MBE Advance Access published April 6, 2012

Group I Intron–Mediated *Trans*-splicing in Mitochondria of *Gigaspora rosea* and a Robust Phylogenetic Affiliation of Arbuscular Mycorrhizal Fungi with Mortierellales

Maryam Nadimi,†,[1] Denis Beaudet,†,[1] Lise Forget,[2] Mohamed Hijri,[1] and B. Franz Lang*,[2]

[1]Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, Université de Montréal, Montréal, Québec, Canada

[2]Département de Biochimie, Centre Robert-Cedergren, Université de Montréal, Montréal, Québec, Canada

†These authors contributed equally to this work.

***Corresponding author:** Email: franz.lang@umontreal.ca.

Associate editor: Andrew Roger

Research article

Abstract

Gigaspora rosea is a member of the arbuscular mycorrhizal fungi (AMF; Glomeromycota) and a distant relative of *Glomus* species that are beneficial to plant growth. To allow for a better understanding of Glomeromycota, we have sequenced the mitochondrial DNA of *G. rosea*. A comparison with *Glomus* mitochondrial genomes reveals that Glomeromycota undergo insertion and loss of mitochondrial plasmid-related sequences and exhibit considerable variation in introns. The gene order between the two species is almost completely reshuffled. Furthermore, *Gigaspora* has fragmented *cox1* and *rns* genes, and an unorthodox initiator tRNA that is tailored to decoding frequent UUG initiation codons. For the fragmented *cox1* gene, we provide evidence that its RNA is joined via group I–mediated *trans*-splicing, whereas *rns* RNA remains in pieces. According to our model, the two *cox1* precursor RNA pieces are brought together by flanking *cox1* exon sequences that form a group I intron structure, potentially in conjunction with the *nad5* intron 3 sequence. Finally, we present analyses that address the controversial phylogenetic association of Glomeromycota within fungi. According to our results, Glomeromycota are not a separate group of paraphyletic zygomycetes but branch together with Mortierellales, potentially also Harpellales.

Key words: arbuscular mycorrhizal fungi (AMF), mitochondrial genome, intron evolution, phylogeny, tRNA structure, genetic code.

Introduction

Arbuscular mycorrhizal fungi (AMF) is a group of ubiquitous soil-borne fungi that form symbiotic associations with the majority of vascular plants (Parniske 2008). AMF are obligate biotrophs, that is, they are unable to grow without a host plant that provides them with carbohydrates; in turn, AMF transfer nutrients such as phosphate to the plant (reviewed in Strack et al. 2003). At the cellular level, AMF are characterized by the formation of large, multinucleate hyphae, and asexual spores (e.g., Marleau et al. 2011). Apparently, the genetic segregation of the hundreds of distinct nuclei that are present in these species does not follow canonical but rather population rules, and recent analyses demonstrate substantial sequence variation in certain nuclear genes (Hijri and Sanders 2005; Croll and Sanders 2009; Croll et al. 2009; Boon et al. 2010). It is therefore no surprise that the *Glomus irregulare* nuclear genome project has turned into a sequence assembly nightmare (Martin et al. 2008). Yet in stark contrast, the first complete *Glomus* mitochondrial DNAs (mtDNAs) that have been deciphered recently by 454 sequencing are homogeneous in sequence (Lee and Young 2009 and GenBank #FJ648425; Bullerwell CE, Forget L, Lang BF, unpublished data), that is, genetic segregation of mtDNA is as effective in *Glomus* as in other fungi. In these two cases, long homopolymer stretches that introduce systematic pyrosequencing error are surprisingly absent. In other, more A+T-rich mtDNAs, however, we have observed intolerable levels of 454 sequence error (close to one per 1 kbp sequence on average in a heterolobosean amoeba; Bullerwell CE, Forget L, Lang BF, unpublished data) causing frameshifts in several protein-coding genes. It therefore remains advisable to carefully examine homopolymer-rich sequences for potential error, for instance by resequencing with Sanger technology.

The taxonomic and phylogenetic identity of AMF have been, and continues to be, controversial. Initially assigned to zygomycetes, a fungal taxon that is strongly suspected to be paraphyletic (e.g., Schwarzott et al. 2001; Seif et al. 2006; Hibbett et al. 2007; Liu, Leigh, et al. 2009), AMF have been recently moved into a separate fungal phylum, Glomeromycota (Hibbett et al. 2007). Yet the underlying published phylogenies are controversial and often lack significant statistical support, either due to a limited amount of sequence data (based on one or only few gene sequences), poor taxon sampling, or a combination of both. In many instances, analyses further suffer from potential phylogenetic artifacts such as long-branch attraction (e.g., Felsenstein 1978).

A phylogenetic data set of complete mtDNA sequences is currently restricted to two *G. irregulare* isolates (Lee and

Young 2009 and GenBank #FJ648425) having identical coding sequences. Likewise, a previous phylogenomic analysis with a large number of nuclear genes (Liu, Leigh, et al. 2009) had only limited taxon sampling. Accordingly, phylogenetic analyses with both mitochondrial and nuclear genome data have provided only a tentative answer to the question of where AMF belong within Fungi. In some cases, they both show a weak affinity of Mortierellales with Glomeromycota (Lee and Young 2009; Liu, Leigh, et al. 2009), unsupported however by strict statistical analysis (such as the AU test, Shimodaira 2002). An updated, comprehensive fungal phylogenomic analysis with nuclear sequence data published in 2011 (Ebersberger et al. 2012) nicely summarizes the confusing state of the art, commenting that "at the moment, available data do not allow to confidently attach glomeromycetes to the phylogenetic backbone of the fungi." In the latter phylogenomic analysis, Mortierellales are shown separate from Mucorales, that is, excluding Mortierellales from a monophyletic taxon Mucoromycotina favored by others and in contradiction to conclusions reached in a previous phylogenomic analysis (Liu, Leigh, et al. 2009). Evidently, better taxon sampling of genomic data sets is required to resolve these questions, in particular by adding to both mitochondrial and nuclear gene data sets members of Mortierellales, and AMF lineages that are distant from Glomeraceae.

Our rationale for sequencing the *Gigaspora rosea* mtDNA is that Gigasporaceae are at a large evolutionary distance to *Glomus* species, with clearly distinct morphological characteristics. Gigasporaceae form auxiliary cells in the extraradical mycelium, and giant spores that are usually larger than 200 μm and visible to the naked eye. These atypical spores are formed individually from hyphae and contain funnel-shaped hyphal attachments that extend from a specialized bulbous sporogenous cell. Another motive for sequencing mtDNAs from additional Glomeromycota relates to the origin and distribution of introns, other endonuclease-based mobile elements, and mitochondrial plasmid-like inserts in mtDNAs that are frequent in *Glomus* species (Lee and Young 2009; Lang BF, unpublished data). In fact, one of the group I introns of a glomeromycotan *cox1* gene might have been transmitted to plant mitochondria (Vaughn et al. 1995; Adams et al. 1998; Seif et al. 2005; Lang and Hijri 2009), potentially as part of symbiotic plant–AMF interactions. Group I introns are mobile genetic elements that may insert into intron-less gene copies (i.e., with respect to a given intron insertion point) by a process called intron homing (Colleaux et al. 1986; Lambowitz and Zimmerly 2004). They are further characterized by an RNA that folds into a distinct secondary structure, consisting of up to nine base-paired helical regions (P1–P9) that are involved in bringing respective exons into close proximity and in the splicing reaction itself (Anziano et al. 1982; Michel et al. 1982; Waring et al. 1982). In many but not all cases (Lang et al. 2007), group I introns are self-splicing in vitro, that is, the RNA is a ribozyme capable of catalyzing its own excision from precursor RNA (Cech et al. 1983; Jacquier and Rosbash 1986; Schmelzer and Schweyen 1986; Van der Veen et al. 1986).

Intron splicing usually occurs in *cis*, but a few cases of *trans*-splicing are known, and we provide evidence in this paper that this also applies to a *G. rosea* mitochondrial gene. For organelle genes, *trans*-splicing is usually mediated by group II introns (for a review, see Bonen 2008) but in a few recently discovered instances also by group I introns (Burger et al. 2009; Grewe et al. 2009; Pombert and Keeling 2010; Hecht et al. 2011). Curiously, all known *trans*-splicing group I introns are located in *cox1*, with identical insertion points in the quilwort *Isoetes* and the spikemoss *Selaginella* but also in the evolutionarily distant *Gigaspora* and the metazoan *Trichoplax* (supplementary fig. S1, Supplementary Material online). *Trans*-splicing entails the joining and ligation of discontinuous coding regions that are transcribed separately, that is, located on distinct RNA molecules. For group I and II introns, these separate transcripts may be brought together by interaction of partial (i.e., fragmented) intron RNA sequences that fold into the typical RNA structure that allows splicing to take place. In more complicated, rare instances, the flanking sequences of precursor transcripts alone do not code for the full intron structure, but additional helper RNAs are involved. In the two known instances, three distinct RNA molecules in *trans* are required to form a complete intron RNA structure (Goldschmidt-Clermont et al. 1991; Knoop et al. 1997). In any case, claims for intron-mediated *trans*-splicing should be based on evidence for 1) the presence of a mature RNA that is ligated exactly at predicted exon–intron junctions (e.g., based on multiple sequence alignments) and 2) an inferred intron RNA structure that is complete, matching previously established conservation rules perfectly. Note that not all discontinuous genes require *trans*-splicing. For instance, known fragmented rRNAs remain in separate transcripts, folding into a functional ribosomal structure, but without the requirement for *trans*-splicing (e.g., Boer and Gray 1988; Schnare and Gray 1990).

In this article, we report the complete mtDNA sequence of *G. rosea*, which encodes two fragmented genes, transcripts of one of which undergo group I intron–mediated *trans*-splicing. We further report the results of a phylogenetic analysis of mitochondrial proteins that allows more confident positioning of Glomeromycota within Fungi.

Materials and Methods

Fungal Material

Spores of *G. rosea* (DAOM194757) were extracted from soil of in vivo pot cultures with leek as a host plant, using a wet-sieve cascade (400, 250, and 60 μM). Spores were further purified by centrifugation through a water/50% sucrose step gradient (50-ml tube; 5 min at 5,000 rpm in an Eppendorf 5804 centrifuge; Esch et al. 1994). The resulting spore-containing layer from the upper interface of the 50% sucrose solution was collected, and clean spores were sorted manually under a binocular.

DNA Purification

Spores were suspended in 400 μl of the DNeasy Plant Mini Kit AP1 buffer (Qiagen) and crushed with a pestle in 1.5-ml

microtubes. DNA was purified with the same kit according to the manufacturer's recommendations, with a final elution volume of 40 µl. Purified DNA was stored at −20 °C until use.

RNA Purification

Freshly harvested *G. rosea* spores from in vivo pot cultures were vortexed with 1 ml TRIzol reagent (Invitrogen) and glass beads for 10 min. Other RNA extraction steps followed the manufacturer's protocol. Total RNA was then purified with the RNeasy mini kit (Qiagen), treated with DNase using the Turbo DNA free kit (Applied Biosystem), and stored at −20 °C. RNA quality was checked by electrophoresis on a 1.5% agarose gel, stained with ethidium bromide, and visualized under UV light.

Reverse Transcriptase–Polymerase Chain Reaction

We have used reverse transcriptase–polymerase chain reaction (RT-PCR) to test for intron excision, *trans*-splicing, and the presence of (apparently nonintron) sequence inserts in rRNAs. In all experiments, PCR amplifications were performed either with or without prior conversion of RNA into cDNA (i.e., addition of RT). As an additional control, genomic DNA was PCR amplified.

To test for potential *trans*-splicing of *cox1* and *rns*, primers were designed within the coding regions that flank predicted break points. In case of *cox1*, primers were located in *cox1* exon 4 (5′-CTGTATTGGTCACTGCCGT-3′) and exon 5 (5′-AAAGGCTGAAATAACATGGCT-3′). For *rns*, the two respective primers were 5′-ACCTTGATC-CAGCCAACTAGA-3′ and 5′-CACACTATCGTATCTCAG-CGTC-3′. To test for functionality of the RT-PCR assay (positive control), an internal region of *cox1* exon 4 was amplified (primers 5′-CTCTAGCAGGGACCCAGTC-3′ and 5′-CCGGATCATAGAAGCAGGT-3′).

To verify whether several inserts in *rnl* and *rns* remain at the RNA level (suspected because of an apparent lack of conserved intron RNA secondary structure), primers were designed within the flanking regions of these inserts. For *rns*, the primers were 5′-GTGCATTGTCATCA-CAGGTG-3′, 5′-CGAGTTACAGAGCACAGTTCG-3′, and 5′-GTCTCGTAACAAGCCTCCTTAAC-3′ (the latter is located inside the predicted insert). In the case of *rnl*, five primers were designed to test inserts 3, 4, and 5 (5′-AAG-TAGAGGCTCCAGAAGCAG-3′, 5′-GAAGGGTTTCATGA-GTAAGGTGA-3′, 5′-CAGGTCTGCAGGTTCACG-3′, 5′-CGTTCAGTCTTAACACTTGGC-3′, and 5′-CCCCTTTTT-AGTGCCGC-3′). We further verified that the two predicted *rnl* introns are indeed excised at the RNA level, using primer (5′-GGGCGCGTCTGTTTACTTA-3′) upstream of the first intron (group II), in conjunction with a primer downstream of the second group IB intron (5′-CGAGTACCGGTACCAGAGTAGGT-3′).

For samples undergoing cDNA synthesis, relevant primers (0.1 µl of 5 µM stock solutions), 2 µl RNA solution (~1 to 10 ng), and 8 µl of RNase-free water were mixed, denatured at 75 °C for 2 min, and immediately placed on ice for 2 min. Then, 4 µl RT buffer, 2 µl of each dNTP (10 mM stocks), 4 µl water, and 0.1 µl (1.5 U) AMV RT enzyme (Roche Applied Science) were added. After incubation at 45 °C for 45 min in the reaction mix, the RNA was denatured again at 75 °C for 2 min, placed on ice, fresh AMV RT enzyme was added, and the samples incubated for another 45 min at 45 °C. Reaction products were stored at −20 °C until use.

Subsequent PCR amplifications were carried out with the Expand High Fidelity PCR system (Roche Applied Science) in a total volume of 40 µl, containing either cDNAs (4 µl) or positive and negative controls (0.4 µl RNA or 0.2 µl genomic DNA), primers (4 µl each from 5 µM stocks), 4 µl PCR buffer (10×, without $MgCl_2$), 3.2 µl $MgCl_2$ (25 mM), 0.4 µl (1.5 U) DNA polymerase, 4 µl dNTPs (2 mM), and 1 µl RNase (5 µg/ml). Cycling parameters were 4 min at 95 °C, followed by 33 cycles of 15 s at 95 °C, 20 s at a 53 to 57 °C temperature gradient, 20 s at 72 °C, and a final elongation step of 4 min at 72 °C. PCR products were separated by electrophoresis in a 1.5% (w/v) agarose gel and visualized with ethidium bromide under UV light.

Sequencing, Assembly, and Gene Annotation

Gigaspora rosea total DNA was sequenced by 454 shotgun technology, and the resulting 587,881 reads were assembled with Newbler (Genome Quebec Innovation Center, McGill University, Montreal; Titanium Flex, ½ plate). The resulting 26 mitochondrial contigs were combined into a single circular-mapping DNA by PCR amplification of total DNA and Sanger sequencing of the PCR fragments. Regions with potential 454 sequence error (due to homopolymer motifs) were resequenced in the same way. Gene annotation was performed with MFannot (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl), followed by manual inspection and addition of missing gene features. Mfannot predicts group I and II introns, tRNAs, RNase P RNA, and 5S rRNA with Erpin (Gautheret and Lambert 2001) as a search engine, based on RNA structural profiles developed in house. Exons of protein-coding genes are inferred in a first round with Exonerate (Slater and Birney 2005) and then for less well-conserved genes with HMM profile searches (Eddy 2008; based on models for all known mtDNA-encoded proteins). Miniexons that are not recognized by Exonerate and that may be as short as 3 nucleotides are inferred by the presence of missing conserved protein regions (with reference to multiple protein alignments including other species) and orphan introns. The precise placement of small exons is based on the best fit of HMM protein profiles and on the fit with conserved nucleotide sequence profiles of respective (group I or II) exon–intron boundaries. Genes encoding the small and large rRNA subunits are predicted with HMM profiles that cover the most highly conserved domains, allowing precise predictions of the small subunit rRNA termini but only approximate positioning of large subunit rRNA ends. The latter termini, as well as the precise exon–intron structure of rRNA genes,

Table 1. Gene and Intron Content in Selected Fungal mtDNAs.

	Genes and Introns										
Species	***rnl, rns***	***atp6, 8, 9***	***cob***	***cox1, 2, 3***	***nad1–6*** [a]	***trn A-W***	***rnpB***	***rps3***	**ORFs** [b]	**Intron I** [c]	**Intron II** [c]
Gigaspora rosea	2	3	1	3	7	25	0	0	4	13	1
***Glomus irregulare* 494**	2	3	1	3	7	25	0	0	8	24	0
***G. irregulare* 197198**	2	3	1	3	7	25	0	0	8	26	0
Smittium culisetae	2	3	1	3	7	26	1	1	3	14	0
Mortierella verticillata	2	3	1	3	7	28	1	1	7	4	0
Rhizopus oryzae	2	3	1	3	7	23	1	0	4	9	0
Allomyces macrogynus	2	3	1	3	7	25	0	1	4	26	2
Saccharomyces cerevisiae [d]	2	3	1	3	0	25	1	1	3	9	4

[a] Includes *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6*.
[b] Only ORFs greater than 100 amino acids in length are listed, not including intronic ORFs and *dpo* and *rpo* fragments.
[c] Intron I and Intron II denote introns of group I and group II, respectively.
[d] *S. cerevisiae* strain FY 1679 (Foury et al. 1998).

are predicted manually (using comparative structure modeling principles). Furthermore, RT-PCR experiments are evaluated to distinguish between true introns and sequence insertions that are not removed by splicing. In any case, automated annotations are curated manually to account for MFannot warnings (e.g., potential *trans*-spliced genes, gene fusions, frameshifts, alternative translation initiation sites, failure to identify miniintrons, etc.), to correct potential errors, and to find features that are not recognized by the automated procedures. The annotated sequences (mtDNA and plasmid) are available at GenBank (JQ693395 and JQ693396).

Phylogenetic Analysis

The data set contains 13 mitochondrion-encoded proteins (*Cox1*, *2*, *3*, *Cob*, *Atp6*, *9* and *Nad1*, *2*, *3*, *4*, *4L*, *5*, *6*) and includes sequences from all zygomycetes for which complete mtDNA sequences are available. Protein collections were managed and automatically aligned, trimmed, and concatenated with Mams (developed in house; Lang BF, Rioux P, unpublished data). Mams uses Muscle (Edgar 2004) for an initial alignment, followed by a refinement step with HMMalign (S. Eddy; http://hmmer.janelia.org). The final data set contains 37 taxa and 3,664 amino acid positions and is available from the authors upon request.

Phylogenetic analyses were performed by Bayesian inference (PhyloBayes; Lartillot and Philippe 2004) using the CAT+Gamma model, six discrete categories, four independent chains, 10,000 cycles (corresponding to ~550, 000 generations), and the −dc parameter to remove constant sites. The site removal reduces the total number of amino acid positions to 3048. The first 1,700 cycles were removed as burn-in. The robustness of internal branches was evaluated based on 100 jackknife (60%) replicates, as modeling of duplicated sequence sites generated by bootstrap analysis is problematic with the Bayesian approach.

Maximum likelihood (ML) analysis was performed with RaxML-HPC v7.2.2 (Stamatakis 2006), using the LG model (PROTGAMMALGF), and the fast bootstrapping option (100 replicates).

Results and Discussion

Comparison of *G. irregulare* and *G. rosea* Mitochondrial Genomes

The mitochondrial genome of *G. rosea* has been completely sequenced by a combination of 454 sequencing from total DNA and filling of the remaining 26 gaps (due to low sequence coverage and 454 sequence error) by PCR amplification and Sanger sequencing. The size of *G. rosea* mtDNA is relatively large (97,349 bp) when compared with those of the two *G. irregulare* isolates (70,606 and 70,799 bp) and contained several homopolymer stretches that required correction of the 454 data by traditional Sanger sequencing. In addition, paired-end sequencing was performed to avoid misassembly of the genome due to the presence of long sequence repeats. Whenever sequence repeats occur in a genome, more than one genome conformation may exist due to recombination events, but we have no evidence that this occurs in *G. rosea*. The large genome size of *G. rosea* is essentially due to extended intergenic regions, not a variation in the number or size of introns (table 1), or the presence of mitochondrial plasmid insertions that are otherwise common in fungi.

The high number of initial sequencing gaps was surprising as according to our experience, the same amount of total 454 sequence information resulted in complete or almost complete coverage of *Glomus* spp. mtDNA. Among possible explanations, a substantial contamination with DNA from foreign species seems unlikely, as spores were carefully cleaned and sorted individually prior to extraction of total DNA. It remains possible either that the copy number of mtDNA is lower in *G. rosea* than in other AMF species, that its nuclear genome is much more complex, or that the spores harbor endosymbionts accounting for a substantial increase in total DNA relative to mtDNA content. Given the available sequence data, we have no reason to favor any of these scenarios.

As in most other fungi, the *G. rosea* mtDNA maps as a circular molecule (fig. 1) but is likely organized as a linear multimeric concatamer in vivo (Bendich 1996). To allow for easier genome comparison, we have opened the circle at a previously introduced standard position, upstream of *rnl*. Conservation of mitochondrial gene order between

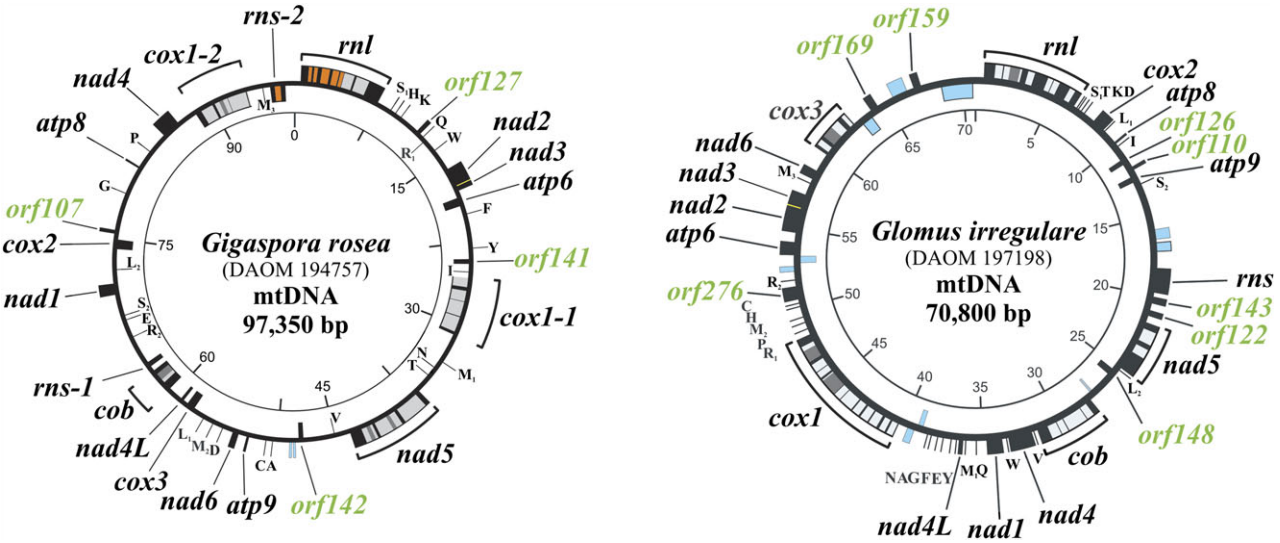

Fig. 1. Comparison of *Gigaspora rosea* and *Glomus irregulare* mitochondrial genomes. The circular-mapping genomes of *G. rosea* and *G. irregulare* were opened upstream of *rnl* to allow for easier comparisons. Genes on the outer and inner circumference are transcribed in clockwise and counterclockwise direction, respectively. Arcs indicate coding regions interrupted by introns. Gene fragments in *G. rosea* are numbered −1 and −2. Boxes of coding regions are filled black, intron ORFs dark gray, and introns light gray. Regions with similarity to mitochondrial plasmid-like DNA polymerase are marked light blue and insertions in *rns* and *rnl* that are not excised from the rRNA (i.e., are not introns) in orange. Gene and corresponding product names are *atp6*, ATP synthase subunit 6; *cob*, apocytochrome b; *cox1–3*, cytochrome c oxidase subunits; *nad1–4*, *4L*, *5–6*, NADH dehydrogenase subunits; *rnl*, *rns*, large and small subunit rRNAs; A-W, tRNAs, the letter corresponding to the amino acid specified by the particular tRNA.

Gigaspora and *Glomus* is virtually nonexistent and, in contrast *to G. irregulare*, genes are encoded on both strands of the *G. rosea* mtDNA. These genes belong to the basic fungal set (table 1), including all required tRNAs. Three tRNAs have a CAU anticodon, one each for initiator and methionine elongator tRNAs and a third one in which a potential lysidine modification of the anticodon C residue would allow decoding of AUA as isoleucine (e.g., Muramatsu et al. 1988; Weber et al. 1990). Notably, the predicted initiator methionine tRNA exhibits structural features in the anticodon loop that never occur in orthodox tRNAs, entailing a most unorthodox modification of decoding preferences (for more details, see below).

As in *G. irregulare* (Lee and Young 2009), genes for the ribosomal protein *rps3* and the RNA component of mitochondrial RNase P (*rnpB*) are absent, although they occur in most zygomycetes (Seif et al. 2005). However, whereas *rps3* may have been transferred to the nuclear genome, this inference remains more contentious for *rnpB* as it would require targeting of a relative large RNA molecule from the cytoplasm back into mitochondria. An alternative solution would be a nucleus-encoded protein-only RNase P activity that replaces the ribozyme, as it occurs in animal and plant mitochondria (Holzmann et al. 2008; Gobert et al. 2010).

Mitochondrial Plasmid Insertions

Mitochondrial autonomously replicating plasmids are widespread. They may carry their own genes for DNA replication (*dpo*) and transcription (*rpo*), and in some cases, plasmids are known to integrate into mtDNA. There is no biochemical or genetic evidence that complete or fragmented versions of *dpo* and *rpo* are functional when inserted into mtDNA; in fact, plasmid insertion into mtDNAs is followed by rapid genome reorganization and loss (e.g., Bertrand et al. 1986; Vierula and Bertrand 1992; Baidyaroy et al. 2000; Hausner 2011), confirming the view that *dpo* and *rpo* are required for plasmid replication and transcription only.

We have used our shotgun data containing both mitochondrial and nuclear sequence reads to search for plasmid-related sequences. *G. irregulare* mtDNA carries many more and larger plasmid-related DNA polymerase (*dpo*) gene fragments than *G. rosea* (filled light blue boxes; two tiny fragments in *G. rosea* are at map position ~50 kbp; fig. 1). In addition, *G. rosea* has an apparently complete, plasmid-like *dpo* (but no *rpo*) in a 3,582-bp contig. Given its relatively high sequence coverage with respect to nuclear genes, this element is likely a freestanding plasmid that only encodes *dpo*. We have not succeeded in demonstrating a circular plasmid version by PCR amplification (including the use of a dedicated long PCR kit; data not shown) and therefore assume that the plasmid is linear. Its cellular location (cytoplasmic or mitochondrial) remains unknown. The *dpo* gene is translated with the standard genetic code, but this would be compatible with both a cytoplasmic and mitochondrial location (for a more detailed discussion of the mitochondrial genetic code, see below). Whatever its cellular location, the *G. rosea* plasmid *dpo* is most closely related to mitochondrial *dpo* fragments, in both *Glomus* and *Gigaspora*, belonging to the family of plasmids that repeatedly invade mitochondrial genomes. A freestanding form of a mitochondrial plasmid has so far not been described for glomeromycotan species.

Evolution of the Mitochondrial Genetic Code in Glomeromycota

Some members of the paraphyletic zygomycete lineages, such as *Rhizopus oryzae*, but also the rapidly evolving *Smittium culisetae*, have retained the standard genetic code for the regular set of mitochondrial protein-coding genes, a trait inherited from their bacterial ancestors (Seif et al. 2005). However, *Mortierella verticillata* reassigns UGA "stop" codons as tryptophan, one each in *nad3* and *nad4*. Likewise, UGA(Trp) codons are also present in the *S. culisetae* group I intronic open reading frame (ORF)283 and ORF248, encoding homing endonucleases of the LAGLI/DADG type (Seif et al. 2005).

In Glomeromycota, deviations from the standard genetic code also vary. A few clear-cut instances of UGA(Trp) codons are found in the two published *G. irregulare* mtDNAs but not in *G. rosea*. In turn, translation initiation with either AUG(Met) or GUG(Met) leads to seven proteins in *G. rosea* (i.e., a surprising 50% of standard protein-coding genes) that are severely truncated at their amino-terminal. In these instances, sequence similarity at the protein level extends further upstream to potential UUG initiation codons (supplementary fig. S2, Supplementary Material online).

UUG initiation is infrequent in fungal mitochondria (e.g., an isolated instance in *Mortierella*; GenBank #AY863211; supplementary fig. S2, Supplementary Material online). It does occur in bacteria where initiation of translation depends on a specific initiator methionine tRNA that recognizes AUG, GUG, and UUG, in descending order of efficiency. In the case of UUG initiation, only the second and third codon positions (U and G) are able to interact directly with the tRNA anticodon (i.e., the C and A of the CAU anticodon), accounting for the very low effectiveness of UUG as an initiation codon, despite stabilization and precise positioning of translation initiation codons by Shine–Dalgarno sequence motifs. These motifs occur at a defined distance upstream of initiation codons, yet in mitochondria, they are only known from the jakobid flagellate *Reclinomonas americana* (Lang et al. 1997). How then are UUG codons recognized in mitochondria of *G. rosea*?

A mechanism for more precise positioning of translation starts is known from mitochondria of monoblepharidalian chytrid fungi and the sea anemone *Metridium senile*, in which almost every protein-coding gene has a guanosine residue upstream of the predicted AUG or GUG start codons (Bullerwell, Forget, et al. 2003). This conserved G residue converts the triplet into a more stable quadruplet initiation codon, correlating with the presence of an unorthodox cytosine residue at position 37 in the anticodon loop of the initiator tRNAs. This feature permits a more stable 4-bp interaction between CAUC anticodons and quartet GAUG/GGUG codons. Inspection of *G. rosea* sequences does not confirm the presence of quadruplet initiation codons, yet its predicted initiator tRNA is highly unorthodox, with a G residue at position 32 that is otherwise always a pyrimidine (for a recent review on mitochondrial tRNA structure, see Lang et al. 2011). In fact, in a 3D structure, position G-32 of the anticodon loop is in close vicinity to both the U-36 of the anticodon and the first U of UUG initiation codons. Interaction of G-32 with both U residues would thus stabilize the U–U pair. In other words, all three nucleotides of the UUG codon are capable of forming stable interactions with this unorthodox initiator tRNA, implying that translation initiation with UUG becomes as effective as with AUG: in accord with the high incidence of inferred UUG translation initiation codons in *G. rosea* mitochondria.

Robust Phylogenetic Association of Glomeromycota with Mortierellales

The Glomeromycota are part of zygomycetes, a phylogenetically heterogenous (paraphyletic) taxon (Seif et al. 2005; James et al. 2006; Liu, Steemkamp, et al. 2009). According to a current proposal, Glomeromycota constitute an independent fungal phylum (Hibbett et al. 2007), yet phylogenomic analyses with nuclear and mitochondrial genes (e.g., Seif et al. 2006; Liu, Leigh, et al. 2009) do not support this view. When the complete mtDNA data of *G. irregulare* are included in the analysis, this fungus associates with *Mortierella* (Lang and Hijri 2009; Lee and Young 2009): an affinity that was also observed with a nuclear data set (Liu, Leigh, et al. 2009), although in both cases based on poor taxon sampling and without significant statistical support. Evidently, additional zygomycete nuclear or mitochondrial genome data are required to address this question and demonstrate beyond reasonable doubt that Glomeromycota merit the status of an independent phylum.

When *G. rosea* data are included in a data set containing all standard mtDNA-encoded proteins, Glomeromycota branch with *Mortierella* as noted above, but with better statistical support (PP 0.99; 66% jackknife value [JV]; fig. 2). In addition, *S. culisetae*, a rapidly evolving species belonging to Harpellales (grouping at a basal position with ML and the applied global model of protein evolution; data not shown), occupies a sistergroup position with Glomeromycota when PhyloBayes and the more realistic CAT model are used for phylogenetic inference (modeling site-wise heterogeneity of amino acid positions, one of the major sources of systematic phylogenetic error in global models) (Lartillot and Philippe 2004; Lartillot et al. 2007). However, the inclusion of *Smittium* in a monophyletic group including *Mortierella* and the two Glomeromycota draws only low statistical support (a posterior probability value of 0.68% and 43% JV), which we attribute to a combination of low phylogenetic signal (use of a single representative for Harpellales) and conflicting long-branch attraction (because rapidly evolving) toward a basal position in the tree. When removing Smittium from the data set, the support for Mortierella forming a monophyletic group with Glomeromycota increases substantially (PP 1.0; 93% JV)

In summary, our results further support the view that Glomeromycota and Mortierellales are sister taxa. This

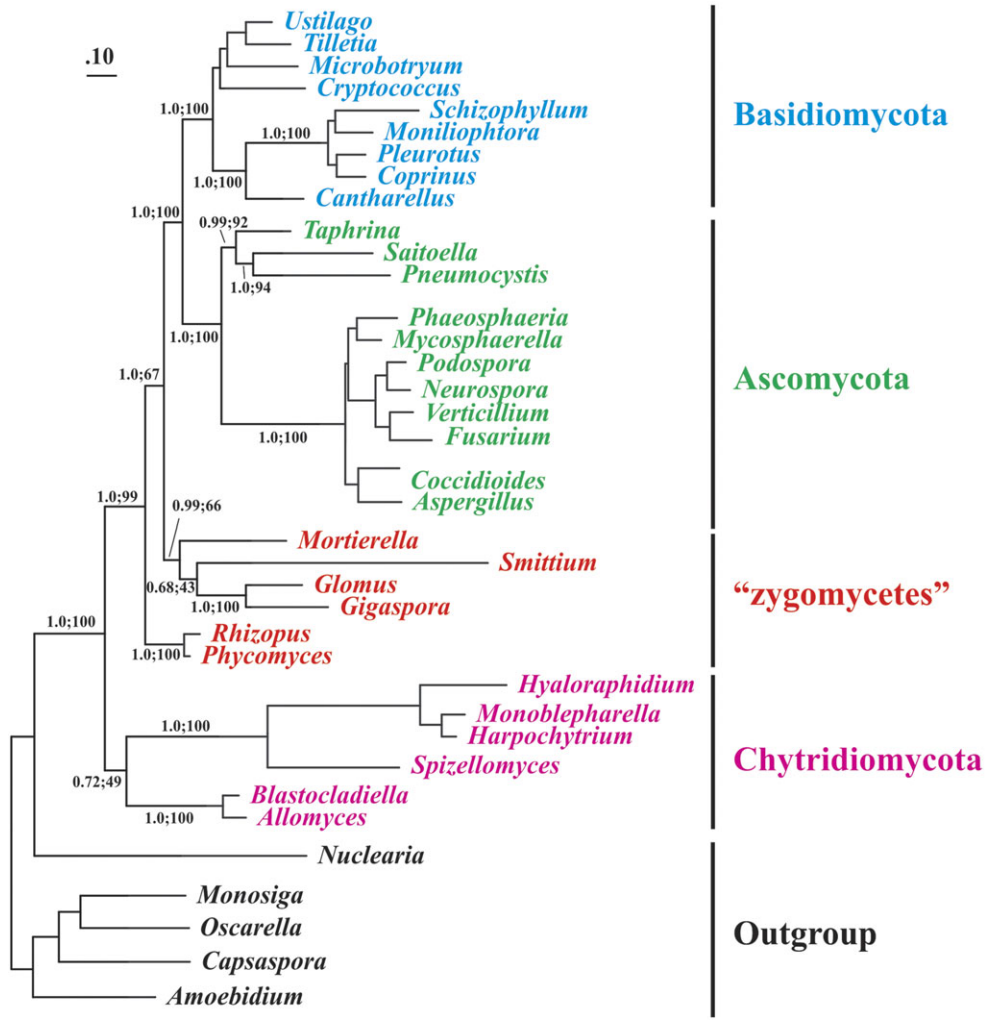

Fig. 2. Phylogenetic positioning of Glomeromycota with mitochondrial protein data. The tree shown was inferred with PhyloBayes, the CAT+Gamma model and six discrete categories (Lartillot and Philippe 2004) based on 13 concatenated proteins. The first number at branches indicates posterior probability values of a PhyloBayes analysis with four independent chains and the second number jackknife supports values. ML inference predicts a similar tree (not shown), except for *Smittium* and Blastocladiales.

inference and the potential inclusion of Harpellales in the same monophyletic group remain to be tested with extended species sampling across "zygomycetes." In particular, representatives of Harpellales with more moderate evolutionary rates need to be identified and data from Mortierellales and Glomeromycota added. Contrary to our previously more pessimistic assessment (Lang and Hijri 2009), broad taxon sampling of mtDNAs alone may turn out to be sufficient to resolve phylogenetic relationships of Glomeromycota with confidence. Whatever the outcome, checking the consistency of mitochondrial phylogenies with phylogenomic analyses based on nuclear gene sequences will be important as a means of informing us about hidden phylogenetic artifacts.

Group I Intron–Mediated *Trans*-splicing of *G. rosea cox1*

The *G. rosea cox1* gene (encoding subunit 1 of respiratory chain complex IV) is broken up into two fragments that are located at a distance of ~30 kbp, encoded on the same strand of the mitochondrial genome. Formally, a *cox1* gene with a 30 kbp intron could be postulated, but this scenario is most unlikely for the following reasons. The two *cox1* fragments are separated by a total of 15 genes, including several tRNAs that are on the same strand. It is known that tRNAs are rapidly and efficiently processed from RNA precursors, so that a 30-kb RNA precursor would be rapidly fragmented within the large intervening region. In addition, the *cox1* break points correspond exactly to a position where *G. irregulare* carries a group I intron, favoring the hypothesis of group I intron–mediated *trans*-splicing.

Examples of group I intron–mediated *trans*-splicing have been reported previously (Burger et al. 2009; Grewe et al. 2009; Pombert and Keeling 2010; Hecht et al. 2011), which in all three cases occur in the most intron-rich mitochondrial gene, *cox1*. The insertion points of *trans*-spliced group I introns vary (supplementary

Table 2. Presence (●) or Absence (○) at Cognate Insertion Site of the Shared Introns between *Gigaspora* and *Glomus* in Representatives of Other Fungi.

Species	Accession	Introns								
Glomeromycota										
Glomus irregulare	FJ648425	*rnli1939*	*nad5i417*	*nad5i717*	○	*cobi429*	*cox1i386*	*cox1i731*	*cox1i867*	*cox1i1108*
Gigaspora rosea		●	●	●	*nad5i954*	●	●	±	●	●
Basidiomycota										
Cryptococcus neoformans	NC004336	○	○	○	○	○	○	○	○	○
Pleurotus ostreatus	EF204913	○	○	○	○	○	●	○	○	●
Ascomycota										
Pneumocystis carinii	GU133622	○	○	○	○	○	○	○	○	○
Podospora anserina	NC001329	○	○	○	○	○	○	○	●	●
Zygomycota										
Mortierella verticillata	AY863211	○	○	○	○	○	○	○	○	○
Rhizopus oryzae	AY863212	○	○	○	○	○	●	○	○	○
Chytridiomycota										
Monoblepharella	AY182007	○	●	●	○	○	○	○	○	●
A. macrogynus	NC001715	○	○	●	○	●	●	○	●	●

NOTE.—The representation ± indicates a partial group I intron secondary structure. The gray columns represent the *cox1* intron that is *trans*-spliced in *G. rosea*, and the *nad5* intron believed to be involved in the *trans*-splicing. Introns were named according to Dombrovska and Qiu (2004). Shared intron positions are based on the coding sequence of the gene in which they are inserted in the first Glomeromycota mitochondrial genome published (Lee and Young 2009). Unique introns to *G. rosea* are named according to their positions in the gene in which they were identified. Of the 13 group I introns in *G. rosea* and the 24 in *G. irregulare* 494, only seven are shared at cognate insertion site. Note that the *G. rosea* nad5i954 intron that may complement a functional secondary structure of the *trans*-spliced *cox1* intron is not present in *G. irregulare* and in the other represented fungal species.

fig. S1, Supplementary Material online; table 2), with exceptions in the related species *Isoetes* and *Selaginella* and in *G. rosea* and *Trichoplax* where introns occupy the same (otherwise frequent) insertion point. In the published examples, a complete canonical group I intron RNA secondary structure can be inferred by pairing the flanking portion of the discontinuous exons and is confirmed by use of group I intron prediction algorithms (Lang et al. 2007; Beck and Lang 2009). Yet in *G. rosea*, certain structural core elements appear to be either missing or to deviate significantly from established rules. As other group I introns in *G. rosea*, protein-coding genes do not display many unusual features, and as a catalytically conserved intron RNA core is essential for intron excision, we predict the use of an additional third helper RNA fragment in *trans*, similar to the situation of fragmented introns in *Chlamydomonas* (Goldschmidt-Clermont et al. 1991) and *Oenothera* (Knoop et al. 1997). According to our assessment, the third intron in the *nad5* gene matches perfectly the two partial discontinuous intron sequences, completing a comprehensive intron RNA secondary structure (fig. 3*A* and *B*).

To provide evidence that the mitochondrial *cox1* is not a pseudogene in *G. rosea* but indeed *trans*-spliced in vivo, we have conducted RT-PCR experiments, using primers located in the respective flanking exons. Sequencing of the resulting PCR product confirms that the *cox1* exons are accurately ligated in vivo (fig. 3*C*). Sequencing of the PCR product confirmed sequence identity with the respective *cox1* exons, without any sequence modification. In the placozoan animal example as mentioned above, *trans*-splicing combines with an additional RNA editing step (Burger et al. 2009). No PCR product was obtained with total genomic DNA, ruling out the possibility that a contiguous *cox1* gene exists in the nuclear genome.

Gene for the Small Subunit rRNA in Two Pieces

The other fragmented *Gigaspora* gene is *rns*, coding for the small subunit rRNA. In similar cases (including both the large and the small subunit rRNAs), *trans*-splicing has never been observed, and break points usually occur in variable loops of the rRNA secondary structure rather than in highly conserved regions as in protein-coding genes (Lang 1984). Secondary and tertiary interactions between the unligated rRNA fragments are sufficient for folding into a complete rRNA structure and formation of functional ribosomes. According to our rRNA secondary structure model, the breakpoint of *G. rosea* is in a variable loop (fig. 4). RT-PCR experiments did not produce DNA fragments of the predicted size and sequence (data not shown), and we explain a few incongruent RT-PCR sequences that we obtained by a low level of template-switching activity of RT. We have also been unsuccessful in folding the flanking regions into a bona fide intron RNA structures (either by inference with intron models that are part of the MFannot annotation or manually).

How Many True Introns in *rns* and *rnl*?

Modeling of rRNA structures and prediction of introns are problematic, without confirmation at the RNA sequence level. It is well known that mitochondrial rRNAs may carry inserts in variable regions of the structure that are not excised at the RNA level (e.g., Lang et al. 1987). In *G. rosea rns* and *rnl*, we find numerous inserts of this kind (figs. 2 and 4) that do not carry evidence for sound group I or group II intron structures. RT-PCR

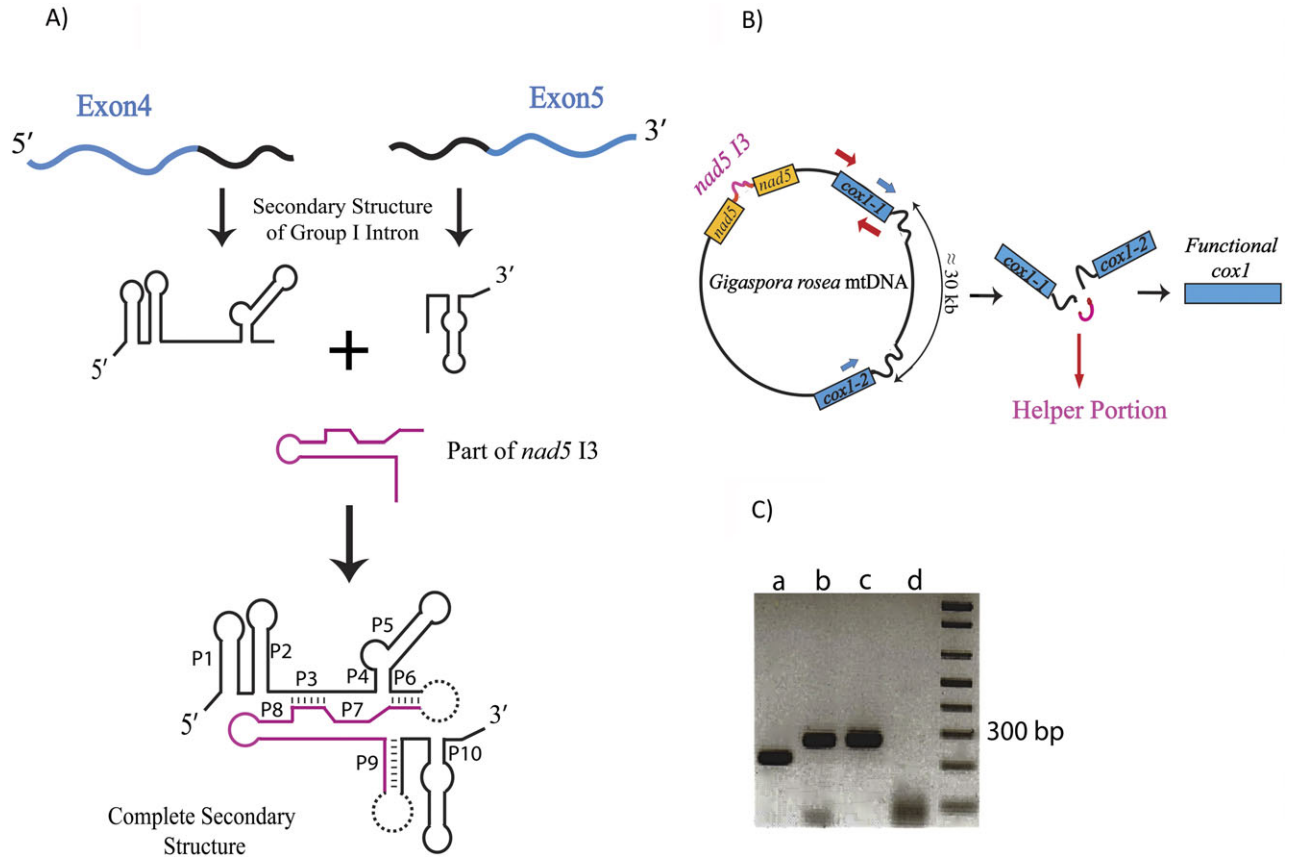

Fig. 3. Model of group I intron–mediated *trans*-splicing in *G. rosea* and demonstration of mRNA *trans*-splicing by RT-PCR. (*A*) Schematic view of the *trans*-spliced *cox1* group IB intron RNA secondary structure. The pink-colored RNA corresponds to the predicted *nad5* intron 3 helper sequence, and dashed loops indicate its interaction with the two intron fragments flanking *cox1* exons 4 and 5 (shown in blue). P1–P10 illustrate the conserved stems of the intron secondary structure according to Burke et al. (1987). Canonical Watson–Crick base pairing is shown by dashes. (*B*) Genomic view of RNAs involved in *cox*1 *trans*-splicing. Blue and red arrows indicate PCR and RT-PCR primers (see Materials and Methods). (*C*) Amplification products obtained with *G. rosea* cDNA and genomic DNA. a, RT-PCR amplification in the presence of total RNA (demonstration of *trans*-splicing and that mitochondrial *cox1* is not a pseudogene); b, RT-PCR amplification of an internal section of exon 4 (positive control); c, PCR amplification of total genomic DNA, with the same primers as in "b" (positive control); and d, PCR amplification of total genomic DNA, with the primers used in "a" (negative control, no presence of a contiguous gene in the nuclear genome). Further negative controls (see Materials and Methods) not shown.

experiments confirm this view: the insert in *rns_2* is not excised at the RNA level nor are five intron-size inserts in *rnl*. According to our results, RNA splicing occurs only for the two *rnl* introns that are predicted by RNAweasel (Beck and Lang 2009; one group II followed by a group I intron; fig. 2). It is noteworthy that one of the *rnl* inserts occurs precisely at a position where *G. irregulare* carries a true intron, as verified in that case by cDNA sequencing (Lee and Young 2009). Evidently, the nature and size of gene and genome insertions in glomeromycotan mtDNAs undergo the most rapid change, similar to what we observe with mitochondrial plasmid insertions.

Conclusion

The *Gigaspora* mitochondrial genome has turned out to be most unusual, with fragmented genes, *trans*-splicing, a ribosomal RNA gene in pieces, and an inflated genome size. The presence of substantial gene order differences between *Glomus* and *Gigaspora* is another intriguing observation, suggesting the capacity for nonhomologous recombination repair. The repeated insertion of plastid-like sequences into the mtDNA is in support of this view, as is a high but variable number of (mostly group I) introns. Group I introns propagate by a homing mechanism that relies on homologous recombination (Perrin et al. 1993; Chevalier et al. 2003), that is, we hypothesize that the biochemical machineries for both homologous and nonhomologous recombination are present in glomeromycotan mitochondria. The effectiveness and mechanisms of recombination may be tested by combining, for instance, compatible *Glomus* species with differences in sequence or intron content via anastomosis and analysis for recombination products in the segregated offspring. Evidently, mtDNAs do segregate effectively in Glomeromycota, as any mtDNA so far sequenced has had a perfectly homogeneous sequence: in contrast to their nuclear genomes, which are multinucleate, most divergent in

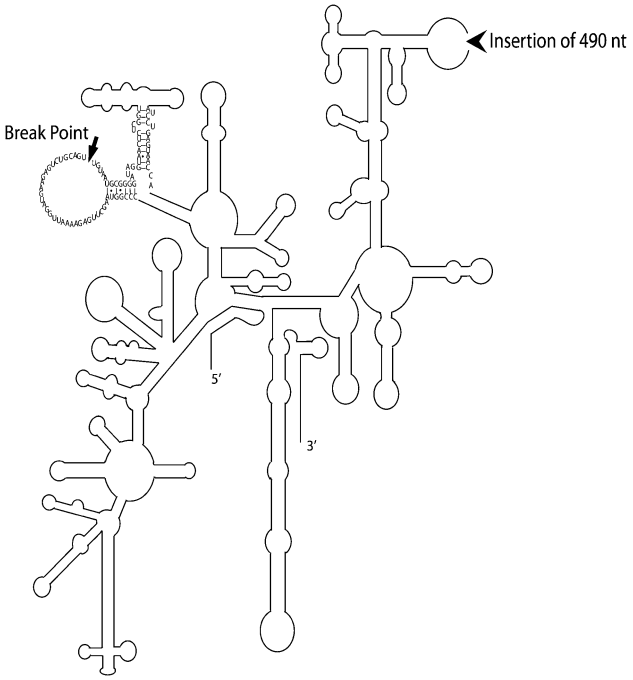

Fig. 4. Secondary structure model of the fragmented *G. rosea* small subunit rRNA. Schematic *rns* structure in which only sequences close to the breakpoint are indicated. The 5′ and 3′ termini at the breakpoint are precisely inferred based on (significant) sequence similarity with the *Glomus irregulare* counterpart. Dashes indicate canonical Watson–Crick base pairings, and dots denote guanine–uracil pairings. The arrowhead points to a 490-nt (nonintron) insertion in a variable loop.

sequence, and seemingly impossible to assemble (Martin et al. 2008). The large differences among glomeromycotan mitochondrial genomes should motivate the future inclusion of other distant lineages, such as *Acaulospora* and *Scutellospora*, for further comparisons. These new sequences will at the same time serve in confirming the phylogenetic positioning proposed here of Glomeromycota within Fungi.

Supplementary Material

Supplementary figures S1 and S2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Dr Michael Gray (Halifax, NS) for critical reading of the manuscript, Natacha Beck for continued development of automated organelle genome annotation software (MFannot), and Dr Yves Piché for kindly providing *G. rosea*. We thank NSERC-CRD and Genome Quebec/Genome Canada for providing financial support to B.F.L. and M.H.

References

Adams KL, Clements MJ, Vaughn JC. 1998. The *Peperomia* mitochondrial coxI group I intron: timing of horizontal transfer and subsequent evolution of the intron. *J Mol Evol.* 46:689–696.

Anziano PQ, Hanson DK, Mahler HR, Perlman PS. 1982. Functional domains in introns: trans-acting and cis-acting regions of intron 4 of the cob gene. *Cell* 30:925–932.

Baidyaroy D, Glynn JM, Bertrand H. 2000. Dynamics of asexual transmission of a mitochondrial plasmid in *Cryphonectria parasitica*. *Curr Genet.* 37:257–267.

Beck N, Lang BF. 2009. RNAweasel, a webserver for identification of mitochondrial, structured RNAs [Internet]. Montreal (Quebec): University of Montreal; 2006. Available from: http://megasun.bch.umontreal.ca/RNAweasel

Bendich AJ. 1996. Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. *J Mol Biol.* 255:564–588.

Bertrand H, Griffiths AJ, Court DA, Cheng CK. 1986. An extrachromosomal plasmid is the etiological precursor of kalDNA insertion sequences in the mitochondrial chromosome of senescent *Neurospora*. *Cell* 47:829–837.

Boer PH, Gray MW. 1988. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. *Cell* 55: 399–411.

Bonen L. 2008. Cis- and trans-splicing of group II introns in plant mitochondria. *Mitochondrion* 8:26–34.

Boon E, Zimmerman E, Lang BF, Hijri M. 2010. Intra-isolate genome variation in arbuscular mycorrhizal fungi persists in the transcriptome. *J Evol Biol.* 23:1519–1527.

Bullerwell CE, Forget L, Lang BF. 2003. Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Res.* 31:1614–1623.

Burger G, Yan Y, Javadi P, Lang BF. 2009. Group I-intron trans-splicing and mRNA editing in the mitochondria of placozoan animals. *Trends Genet.* 25:381–386.

Burke JM, Belfort M, Cech TR, Davies RW, Schweyen RJ, Shub DA, Szostak JW, Tabak HF. 1987. Structural conventions for group I introns. *Nucleic Acids Res.* 15:7217–7221.

Cech TR, Tanner NK, Tinoco I Jr, Weir BR, Zuker M, Perlman PS. 1983. Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence: structural homology with fungal mitochondrial intervening sequences. *Proc Natl Acad Sci U S A.* 80:3903–3907.

Chevalier B, Turmel M, Lemieux C, Monnat RJ Jr, Stoddard BL. 2003. Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. *J Mol Biol.* 329:253–269.

Colleaux L, d'Auriol L, Betermier M, Cottarel G, Jacquier A, Galibert F, Dujon B. 1986. Universal code equivalent of a yeast mitochondrial intron reading frame is expressed into E. coli as a specific double strand endonuclease. *Cell* 44:521–533.

Croll D, Giovannetti M, Koch AM, Sbrana C, Ehinger M, Lammers PJ, Sanders IR. 2009. Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus *Glomus intraradices*. *New Phytol.* 181:924–937.

Croll D, Sanders IR. 2009. Recombination in *Glomus intraradices*, a supposed ancient asexual arbuscular mycorrhizal fungus. *BMC Evol Biol.* 9:13.

Dombrovska O, Qiu YL. 2004. Distribution of introns in the mitochondrial gene nad1 in land plants: phylogenetic and molecular evolutionary implications. *Mol Phylogenet Evol.* 32:246–263.

Ebersberger I, de Matos Simoes R, Kupczok A, Gube M, Kothe E, Voigt K, von Haeseler A. A consistent phylogenetic backbone for the fungi. *Mol Biol Evol.* Advance Access published March 12, 2012, doi:10.1093/molbev/msr285

Eddy S. 2008. HMMER website [Internet]. Ashburn (VA): HHMI Janelia Farm Research Campus. Available from: http://hmmer.janelia.org

Edgar RC. 2004. Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.

Esch H, Hundeshagen B, Schneider-Poetsch H, Bothe H. 1994. Demonstration of abscisic acid in spores and hyphae of the arbuscular-mycorrhizal fungus *Glomus* and in the N2-fixing cyanobacterium *Anabaena variabilis*. *Plant Sci.* 99:9–16.

Felsenstein J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. *Syst Zool.* 27:27–33.

Foury F, Roganti T, Lecrenier N, Purnelle B. 1998. The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. *FEBS Lett.* 440:325–331.

Gautheret D, Lambert A. 2001. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *J Mol Biol.* 313:1003–1011.

Gobert A, Gutmann B, Taschner A, Gossringer M, Holzmann J, Hartmann RK, Rossmanith W, Giegé P. 2010. A single *Arabidopsis* organellar protein has RNase P activity. *Nat Struct Mol Biol.* 17:740–744.

Goldschmidt-Clermont M, Choquet Y, Girard-Bascou J, Michel F, Schirmer-Rahire M, Rochaix JD. 1991. A small chloroplast RNA may be required for trans-splicing in *Chlamydomonas reinhardtii*. *Cell* 65:135–143.

Grewe F, Viehoever P, Weisshaar B, Knoop V. 2009. A trans-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte *Isoetes engelmannii*. *Nucleic Acids Res.* 37:5093–5104.

Hausner G. 2011. Introns, mobile elements, and plasmids. In: Bullerwell CE, editor. Organelle genetics. Berlin (Germany): Springer.

Hecht J, Grewe F, Knoop V. 2011. Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of *Selaginella moellendorffii* mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes. *Genome Biol Evol.* 3:344–358.

Hibbett DS, Binder M, Bischoff JF, et al. 2007. A higher-level phylogenetic classification of the fungi. *Mycol Res.* 111:509–547.

Hijri M, Sanders IR. 2005. Low gene copy number shows that arbuscular mycorrhizal fungi inherit genetically different nuclei. *Nature* 433:160–163.

Holzmann J, Frank P, Loffler E, Bennett KL, Gerner C, Rossmanith W. 2008. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* 135:462–474.

Jacquier A, Rosbash M. 1986. Efficient trans-splicing of a yeast mitochondrial RNA group II intron implicates a strong 5′ exon-intron interaction. *Science* 234:1099–1104.

James TY, Kauff F, Schoch CL, et al. 2006. Reconstructing the early evolution of fungi using a six-gene phylogeny. *Nature* 443:818–822.

Knoop V, Altwasser M, Brennicke A. 1997. A tripartite group II intron in mitochondria of an angiosperm plant. *Mol Gen Genet.* 255:269–276.

Lambowitz AM, Zimmerly S. 2004. Mobile group II introns. *Annu Rev Genet.* 38:1–35.

Lang BF. 1984. The mitochondrial genome of the fission yeast *Schizosaccharomyces pombe*: highly homologous introns are inserted at the same position of the otherwise less conserved *cox1* genes in *Schizosaccharomyces pombe* and *Aspergillus nidulans*. *EMBO J.* 3:2129–2136.

Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387:493–497.

Lang BF, Cedergren R, Gray MW. 1987. The mitochondrial genome of the fission yeast, *Schizosaccharomyces pombe*. Sequence of the large-subunit ribosomal RNA gene, comparison of potential secondary structure in fungal mitochondrial large-subunit rRNAs and evolutionary considerations. *Eur J Biochem.* 169:527–537.

Lang BF, Hijri M. 2009. The complete *Glomus intraradices* mitochondrial genome sequence—a milestone in mycorrhizal research. *New Phytol.* 183:3–6.

Lang BF, Laforest MJ, Burger G. 2007. Mitochondrial introns: a critical view. *Trends Genet.* 23:119–125.

Lang BF, Lavrov D, Beck N, Steinberg V. 2011. Mitochondrial tRNA structure, identity and evolution of the genetic code. In: Bullerwell CE, editor. Organelle genetics. Berlin (Germany): Springer.

Lartillot N, Brinkmann H, Philippe H. 2007. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evol Biol.* 7(Suppl 1):S4.

Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol Biol Evol.* 21:1095–1109.

Lee J, Young JP. 2009. The mitochondrial genome sequence of the arbuscular mycorrhizal fungus *Glomus intraradices* isolate 494 and implications for the phylogenetic placement of *Glomus*. *New Phytol.* 183:200–211.

Liu Y, Leigh JW, Brinkmann H, Cushion MT, Rodriguez-Ezpeleta N, Philippe H, Lang BF. 2009. Phylogenomic analyses support the monophyly of Taphrinomycotina, including *Schizosaccharomyces* fission yeasts. *Mol Biol Evol.* 26:27–34.

Liu Y, Steenkamp ET, Brinkmann H, Forget L, Philippe H, Lang BF. 2009. Phylogenomic analyses predict sistergroup relationship of nucleariids and fungi and paraphyly of zygomycetes with significant support. *BMC Evol Biol.* 9:272.

Marleau J, Dalpe Y, St-Arnaud M, Hijri M. 2011. Spore development and nuclear inheritance in arbuscular mycorrhizal fungi. *BMC Evol Biol.* 11:51.

Martin F, Gianinazzi-Pearson V, Hijri M, Lammers P, Requena N, Sanders IR, Shachar-Hill Y, Shapiro H, Tuskan GA, Young JP. 2008. The long hard road to a completed *Glomus intraradices* genome. *New Phytol.* 180:747–750.

Michel F, Jacquier A, Dujon B. 1982. Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie* 64:867–881.

Muramatsu T, Yokoyama S, Horie N, Matsuda A, Ueda T, Yamaizumi Z, Kuchino Y, Nishimura S, Miyazawa T. 1988. A novel lysine-substituted nucleoside in the first position of the anticodon of minor isoleucine tRNA from Escherichia coli. *J Biol Chem.* 263:9261–9267.

Parniske M. 2008. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nat Rev Microbiol.* 6:763–775.

Perrin A, Buckle M, Dujon B. 1993. Asymmetrical recognition and activity of the I-Scei endonuclease on its site and on intron-exon junctions. *EMBO J.* 12:2939–2947.

Pombert JF, Keeling PJ. 2010. The mitochondrial genome of the entomoparasitic green alga *Helicosporidium*. *PLoS One* 5:e8954.

Schmelzer C, Schweyen RJ. 1986. Self-splicing of group II introns in vitro: mapping of the branch point and mutational inhibition of lariat formation. *Cell* 46:557–565.

Schnare MN, Gray MW. 1990. Sixteen discrete RNA components in the cytoplasmic ribosome of *Euglena gracilis*. *J Mol Biol.* 215:73–83.

Schwarzott D, Walker C, Schüssler A. 2001. *Glomus*, the largest genus of the arbuscular mycorrhizal fungi (glomales), is nonmonophyletic. *Mol Phylogenet Evol.* 21:190–197.

Seif E, Cadieux A, Lang BF. 2006. Hybrid *E. coli*—mitochondrial ribonuclease P RNAs are catalytically active. *RNA* 12:1661–1670.

Seif E, Leigh J, Liu Y, Roewer I, Forget L, Lang BF. 2005. Comparative mitochondrial genomics in zygomycetes: bacteria-like RNase P RNAs, mobile elements and a close source of the group I intron invasion in angiosperms. *Nucleic Acids Res.* 33:734–744.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst Biol.* 51:492–508.

Slater GS, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6:31.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.

Strack D, Fester T, Hause B, Schliemann W, Walter MH. 2003. Arbuscular mycorrhiza: biological, chemical, and molecular aspects. *J Chem Ecol.* 29:1955–1979.

Van der Veen R, Arnberg AC, van der Horst G, Bonen L, Tabak HF, Grivell LA. 1986. Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* 44:225–234.

Vaughn JC, Mason MT, Sper-Whitis GL, Kuhlman P, Palmer JD. 1995. Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric *cox1* gene of *Peperomia*. *J Mol Evol.* 41:563–572.

Vierula PJ, Bertrand H. 1992. A deletion derivative of the kalilo senescence plasmid forms hairpin and duplex DNA structures in the mitochondria of *Neurospora*. *Mol Gen Genet.* 234:361–368.

Waring RB, Davies RW, Scazzocchio C, Brown TA. 1982. Internal structure of a mitochondrial intron of *Aspergillus nidulans*. *Proc Natl Acad Sci U S A.* 79:6332–6336.

Weber F, Dietrich A, Weil JH, Marechal-Drouard L. 1990. A potato mitochondrial isoleucine tRNA is coded for by a mitochondrial gene possessing a methionine anticodon. *Nucleic Acids Res.* 18:5027–5030.