the \textit{Nephridiophagida}, a clade of fungal parasites of insects and myriapods\textsuperscript{89,90} without clear affinity with established fungal clades\textsuperscript{90}. Recently, a possible relationship to chytrids has been suggested\textsuperscript{91}, although genomic or transcriptomic data to clarify their phylogenetic position are still missing.

\textbf{Fungal ‘vision’ and flagellum exaptation.} Why do sanchytrids retain a non-motile flagellum with a simplified but very long kinetosome? Since the primary flagellar function has been lost in
favour of the amoeboid movement, other selective forces must be acting to retain this atypical structure for a different function in zoospores. In bacteria, the exaptation of the flagellum for new roles in mechanosensitivity92,93 and wetness sensing94 has been documented. Microscopy observations of sanchytrid cultures showed that the flagellum is rather labile and can be totally retracted within the cell cytoplasm, the long kinetosome likely being involved in this retraction capability.33,34 Interestingly, a conspicuously curved rosary chain of lipid globules has been observed near the kinetosome in A. gromovi zoospores, often also close to mitochondria33,34. In the blastocladiomycete Blastocladiella emersonii, similar structures tightly associated with mitochondria are known as ‘side-body complexes’95. Blastocladiella emersonii possesses a unique bacterial type-1-rhodopsin + guanylyl-cyclase domain fusion (BeGC1, 626 amino acids), which, together with a cyclic nucleotide-gated channel (BeCNG1), controls zoospore phototaxis in response to cGMP levels after exposure to green light96. BeGC1 was localized by immunofluorescence on the external membrane of the axoneme-associated lipid droplets, which function as an anchor at the base of the flagellum and control its beating96–98. The BeGC1 fusion and the channel BeCNG1 proteins have also been found in other blastocladiomycetes (A. macrogynus and C. anguillulae). Both A. gromovi and S. tribonematis possessed the BeGC1 fusion (532 and 535 amino acids, respectively) and the gated channel BeCNG1 (Supplementary Fig. 12a–c). Therefore, this fusion constitutes a shared trait in Blastocladiomycota and Sanchytriomycota. Despite some ultrastructural differences and the need for functional studies to confirm their role, the presence of lipid threads in the vicinity of the kinetosome and mitochondria, together with the BeGC1 and BeCNG1 homologues, suggest the existence of a comparable light-sensing organelle in Amoeobadix and Sanchytrium. We hypothesize that, as in B. emersonii, the sanchytrid reduced flagellum could be involved in phototactic response, at least as a structural support for the lipid droplets. Interestingly, sanchytrids showed considerably shorter branches in rhodopsin and guanylyl-cyclase domain phylogenetic trees (Supplementary Fig. 12a, b) than in multigene phylogenies (Fig. 2), indicating that these proteins (and their functions) are subjected to strong purifying selection as compared to other proteins encoded in their genomes.

Since rhodopsins capture light by using the chromophore retinal99, we looked for the carotenoid (β-carotene) biosynthesis enzymes96,100 necessary for retinal production. Surprisingly, the enzymes involved in the classical pathway (bifunctional lycopene cyclase/phytene synthase, phytoene dehydrogenase and carotenoid oxygenase)96,100 were missing in both sanchytrid genomes, suggesting that they are not capable of synthesizing their own retinal (Supplementary Data 5). We only detected two enzymes (isopenetyl diphasphate isomerase and farnesyldiphasphate synthase) that carry out early overlapping steps in the biosynthesis of both sterol and carotenoids. By contrast, the β-carotene biosynthesis pathway is widely distributed in Fungi, including chytrids and blastocladiomycetes (Alomyces and Blastocladiella96,101). Therefore, sanchytrids, like most heterotrophic eukaryotes, seem unable to synthesize β-carotene and must obtain carotenoids through their diet100. Indeed, we detected all carotenoid and retinal biosynthesis genes in the transcriptome of the yellow-brown alga Trinobema gayanum (Supplementary Data 5), which is their host and likely retinal source during infection.

**Evolution of multicellularity in Holomycota.** Fungal multicellularity results from connected hyphae102. Diverse genes involved in hyphal multicellularity were present in the ancestors of three lineages of unicellular fungi (Blastocladiomycota, Chytridiomycota and Zoopagomycota; BCZ nodes)103. To ascertain whether they were also present in other deep-branching Holomycota with unicellular members, we reconstructed the evolutionary history of 619 hyphal morphogenesis proteins105, which were grouped into ten functional categories (see ‘Methods’, Supplementary Data 6 and Supplementary Fig. 13). Our results showed that most hyphal morphogenesis genes were not only present in the last common fungal ancestor but also in other unicellular holomycotan relatives, indicating that they evolved well before the origin of fungal multicellularity (Fig. 5a). This pattern could be observed for all functional categories with the clear exception of the adhesion proteins, most of which only occur in Dikarya (Supplementary Fig. 13), reinforcing previous conclusions that adhesion proteins played a marginal role in the early evolution of hyphae103 but highlighting their crucial role in the current (fruiting body-producing) hyphal-based multicellularity.

The common ancestor of sanchytrids and blastocladiomycetes possessed a high percentage of hyphae-related proteins (88.4%, node B in Fig. 5a), although this ancestral repertoire became secondarily reduced in sanchytrids (66.6%). Likewise, many yeasts, which are also secondarily reduced organisms, retained most of the genetic repertoire needed for hyphal development. Many of these proteins were also present in nuclearids, Rozellidae–Microsporidia and Aphelida (nodes N, R and A in Fig. 5a and Supplementary Data 7). Consequently, clustering analysis based on the presence/absence of hyphal morphogenesis proteins did not clearly segregate unicellular and multicellular lineages (Fig. 5b) and retrieved very weak intragroup correlation (Fig. 5c). Our results extend previous observations of hyphal morphogenesis genes from Fungi103 to much more ancient diversifications in the Holomycota. The holomycotan ancestor already possessed a rich repertoire of proteins, notably involved in ‘actin cytoskeleton’ and ‘microtubule-based transport’, which were later recruited for hyphal production (Supplementary Data 7). Most innovation concerned the proteins involved in the ‘cell wall biogenesis/remodelling’ and ‘transcriptional regulation’ functional categories, which expanded since the common ancestor of Aphelida and Fungi. This pattern is consistent with the enrichment of gene duplications in these two categories in all major fungal lineages103. Nevertheless, genome and transcriptome data remain very scarce for Aphelida and we expect that part of these duplications will be inferred to be older when more data for this sister lineage of Fungi become available.

**Conclusions** We generated the first genome sequence data for the two known species of sanchytrids, a group of atypical fungal parasites of algae. The phylogenetic analysis of two independent datasets of conserved proteins showed that they form a new fast-evolving fungal phylum, the Sanchytriomycota, sister to the Blastocladiomycota. Our phylogenetic analyses also provided strong support for Chytridiomycota being the sister group to all other fungi. Sanchytrids have a complex life cycle that includes a flagellated phase (zoospores) with non-motile flagella that are engaged in an ongoing reductive process. The inclusion of sanchytrids and a wide taxon sampling of fungi in our multigene phylogeny allowed the inference of three large independent flagellum loss events across Holomycota. Interestingly, the sanchytrid residual flagellum endowed with a long kinetosome might represent an expatation of this structure as support for a lipid organelle probably involved in light sensing. Our taxon-rich dataset of deep-branching Holomycota also provided evidence for a very ancient origin of most genes related to hyphal
morphogenesis, well before the evolution of the multicellular fungal lineages.

Taxonomic appendix

Sanchytirymycota phy1. nov. Monocentric thallus, epibiotic; usually amoeboid zoospores with longest-known kinetosome in fungi (1–2 µm) and immobile pseudocell; centrosome in sporangium with two centrioles composed of nine microtubular singlets.

Index Fungorum ID: IF558519

Class Sanchytirymycetes (Tedersoo et al., 2018)¹⁰⁴ emend.

Order Sanchytiriales (Tedersoo et al., 2018)¹⁰⁴ emend.

Family Sanchytiriacae (Karpov et al., 2017)³⁹ emend.

Amoeboid zoospores with anterior lamellipodium producing subfilopodia, and lateral and posterior filopodia; with (rarely without) posterior pseudocell; kinetosome composed of nine microtubule singlets or singlets/doublets, 1–2 µm in length. Zoospores attach to the algal cell wall, encyst and penetrate the host wall with a short rhizoid. Interphase nuclei in sporangia have a centrosome of two centrioles composed of nine microtubular singlets. Predominantly parasites of freshwater algae.

Type: Sanchytium (Karpov et al., 2017)³⁹ emend. Karpov et al. (2019)³⁴.

Parasite of algae. Epibiotic, spherical to ovate, sporangia with one (rarely more) discharge papillae. Amoeboid zoospores with anterior lamellipodium; with (rarely without) pseudocell; kinetosomes composed of nine single microtubules 1–1.2 µm in length. Interphase nuclei in sporangia have a centrosome with two orthogonal centrioles composed of nine microtubular singlets and with an internal fibrillar ring.

Type: Sanchytium tribonematis (Karpov and Aleoshin, 2017)³⁹ emend. Karpov et al. (2019)³⁴.

Sanchytium tribonematis (Karpov and Aleoshin, 2017)³⁹ emend. Karpov et al. (2019)³⁴.

Fig. 5 Hyphae-related genes across Holomycota. a Cladogram of Holomycota depicting the phylogenetic relationships according to the GBE59 phylogenomic reconstruction. Bubble size on the nodes and tips represents the total number of reconstructed ancestral and extant hyphal multicellularity-related proteins. CBZ (Chytriomyces, Blastocladiomycota and Zoopagomyces) and NRA (nucleoids, RozellidaeMicrosporida and Aphelidae) nodes are indicated with letters within the corresponding bubbles. Unicellular lineages are highlighted in colours according to their taxonomic affiliation (bottom to top: Amoebobazoa [yellow], Apusomonadida [light green], Holozo [dark green], biotids [light blue], RozellidaeMicrosporida [purple], Aphelida [pink], Chytriomyces [violet], Sanchytium–Blastocladiomycota [red] and Zoopagomyces [orange]). b Presence/absence heatmap of 619 hyphal morphogenesis proteins in 59 unicellular (taxa are colour-highlighted as in a) and multicellular (not colour-highlighted) eukaryotic proteomes. Gene presence is depicted in blue and absence in white. c Heatmap clustered by similarity showing the correlation between unicellular (taxa are colour-highlighted as in a) and multicellular (not colour-highlighted) taxa according to the presence/absence of hyphal morphogenesis gene proteins.
Round to ovate smooth sporangium, ~10 µm diameter, without or with one discharge papilla; sessile on the algal surface. Slightly branched rhizoid, almost invisible inside host. Amoeboid zoospores 5.4–3.3 µm (maximum) with anterior lamellipodium producing sublifodia, and lateral and posterior filopodia; normally with posterior pseudocilium up to 5 µm in length supported by up to four microtubules.

Amano et al. (2018)89.

Zooporous fungus with monocentric, epibiotic sporangia and amoeboid zoospores having posterior pseudocilium that emerges from long kinetosome (ca. 2 µm) composed of microtubular singlets or doublets.

Type species: A. gromovi (Karpov et al., 2018)89

Amano et al. (2018)89.

Amano et al. (2018)89.

Methods

Biological material. Sanphytrium tribonematis strain X-128 and A. gromovi strain X-113, isolated from freshwater sampling locations in Russia,39,40, were maintained in culture with the freshwater yellow-green alga T. gayanum Pasch. strain 20 CALU as host.39 The algal host was grown in mineral freshwater medium at room temperature under white light. After inoculation with Sanphytrium or Amanoeradix, cultures were incubated for 2 weeks to reach the maximum infection level. We then collected both individual zoospores and sporangia full of moving zoospores by micromanipulation with an Eppendorf PatchMan NP2 micromanipulator using 19 µm VacuTip micropipettes (Eppendorf) on an inverted Leica DIIl3000 B microscope. Sporangia were separated from the algal host cells using a microblade mounted on the micromanipulator. Zoospores and sporangia were washed twice in clean sterile water drops before storing them into individual tubes for further analyses.

Whole-genome amplification and sequencing. DNA extraction from single zoospores and sporangia was done with the PicPure kit (Thermo Fisher Scientific) according to the manufacturer’s protocol. WGA was carried out by multiple displacement amplification with the single-cell REPLI-g kit (Qiagen). DNA amplification was quantified using a Qubit fluorometer (Life Technologies). We retained WGA products that yielded high DNA concentrations. As expected, WGA from sporangia (many zoospores per sporangium) yielded more DNA than individual zoospores and was selected for sequencing (K1-9,WGA for A. gromovi; SC-2,WGA for S. tribonematis). TrueSeq paired-end, single-cell libraries were prepared from these samples by sequencing on a HiSeq 2500 Illumina instrument. The predicted protein sequences were searched by BLASTp against two predicted yellow-green algae proteomes inferred from the T. gayanum transcriptome13 and the sanchytrids transcriptome33. We carried out statistical multivariate analyses to get insights into the metabolic capabilities of sanchytrids in
comparison with other Holomyctoa. We searched in both sanchytrid 1206 egg-
NOG orthologous groups corresponding to eight primary metabolism categories
The raw sequence data and assembled genomes generated in this study have been deposited at the National Center for Biotechnology Information (NCBI) sequence databases under Bioproject accession codes PRJNA66893 and PRJNA66894. Additional data generated in this study (including alignments, phylogenetic trees and assembled genomes) are available in the Fungal Public database [https://www.ncbi.nlm.nih.gov/nuccore/ and https://www.ncbi.nlm.nih.gov/genome/], the JGI genome database [https://www.jgi.doe.gov/portal/], the AZY database [cazy.org] and the mycCLAP database [https://mycoclap.fungalgenomics.ca/mycCLAP/]; for more details see Supplementary Data 8.

Received: 29 November 2020; Accepted: 29 July 2021; Published online: 17 August 2021

References

1. Cavalier-Smith, T. A revised six-kingdom system of life. Biol. Rev. Camb. Philos. Soc. 73, 203–266 (1998).
2. Douzery, E. J. P., Snell, E. A., Baptiste, E., Deluc, F. & Philippe, H. The timing of early eukaryotic diversification: Evidence from rDNA ribosomal molecules. Proc. Natl Acad. Sci. USA 101, 15386–15391 (2004).
3. Parfrey, L. W., Lahr, D. J. G., Knoll, A. H. & Katz, L. A. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc. Natl Acad. Sci. USA 108, 13624–13629 (2011).
4. Emes, L., Sharpe, S. C., Brown, M. V. & Roger, A. J. On the age of eukaryotes: Evaluating evidence from fossils and molecular clocks. Cold Spring Harb. Perspect. Biol. 6, 1–16 (2014).
5. Bebee, M. L., James, T. Y. & Strullu-Derrien, C. Early diverging fungi: diversity and impact at the dawn of terrestrial life. Ann. Rev. Microbiol. 71, 41–67 (2017).
6. Spataro, F. J. et al. The Fungal Tree of Life: from molecular systematics to genome-scale phylogenies. Microbiol. Spectr. 5, 1–32 (2017).
7. Lara, E., Moreira, D. & López-Garcia, P. The environmental clade LKM11 and Rozella form the deepest branching clade of fungi. *Protist* **161**, 116–121 (2010).
8. Jones, M. D. M. et al. Discovery of novel intermediate forms redefines the Fungal Tree of Life. *Nature* **474**, 200–203 (2011).
9. James, T. Y. et al. Shared signatures of parasitism and phylogenomics unite cryptomycota and microsporidia. *Curr. Biol.* **23**, 1548–1553 (2013).
10. Bidartondo, M. I. et al. Morphology, phylogeny, and ecology of the aphelids (Aphelidea, Opisthokonta) and proposal for the new superphylum Opisthosphorida. *Front. Microbiol.* **5**, 112 (2014).
11. Haag, K. L. et al. Evolution of a morphological novelty occurred before genome compaction in a lineage of extreme parasites. *Proc. Natl Acad. Sci. USA* **111**, 15480–15485 (2014).
12. Bass, D. et al. Clarifying the relationships between Microsporidia and Cryptomycota. *J. Eukaryot. Microbiol.* **56**, 755–782 (2018).
13. Torruella, G. et al. Global transcriptome analysis of the aphelid *Paraphelodium tribonematis* supports the phagotrophic origin of fungi. *Commun. Biol.* **1**, 1–11 (2018).
14. Ustinova, I., Krienitz, L. & Huss, V. A. R. Hyaloraphidium curvatum is not a fungus, but a member of the DRIPs. *Protist* **151**, 253–262 (2000).
15. Powell, M. J. Looking at mycology with a Janus face: a glimpse at the ultrastructure of *Sanchytrium tribonematis* and saprotrophic species with highly intronized rDNA. *Front. Microbiol.* **10**, 267–270 (1991).
16. Youssif, N. H. et al. The genome of the anaerobic fungus *Ophoromyces sp.* strain c1a reveals the unique evolutionary history of a remarkable plant biomass degrader. *Appl. Environ. Microbiol.* **79**, 4620–4634 (2013).
17. Videvall, E. Plasmodium parasites of birds have the most AT-rich genes of eukaryotes. *Microb. Genom.* **4**, 202–211 (2018).
18. Chen, Yping et al. Genome sequencing and comparative genomics of honey bee microsporidia, *Nosema apis* reveal novel insights into host-parasite interactions. *BMJ Genomics* **14**, 451 (2013).
19. Thomarat, F., Vivarés, C. P. & Gouy, M. Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. *J. Mol. Evol.* **59**, 78–790 (2004).
20. Mikhailov, K. V., Simdyanov, T. G. & Aleoshin, V. V. Genomic survey of a hyperparasitic microsporidian *Amphialys sp.* (*Metchnikovellidiaceae*). *Genome Biol. Evol.* **9**, 454–467 (2017).
21. Galindo, L. J. et al. Evolutionary genomics of *Metchkovella incurvata* (*Metchkovellidaceae*): an early branching microsporidium. *Genome Biol. Evol.* **10**, 2736–2748 (2018).
22. Aguilera, G. et al. High variability of mitochondrial gene order among fungi. *Genome Biol. Evol.* **6**, 451–465 (2014).
23. Galindo, L. J. et al. Combined cultivation and single-cell approaches to the phylogenomics of nuclearid amoebae, close relatives of fungi. *Philos. Trans. R. Soc. B* **374**, 20190094 (2019).
24. Bartolomucci, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095–1109 (2004).
25. Wang, H. C., Minh, B. Q., Susko, E. & Roger, A. J. Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenetic estimation. *Syst. Biol.* **67**, 216–235 (2018).
26. Leipe, D. D., Gunderson, J. H., Nerad, T. A. & Sogin, M. L. Small subunit ribosomal RNA + of *Hexamita infausta* and the quest for the first branch in the eukaryotic tree. *Biochem. Parasitol.* **59**, 41–48 (1993).
27. Kamaishi, T. et al. Complete nucleotide sequences of the genes encoding translation elongation Factors 1 and 2 from a microsporidian parasite, *Glaeocloga pleoglossis*: implications for the deepest branching of eukaryotes. *J. Bacteriol.* **120**, 1095–1103 (1996).
28. Philippe, H. et al. Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. *Proc. R. Soc. B* **267**, 1213–1221 (2000).
29. Brinkmann, H. & Philippe, H. Archaea sister group of Bacteria? Indications from tree reconstruction artifacts in ancient phylogenies. *Mol. Biol. Evol.* **16**, 817–825 (1999).
30. Capella-Gutiérrez, S., Marcati-Houben, M. & Gabaldón, T. Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. * BMC Biol.* **10**, 47 (2012).
31. Steenwyk, I. L. et al. Extensive loss of cell-cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts. *PloS Biol.* **17**, e0002355 (2019).
32. Naranjo-Ortiz, M. A. & Gabaldón, T. Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. *Biol. Rev.* **94**, 2101–2137 (2019).
33. Vargas, M. M., Aronson, J. M. & Roberson, R. W. The cytoplasmic organization of hyphal tip cells in the fungus *Allomyces macrognosus*. *Protozoa* **176**, 43–52 (1993).
34. Stajich, J. E. et al. The Fungi. *Curr. Biol.* **19**, 840–845 (2009).
35. Archibald, J. M., Simpson, A. G. B. & Slomovits, C. H. Handbook of the Protists. *Handbook of the Protists* (Springer International Publishing, 2017).
36. Field, K. J. et al. First evidence of mutualism between ancient plant lineages (*Haplopappus lividus*) and *Mucoromycotina* fungi and its response to simulated Palaeozoic changes in atmospheric CO2. *N. Phytol.* **205**, 743–756 (2015).
129. Wickstead, B. & Gull, K. Evolutionary biology of dyneins. Dyneins 88–121. 
https://doi.org/10.1016/B978-0-12-382004-4 (2012).

130. Van Dam, T. J. P. et al. Evolution of modular intraflagellar transport from a 
coatomer-like progenitor. Proc. Natl Acad. Sci. USA 110, 6943–6948 (2013).

131. Kubicek, C. P., Starr, T. L. & Glass, N. L. Plant cell wall - degrading enzymes 
and their secretion in plant-pathogenic fungi. Annu. Rev. Phytopathol. 52, 
427–451 (2014).

132. Sista Kameshwar, A. K. & Qin, W. Systematic review of publicly available non-
Dikarya fungal proteomes for understanding their plant biomass-degrading 
and bioremediation potential. Bioreour. Bioprocess. 6, 30 (2019).

133. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: a 
fast and effective stochastic algorithm for estimating maximum-likelihood 
phylogenies. Mol. Biol. Evol. 32, 268–274 (2015).

134. Csurös, M. Count Evolutionary analysis of phylogenetic profiles with 
parsimony and likelihood. Bioinformatics 26, 1910–1912 (2010).

Acknowledgements
We thank the UNICELL single-cell genomics platform (https://www.deemteam.fr/en/
unicell). This work was funded by the European Research Council Advanced Grants 
ProtistWorld (No. 322669, to P.L.-G.) and Plast-Evol (No. 787904, to D.M.) and the 
Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie 
ITN project SINGEK (http://www.singek.eu/), grant agreement no. H2020-MSCA-ITN-
2015-675752, to L.J.G. and P.L.-G.). The contribution by S.K. was supported by the RSF 
Grant No. 16-14-10302. We also thank Christina Cuomo and the Broad Institute for 
allowing us to use the unpublished fungal genome sequences in this study.

Author contributions
P.L.-G. and D.M. conceived and supervised the study. S.K. and D.M. prepared the biological 
material. L.J.G., G.T. and D.M. analysed the sequence data. L.J.G., P.L.-G. and D.M. supervised 
the study and wrote the manuscript, which was edited and approved by all authors.

Competing interests
The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material 
available at https://doi.org/10.1038/s41467-021-25308-w.

Correspondence and requests for materials should be addressed to L.J.G. or D.M.

Peer review information Nature Communications thanks Timothy James, László Nagy, 
and Xing-Xing Shen for their contributions to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in 
published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons 
Attribution 4.0 International License, which permits use, sharing, 
adaptation, distribution and reproduction in any medium or format, as long as you give 
appropriate credit to the original author(s) and the source, provide a link to the Creative 
Commons license, and indicate if changes were made. The images or other third party 
material in this article are included in the article’s Creative Commons license, unless 
indicated otherwise in a credit line to the material. If material is not included in the 
article’s Creative Commons license and your intended use is not permitted by statutory 
regulation or exceeds the permitted use, you will need to obtain permission directly from 
the copyright holder. To view a copy of this license, visit http://creativecommons.org/
licenses/by/4.0/.

© The Author(s) 2021