

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

The mitochondrial genome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals two unsuspected trans-splicing events of group I introns

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/96696> since 2017-05-15T11:27:05Z

Published version:

DOI:10.1111/j.1469-8137.2012.04072.x

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

This is the author's final version of the contribution published as:

A. Pelin; J.F. Pombert; A. Salvioli; L. Bonen; P. Bonfante; N. Corradi. The mitochondrial genome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals two unsuspected trans-splicing events of group I introns. *NEW PHYTOLOGIST*. 194(3) pp: 836-845.
DOI: 10.1111/j.1469-8137.2012.04072.x

The publisher's version is available at:
<http://doi.wiley.com/10.1111/j.1469-8137.2012.04072.x>

When citing, please refer to the published version.

Link to this full text:
<http://hdl.handle.net/2318/96696>

The mitochondrial genome of the arbuscular mycorrhizal fungus *Gigaspora margarita* reveals two unsuspected trans-splicing events of group I introns

Adrian Pelin^{1*}, Jean-François Pombert^{2*}, Alessandra Salvioli³, Linda Bonen¹, Paola Bonfante³ and Nicolas Corradi¹

¹Department of Biology, University of Ottawa, Ottawa, K1N 6N5, ON, Canada;

²Department of Botany, University of British Columbia; Vancouver, V6T 1Z4, BC, Canada;

³Dipartimento di Biologia Vegetale, Università di Torino, Torino, I-10125, Italy

Summary

- Arbuscular mycorrhizal fungi (AMF) are ubiquitous organisms that benefit ecosystems through the establishment of an association with the roots of most plants: the mycorrhizal symbiosis. Despite their ecological importance, however, these fungi have been poorly studied at the genome level.
- In this study, total DNA from the AMF *Gigaspora margarita* was subjected to a combination of 454 and Illumina sequencing, and the resulting reads were used to assemble its mitochondrial genome *de novo*. This genome was annotated and compared with those of other relatives to better comprehend the evolution of the AMF lineage.
- The mitochondrial genome of *G. margarita* is unique in many ways, exhibiting a large size (97 kbp) and elevated GC content (45%). This genome also harbors molecular events that were previously unknown to occur in fungal mitochondrial genomes, including *trans*-splicing of group I introns from two different genes coding for the first subunit of the cytochrome oxidase and for the small subunit of the rRNA.
- This study reports the second published genome from an AMF organelle, resulting in relevant DNA sequence information from this poorly studied fungal group, and providing new insights into the frequency, origin and evolution of *trans*-spliced group I introns found across the mitochondrial genomes of distantly related organisms.

Introduction

Arbuscular mycorrhizal fungi (AMF) are ecologically relevant organisms that are involved in one of the most widespread terrestrial symbioses with the roots of most land plants: the mycorrhizal symbiosis. In this symbiosis, the plant partners provide AMF with the sugars necessary for their survival, and the fungal partner increases the acquisition and uptake of several nutrients (e.g. phosphates and nitrates) by the plant from the soil. The presence of AMF in the soil therefore benefits many plants, and has been repeatedly correlated with an increase in the overall biodiversity of terrestrial ecosystems worldwide ([Harrison, 1999](#); [Sanders, 2002, 2003](#); [Rosendahl, 2008](#); [Bonfante & Genre, 2010](#); [Corradi & Charest, 2011](#)). In addition to their tremendous ecological importance, AMF have also been studied for their potentially atypical evolutionary history and genetic system. Indeed, AMF spores and hyphae are coenocytic and multinucleated throughout their entire life cycle, a feature unknown from other fungal relatives, and the real genetic structure of these co-existing nuclei (heterokaryons vs homokaryons) is currently a matter of extensive debate ([Kuhn *et al.*, 2001](#); [Pawlowska & Taylor, 2004](#); [Hijri & Sanders, 2005](#); [Stukenbrock & Rosendahl, 2005](#); [Rosendahl, 2008](#)). AMF are also thought to have propagated in the absence of sexual

reproduction for at least 500 million years (My) ([Redecker et al., 2000](#); [Humphreys et al., 2010](#)), a controversial view that has been challenged recently by reports of recombination in some of their populations ([Vandenkoornhuyse et al., 2001](#); [Croll & Sanders, 2009](#); [den Bakker et al., 2010](#); [Sanders & Croll, 2010](#)), and by the presence of meiosis-related genes in many AMF species ([Halary et al., 2011](#)).

To date, publicly available sequence data from AMF include two mitochondrial genomes from species in the genus *Rhizophagus* ([Lee & Young, 2009](#); [Krüger et al., 2012](#)) and a number of expressed sequence tag (EST) sequences ([Tisserant et al., 2012](#)). These data have been most instrumental in analyzing the position of the different fungal phyla ([Liu et al., 2006, 2009](#); [Lee & Young, 2009](#)), and have revealed some of the genetic and molecular peculiarities of this ecologically critical lineage ([Tisserant et al., 2012](#)). In this study, we expanded the current sampling of available AMF sequence data by acquiring the complete mitochondrial genome sequence from another representative of this group: *Gigaspora margarita*. This species represents an additional model in AMF research, but belongs to a lineage (Diversisporales) for which very little nuclear and organellar sequence information is currently available. This evolutionary lineage is also very distinct from that of better studied species in the genus *Rhizophagus* (Glomerales) ([Krüger et al., 2012](#)), so that inspection of DNA sequences from its members may reveal new insights into their phylogenetic relationships.

Mitochondrial genomes often harbor group I and/or group II introns, self-splicing catalytic ribozymes that must be precisely excised from their surroundings (e.g. mRNA, rRNA or tRNA) for interrupted genes to regain their function. Both introns splice via a series of transesterification reactions ([Johansen et al., 2006](#); [Lambowitz & Zimmerly, 2011](#)). Group I intron splicing starts with the binding of an exogenous guanosine to the catalytic core of the intron, whereas, in group II introns, the splicing mechanism starts with a nucleophilic attack on the 5' splice site of the intron, either by an endogenous bulged adenosine or an exogenous hydroxyl residue. *Trans*- and *cis*-splicing of introns can occur ([Glanz & Kuck, 2009](#)). In *trans*-splicing, polynucleotide chains that are separated at the DNA level are joined together at the RNA level. The *trans*-splicing mechanism has long been thought to be exclusive to group II introns, but the first cases of *trans*-spliced group I introns have been reported recently, almost concomitantly, in three distantly related organisms, that is, a placozoan animal, a green plant and a green alga ([Burger et al., 2009](#); [Grewe et al., 2009](#); [Pombert & Keeling, 2010](#)).

Given the occurrence of *trans*-spliced group I introns in widely dispersed lineages, it is surprising that such events have yet to be found in the mitochondrial genomes of fungi (reviewed in [Bullerwell & Lang, 2005](#); [Lang et al., 2007](#)). A probable reason for this may be the presence of sampling biases across the fungal phylogeny, so that *trans*-spliced introns may actually be present in some of the least-studied fungal lineages. Interestingly, AMF (phylum Glomeromycota) represent one of the fungal lineages whose mitochondrial genomes have been largely unexplored so far ([Raab et al., 2005](#); [Borstler et al., 2008](#); [Lee & Young, 2009](#)). The acquisition of the complete mitochondrial genome from *G. margarita* may therefore reveal the presence of a number of molecular features that were previously unsuspected to occur in fungal mitochondrial genomes.

Materials and Methods

Gigaspora margarita spore cultivation and harvesting

Spores of *Gigaspora margarita* Becker and Hall (isolate BEG34; deposited at the European Bank of Glomeromycota) were used for all experiments. These spores contain endobacteria, according to [Lumini et al. \(2007\)](#), and were produced in pot cultures in a climatic chamber, by inoculation on clover (*Trifolium repens* L.) pots, or *in-vitro* propagated in Petri dishes, on minimal medium, by inoculation of T-DNA-transformed chicory (*Cichorium intybus*) roots with a multispore inoculum of sterile spores ([Lumini et al., 2007](#)). Spores were manually collected, rinsed three times with sterile water and surface sterilized with 3% (w/v) chloramine T and 0.03% (w/v) streptomycin, twice for 10 min, with one round of 10 s of sonication and, finally, rinsed three times for 10 min with sterile water, and stored at 4°C.

Nucleic acid extraction and reverse transcription-polymerase chain reaction (RT-PCR) procedures

Genomic DNA was extracted from *c.* 500 *G. margarita* spores, resulting in a total of 10 µg of total DNA. Briefly, spores were pelleted by centrifugation, resuspended in 300 µl of lysis solution (Epicentre Biotechnologies) containing proteinase K, mixed thoroughly using a vortex and crushed using a plastic pestle. The resulting mixture was immediately incubated at 65°C for 15 min, vortexed for 30 s every 5 min, and then cooled and incubated to 37°C for 30 min at the same temperature on addition of 2 µl of 5 µg µl⁻¹ RNase A. Following RNase treatment, the sample was placed on ice for 5 min, 150 µl of MPC Protein Precipitation Reagent (Epicentre Biotechnologies) were added and the solution was vortexed vigorously for 10 s. Protein debris was pelleted at 4°C for 10 min at a speed of $\geq 10\,000\text{ g}$ and the supernatant was transferred to a new tube. DNA was then precipitated using isopropanol, rinsed twice using 70% ethanol and the DNA was finally suspended in TE buffer.

Total RNA was isolated using the QIAGEN Plant RNA extraction kit following the manufacturer's procedures. In this case, a total of 100 spores was pelleted by centrifugation, crushed using a plastic mortar and treated with a mixture of β -mercaptoethanol and a series of specific buffers supplied by the manufacturer. On extraction, total RNA was treated with RNase-Free DNase I (Epicentre Biotechnologies, Madison, WI, USA) to remove any traces of DNA contamination. In this case, the mixture was gently shaken at 37°C for 1 h in the presence of DNase I and specified buffer.

DNase-treated RNA was subjected to RT-PCR using the iScript kit (Bio-Rad Laboratories) following the manufacturer's recommendations in order to produce cDNA. The cDNA was used for PCR with specific primers, and the resulting amplicons were subjected to Sanger DNA sequencing. This procedure was instrumental to detect and confirm the presence of *trans*-splicing in the *coxI* and *rns* genes of *G. margarita*, and to determine the exact location of several introns within the *rnl* gene. Specifically, a combination of primers located on exons 4 and 5 of the gene *coxI*, and on exons 1 and 2 of the *rns* gene, were used with DNase-treated cDNA as a template. PCRs produced amplicons whose sequences reconstructed a fully functional RNA molecule for both genes. The primers used to identify the *trans*-splicing events were as follows: *cox2F*, 5'-TCCTAGCTGGTGGTATCACTATGCT-3'; *cox5R*, 5'-CGTTGAAGTGATGAAAGGGAGCTG-3'; *rnsF*, 5'-TCACACATGCGAGTCCAATGA-3'; *rnsR*, 5'-TAAGGTGTGGACTAGGAGGGTATCTA-3';

Genome sequencing and *de novo* assembly

A total of 5 µg of *G. margarita* DNA was used to produce one Titanium DNA pyrosequencing library from Roche 454 Life Sciences (Branford, CT, USA) located at the Génome Québec Innovation Centre (McGill University, QC, Canada). A quarter of a Titanium 454 pyrosequencing run was performed, and 23% of the resulting reads were assembled using gsAssembler 2.6 (formerly known as Newbler) into 2828 contigs larger than 500 bp. The largest contigs were found to belong to the bacterial symbiont, *Candidatus Glomeribacter gigasporarum*, whose genome is publicly available ([Ghignone et al., 2012](#)). Among all the contigs obtained, however, a total of six had clear homology with known mitochondrial sequences from other fungi.

In parallel, 500 ng of DNA from *G. margarita* was subjected to deep sequencing using the Illumina technology, with 100-bp paired-end reads (with an average insert of 337 bp), and a quarter of a channel on the GAIIx instrument located at Fasteris S.A. (Geneva, Switzerland). This procedure was used to increase the coverage for many regions of the mitochondrial genome, as well as to confirm the source and validate the assembly of the contigs assembled using 454 pyrosequencing data (i.e. additional quality check based on spatial information from the paired-end reads). The Illumina and 454 reads were reassembled together using the hybrid assembler MIRA 4 ([Chevreux et al., 2004](#)), in combination with the mitochondrial contigs previously identified with the gsAssembler. The contigs resulting from this hybrid assembly were converted to, edited and re-assembled with CONSED 19.0 ([Gordon et al., 1998, 2001](#); [Gordon, 2003](#)), resulting in the assembly of three contigs of obvious mitochondrial origin (i.e. no significant homology with the endosymbiotic bacterial genome or with nuclear homologs from other species) with similar coverage. None of the 454 reads assembling onto the genome data of *Candidatus Glomeribacter gigasporarum* align with the mitochondrial genome of *G. margarita*.

Genome completion, annotation and analysis

PCR with a number of primer sets, bacterial cloning and Sanger sequencing were used in combination with CONSED to link the three contigs of mitochondrial origin into a single, circular molecule with an average coverage of 20× (Supporting information Fig. S1a). The mitochondrial genome of *G. margarita* was annotated manually using Artemis 13.2.0 ([Carver et al., 2008](#)) in combination with BLAST ([Altschul et al., 1997, 2005](#)) procedures against the ‘nonredundant’ repository at the National Center for Biotechnology Information (NCBI). tRNAs were detected using RNaseasel 5.1.2 and tRNAscan-SE 1.21 ([Lowe & Eddy, 1997](#); [Lang et al., 2007](#)). Group I and group II introns were manually detected, and repeats were searched with Tandem Repeats Finder ([Benson, 1999](#)). The mitochondrial genome of *G. margarita* is deposited in GenBank under the accession number [JQ041882](#).

Phylogenetic analyses

In the present study, a phylogeny based on 14 mitochondrial genes from representatives of all extant fungal phyla was reconstructed. The mitochondrial genes include seven NADH dehydrogenase subunit units (*nad1–6* and *nad4L*), three cytochrome oxidase subunits (*cox1*, 2 and 3), three ATP synthase subunits (*atp1*, 6, 8, 9/c) and the apocytochrome b (*cob*). Amino acid sequences of these genes were also obtained from representatives of the ascomycetes, basidiomycetes, chytridiomycetes, blastocladiomycetes, zygomycetes and the AMF *Rhizophagus irregularis* and *G. margarita*.

The amino acid sequences of the different genes were aligned using MUSCLE 3.6 ([Edgar, 2004](#)), and concatenated using bioinformatics tools implemented in Geneious 5.5. Best-fit models of

evolution were chosen using ProtTest 3 ([Darriba et al., 2011](#)). Phylogenies were reconstructed using the maximum likelihood algorithm implemented in PhyML 3.0 ([Guindon et al., 2010](#)) with 1000 bootstraps and the LG + I + Γ model of evolution. Bayesian analyses were performed using MrBayes 3.1.2 ([Huelsenbeck & Ronquist, 2001](#)), the model WAG + I + Γ , with 1 000 000 generations and a 'burn in' of 100 000. In parallel, the CAT + Γ 4 model of amino acid substitution implemented in PhyloBayes 3 ([Lartillot et al., 2009](#)) was also used. In this case, two concurrent chains were run in parallel and were terminated using the PhyloBayes automatic stopping rule (maxdiff < 0.1).

Results

High-throughput DNA sequencing using both the Illumina and 454 pyrosequencing platforms was performed using total DNA extracted from *G. margarita*, resulting in the acquisition of 233 365 400 bp (in 4 667 308 reads) and 95 500 945 bp (in 266 716 reads) of raw sequence data, respectively. Subsequent hybrid Illumina and 454 assemblies resulted in three contigs with clear homology with known mitochondrial genomes, which were linked by PCR and confirmed by Sanger sequencing. These systematic procedures resulted in the acquisition of a single, circular molecule, the mitochondrial map of *G. margarita*, which we annotated using available programs ([Lowe & Eddy, 1997](#); [Lang et al., 2007](#); [Carver et al., 2008](#)). The *G. margarita* mitochondrial genome (mtDNA) maps as a single circular molecule of 96 998 bp ([Fig. 1](#)) featuring strong sequence homology with that of other fungal relatives. The overall GC content (45%) of the *G. margarita* mtDNA is higher than that of many fungi, however, and this increase affects both its coding and noncoding regions equally (e.g. intergenic regions are 45.15% GC vs 43.55% for gene features; tRNA is 49.52% GC and rRNA is 47.12% GC; [Table 1](#)). The genes identified in the *G. margarita* mtDNA ([Table 1](#)) are scattered among the two DNA strands. We took advantage of this genome annotation to test recent hypotheses regarding the placement of AMF within the fungal phylogeny, and phylogenetic analyses resulted in very high statistical support for most branches ([Fig. 2](#)).

Table 1. Genome features of *Gigaspora margarita* and other fungal mtDNAs

	<i>Gigaspora margarita</i>	<i>Rhizophagus irregularis</i>	<i>Saccharomyces cerevisiae</i> ^a	<i>Allomyces macrogyne</i>
1.	^a Corresponds to strain FY1674.			
2.	^b Includes protein-encoding genes with known functions, structural RNAs, as well as free-standing and intronic open reading frames (ORFs).			
3.	^c <i>Cis/trans</i> -spliced.			
4.	[*] <i>Rhizophagus irregularis</i> harbors two ORFs (GIY-YIG endonucleases) that are not present in <i>G. margarita</i> .			
Size (bp)	96 998	70 606	85 779	57 473
GC content (%)				
Genes	43.93	38.33	22.79	33.51
Intergenes	45.15	36.79	14.72	42.41
Total	45.02	37.18	17.11	39.5
Genes ^b				
tRNAs	25	26	24	25
rRNAs	2	2	2	2
Proteins	18	20*	19	25
Total	45	49	46	57
Introns ^c	22	26	13	28
Group I	21	26	9	26
Group II	1	0	4	2
Gene density (kb ⁻¹)	0.185	0.283	0.221	0.434
Coding %	15.5	24.3	28.5	38.5

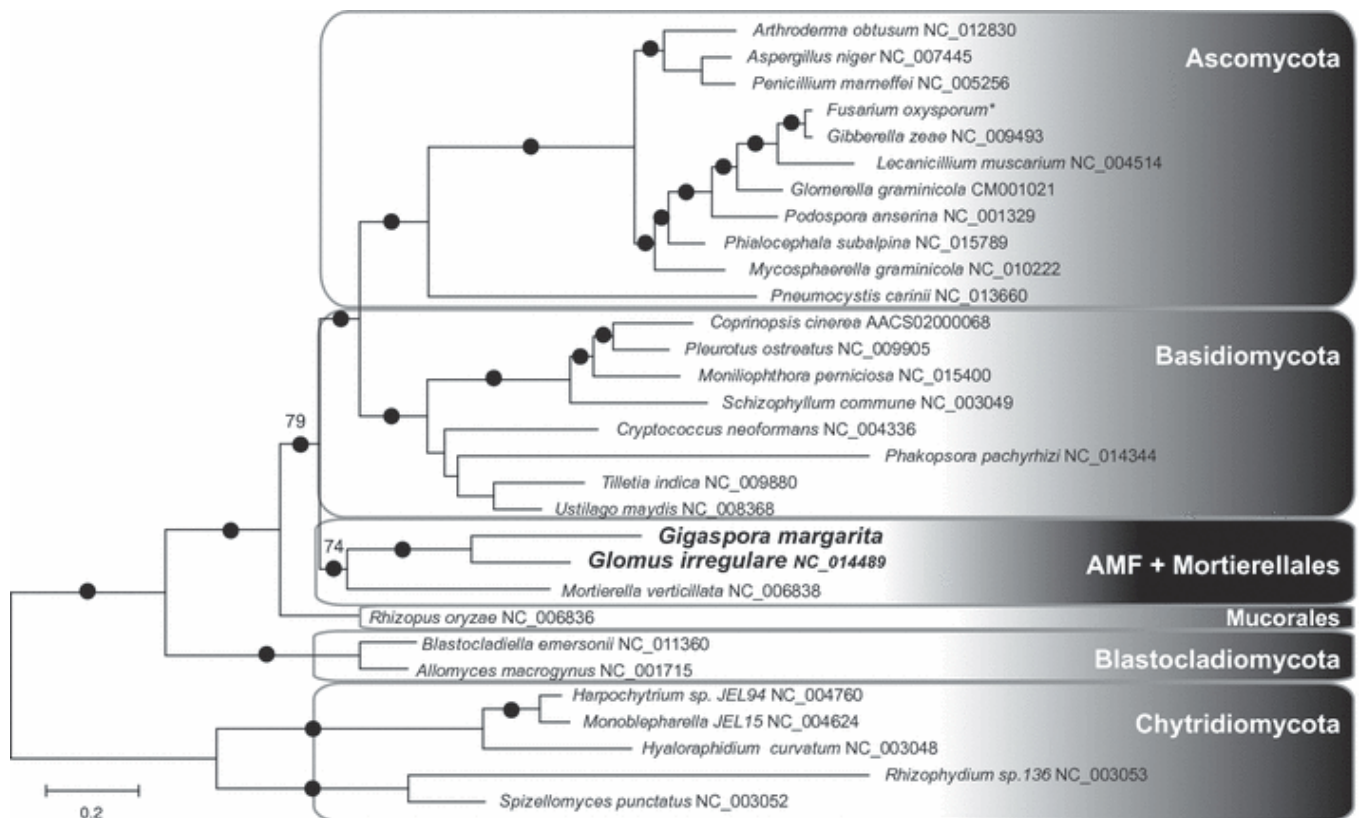


Figure 2. Fungal phylogeny inferred from the concatenation of 14 mitochondrial genes. The major fungal phyla are surrounded by rounded rectangles. The tree is based on an unambiguous alignment of 3668 amino acid positions, and rooted using Chytridiomycota as outgroup. Black circles highlight branches with posterior probabilities of unity for Bayesian analyses performed with MrBayes (WAG + I + Γ model of evolution) and PhyloBayes (CAT + Γ 4 model of evolution). Black circles also indicate a bootstrap support greater than 98% for maximum likelihood analyses performed using PhyML (LG model of evolution), unless a numerical value is present above the black circle. In these cases (two in total), the numerical values represent the bootstrap support for the respective nodes. Scale bar corresponds to 0.2 amino acid substitutions per site. **Fusarium* and *Gibberella* represent different anamorphs.

The coding capacity of the *G. margarita* mitochondrial genome is very similar to that of its closely related species *Rhizophagus irregularis*, with which it shares a complement of 14 protein-encoding genes with known functions, 25 tRNAs and two rRNAs. The *G. margarita* mtDNA features four intronic open reading frames (ORFs) located within the *cox1* and *nad5* genes, and two putative free-standing endonucleases (*orf129* and *orf132*) displaying degenerated single LAGLIDADG motifs (Fig. 1). A second LAGLIDADG motif found downstream of the *orf129* reading frame and separated by a stop codon suggests that this ORF is a pseudo-gene, on its way out of the genome. The *G. margarita* mitochondrial genome also harbors one group II intron within the *rnl* (large subunit of the rRNA) gene. Very little gene order conservation is observed between the two species, reflecting the rapid shuffling of genomic material typical of fungal mitochondrial genomes. Genes are encoded on the different strands in *G. margarita*, but not in *G. intraradices*, and most of the increased size in the *G. margarita* mtDNA relative can be accounted for by the presence of much larger intergenic regions. No evidence for the presence of extensive interspersed repeated elements (Benson, 1999) or heteroplasmy was found. In particular, all reads assembled on our reference map appeared to be monomorphic, and no relevant polymorphism could be identified.

The *G. margarita* mtDNA features a total of 22 introns (21 group I and one group II; [Table 1](#)) inserted within five genes. Of these, two are group I introns that are surprisingly spliced *in trans* ([Fig. 3](#)). These *trans*-spliced group I introns are inserted within the cytochrome oxidase I (*cox1*) and the small subunit of the rRNA (*rns*) genes, respectively. In both cases, exons from these genes are found scattered across the genome on opposite strands. These unusual features were confirmed both *in vitro* by specific PCR products and sequencing, and *in silico*, as these regions are covered by many reads from 454, Illumina and Sanger sequencing technologies ([Fig. S1a,b](#)). The first four *cox1* exons (of eight in total) are located 28 kb apart from the rest of the coding sequence on opposite strands, whereas the two *rns* exons are also on opposite strands, but separated by 24 kb ([Fig. 3a](#)). RT-PCRs using specific primers, followed by sequencing of the resulting products, confirmed that exons of both genes are linked at the RNA level *in trans* ([Fig. 3b](#)). Sequence inspection revealed the presence of potentially *trans*-spliced group I introns, which appear to be broken in the L8 loop in between the P8 pairings ([Fig. 3c](#)). Investigation of *cox1* intron insertion sites revealed that the *G. margarita* *trans*-spliced intron is inserted at the same location as the first of two *trans*-spliced group I introns recently found in *Trichoplax adhaerens* (a primitive metazoan) ([Fig. 4](#)). The *G. margarita* and *T. adhaerens* introns share only traces of sequence similarity within their conserved core. To our knowledge, no other *trans*-spliced group I intron has been reported in *rns*.

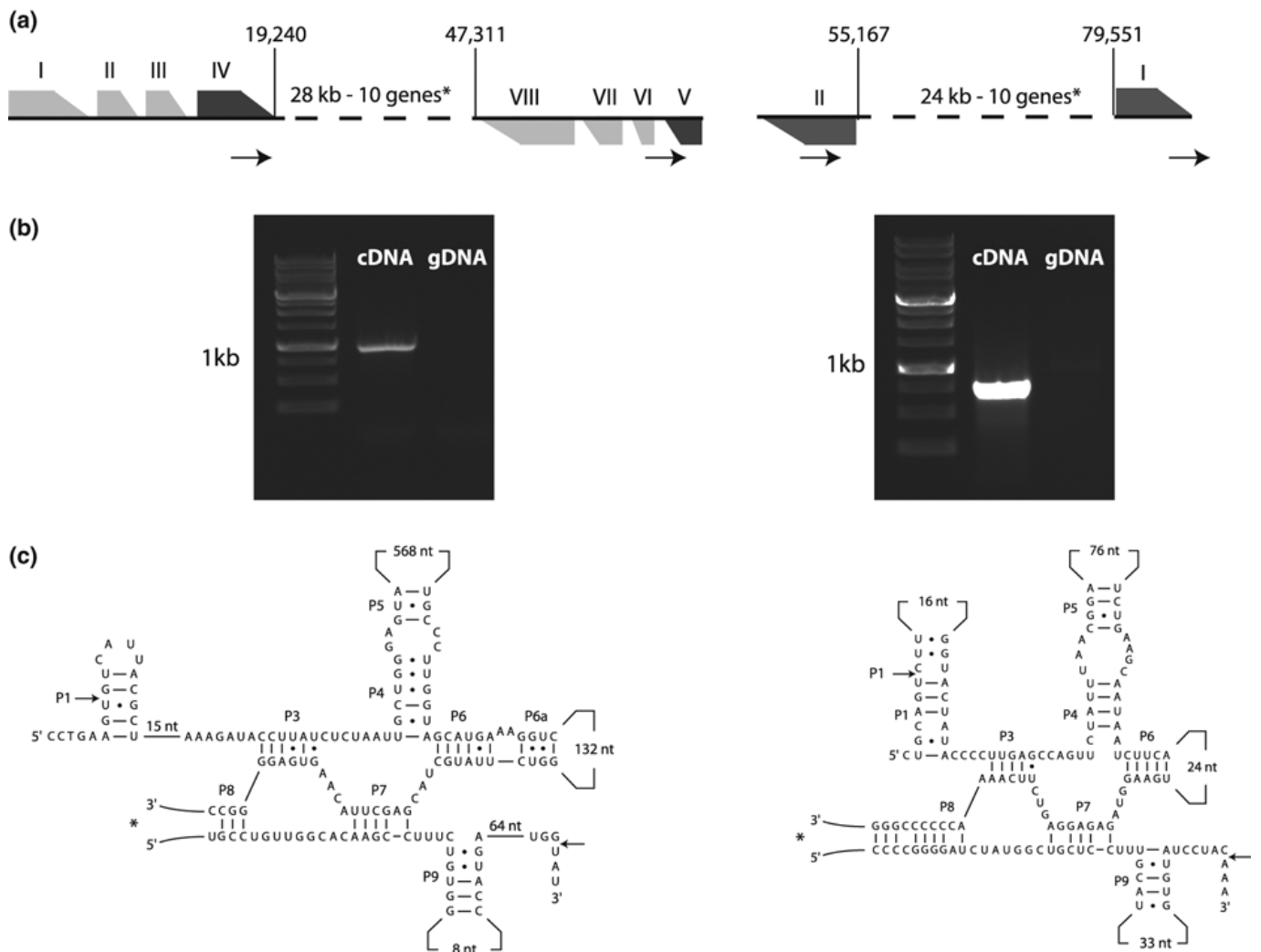


Figure 3. *Trans*-splicing of group I introns in *cox1* and *rns*. (a) Structure of the *cox1* (left) and *rns* (right) exons in the mitochondrial genome of *Gigaspora margarita*. Roman numbers indicate the order and location of the different exons (eight exons in total for *cox1*, left; two exons in total for *rns*, right). Arrows indicate the location and direction of the primers used in reverse transcription-

polymerase chain reaction (RT-PCR) procedures. (b) RT-PCR procedures demonstrating the presence of *trans*-splicing in the *coxI* and *rns* genes. Specific primers produced a band of the expected size and sequence only when used on cDNA. gDNA was used as a control. An extremely faint band appearing using gDNA and *rns*-specific primers represents an unspecific product, which we confirmed by DNA sequencing. (c) Left, *coxI* intron; right, *rns* intron. The GI structures were modeled according to the conventions described in [Burke & RajBhandary \(1982\)](#). The exon/intron splice junctions are denoted by arrows. Canonical Watson–Crick base pairings are indicated by dashes. Guanine–uracyl pairings are indicated by dots. Numbers inside variable loops and above straight lines indicate their length. The putative junctions between the segments of the *trans*-spliced introns are denoted by asterisks (located in L8).

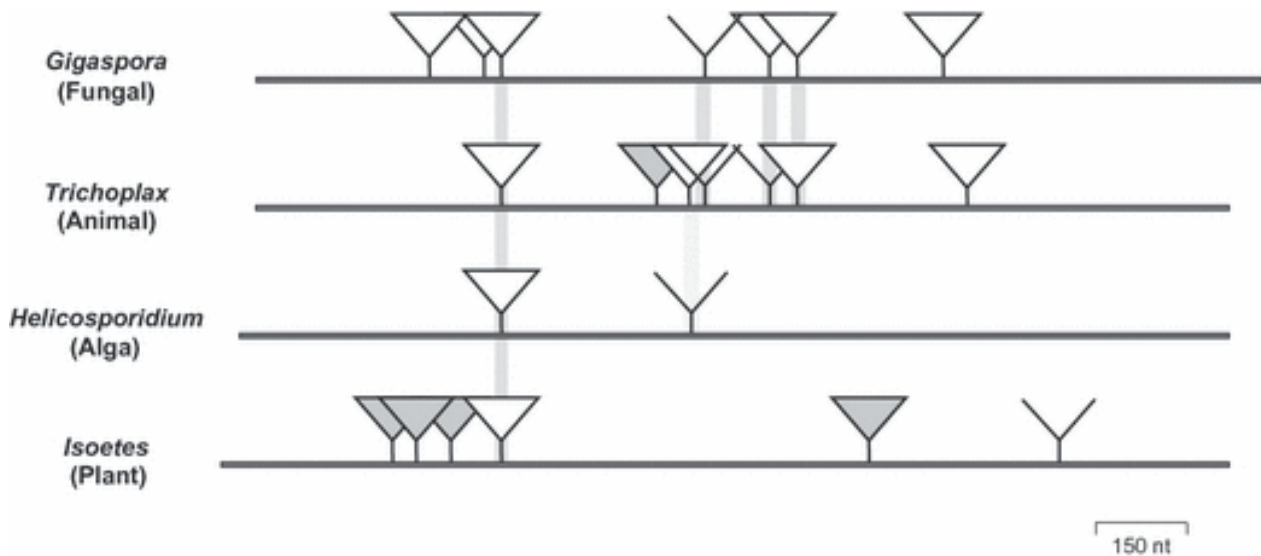


Figure 4. Intron insertion sites in *coxI* from selected organisms of various evolutionary lineages. Location of introns in the mitochondrial *coxI* genes of *Gigaspora margarita* (fungus; this article), *Trichoplax adhaerens* (placozoan animal; [Burger et al., 2009](#)), *Helicosporidium* sp. (nonphotosynthetic green alga; [Pombert & Keeling, 2010](#)) and *Isoetes engelmannii* (vascular plant; [Grewe et al., 2009](#)). The *coxI* coding regions are shown as dark gray bars; cis-splicing and trans-splicing group I introns are depicted by white triangles and Y-shaped symbols, respectively; group II introns are shown by gray triangles. Light gray shading highlights introns located at homologous sites among different organisms. nt, nucleotides. The *coxI* gene in *Selaginella moellendorffii* ([Hecht et al., 2011](#)) also contains a trans-splicing intron at the same position as in the sister lycophyte, *Isoetes engelmannii*.

Discussion

Structure and content of the *G. margarita* mitochondrial genome

The mitochondrial genome of *G. margarita* represents the third AMF organelle genome ever sequenced, and the first from a representative of the Diversiporales. Its coding capacity is similar to that found in other fungi, and is almost identical to that of the AMF *G. intraradices*, the only exceptions being a few unknown, and possibly spurious, ORFs. In particular, the proteome encoded by the genome includes 14 proteins with known functions, involved in a restricted number of processes that are typically found in fungal mitochondria (i.e. electron transport, oxidative phosphorylation) ([Bullerwell & Lang, 2005](#)).

Surprisingly, however, this apparent conservation in the gene repertoire is not mirrored at other levels. For instance, at 97 kbp, the mitochondrial genome of *G. margarita* is at the high end of the size spectrum for the kingdom ([Bullerwell & Lang, 2005](#)). It is also intriguing that the *G. margarita* mitochondrial genome is very GC rich compared with those of most fungi. This variation in GC content is reflected in both coding and noncoding regions, and so this characteristic does not appear to be related to translational constraints. The underlying mechanism giving rise to this nucleotide bias is difficult to assess, but recent analyses of mitochondrial genomes from unicellular organisms characterized by a high GC content have led to suggestions that, rather than being under selective (i.e. adaptive) pressure, GC bias may be a result of neutral processes, such as biased gene conversion or ‘GC drive’ (loss of certain repair functions) ([Barth & Berendonk, 2011](#); [Smith et al., 2011](#)).

Phylogenetic relationships of extant fungal phyla based on mitochondrial protein-encoding genes

The phylogenetic placement of AMF within the fungal tree of life has long been a matter of debate. These plant symbionts were recognized for some time as a potential sister lineage to the higher fungal phyla (i.e. Ascomycota and Basidiomycota; [Schwarzott et al., 2001](#); [James et al., 2006](#)), but recent analyses of concatenated nuclear and mitochondrial sequences have suggested that AMF may represent one of many monophyletic phyla (Glomeromycota; [Hibbett et al., 2007](#)) composing the base of the fungal tree of life ([Corradi & Sanders, 2006](#); [Liu et al., 2006, 2009](#); [Lee & Young, 2009](#)).

In this study, the amino acid sequences of fungal mitochondrial protein-encoding genes were used to determine the relationships of AMF mitochondrial genomes (i.e. *G. margarita* and *G. intraradices*) with those of other members of the kingdom. Our phylogenetic analyses resulted in a robust fungal phylogeny (i.e. posterior probabilities of 1.0 and bootstrap support > 98% for most branches) that appears to be consistent with the most recent reconstructions of the fungal tree based on mitochondrial and nuclear sequence data ([Liu et al., 2006, 2009](#); [Lee & Young, 2009](#)). Specifically, our phylogeny placed the AMF as a potential sister lineage to a basal fungal lineage, the *Mortierellales*, with high statistical support in most cases. Indeed, Bayesian reconstructions using different models resulted in a posterior probability of 1.0 for this particular node, whereas maximum likelihood analyses resulted in a statistical support higher than that reported by others using very similar datasets ([Lee & Young, 2009](#)).

Unfortunately, this strong phylogenetic signal could not be linked to the presence of shared genomic signatures (i.e. unique genes shared between the groups or conserved gene order) between the AMF and the *Mortierellales*, which is most probably a result of the high level of gene shuffling that characterizes fungal mitochondrial genomes. Certainly, similar analyses could strongly benefit from the use of concatenated mitochondrial and nuclear datasets, as a number of discrepancies have been reported between phylogenies reconstructed using organellar and nuclear genes. Together with the addition of many more species representative of the diversity of Glomeromycota and Mucoromycota, such analyses will be critical to resolve the origin and evolution of the most basal lineages of the fungal tree of life.

***Gigaspora margarita* shares a *trans*-spliced group I intron at a cognate site with a placozoan animal**

The phenomenon of *trans*-splicing in organelles has long been thought to be restricted to the group II introns found in plant/algal mitochondria and chloroplasts ([Glanz & Kuck, 2009](#)), but we now know that similar *trans*-splicing events can also involve group I introns within the *coxI* genes of a

metazoan animal (*Trichoplax*), a vascular lycophytic plant (*Isoetes engelmannii*) and a nonphotosynthetic alga (*Helicosporidium*) ([Burger et al., 2009](#); [Grewe et al., 2009](#); [Pombert & Keeling, 2010](#)). In this study, we newly demonstrate that group I *trans*-splicing also occurs in fungal organisms, and that it can also affect the maturation of the SSU rRNA; therefore, the identification of these previously unsuspected events offers a unique opportunity to better comprehend the origin and evolution of these rare mechanisms.

The *G. margarita cox1 trans*-spliced intron (intron 4) is located at the same position as a *cis*-spliced intron 8 homolog in the related fungus *G. intraradices* ([Lee & Young, 2009](#)). Introns of both species share sequence similarity in their core region (e.g. can be folded into a similar group I structure, but the sequence similarity is relatively low), but, in *G. intraradices*, the intron is characterized by the presence of a complete LAGLIDADG ORF, suggesting that it has been acquired recently. Homologous genomic locations have been invaded by *cis*-spliced group I introns in a number of diverse fungi (e.g. *Schizosaccharomyces pombe*, *Gibberella zeae*) ([Lang, 1984](#)), but are intronless in the more closely related species *Rhizopus* and *Mortierella*. Together, these features indicate that this specific genomic location is very prone to intronic invasions, and that the group I intron at this location is likely to have originated recently in the AMF, undergoing subsequent disruption and *trans*-splicing in the lineage leading to *G. margarita*.

Perhaps even more interesting is the fact that the *G. margarita trans*-spliced intron is located at a position homologous to one of two *trans*-spliced introns identified in the *cox1* gene of *Trichoplax*, a primitive animal ([Burger et al., 2009](#)). The *cis*-spliced *cox1* introns found at the same position in other species share significant sequence similarity with their *trans*-counterparts. The structural similarity between *G. margarita* and *T. adhaerens* also extends to the number of introns, as both species share a total of four introns: that is, the same number seen between the two closely related fungi *Gigaspora* and *Rhizophagus* (although only three are at cognate sites in all three organisms). Together, these features are consistent with the view that, after invasion of a genome, copies of the group I mobile element spread to additional compatible sites.

The *trans*-splicing event common to the fungus *G. margarita* and the placozoan *T. adhaerens* has yet to be identified at the same location in other species, but it is worth noting that a similar event has been observed in close vicinity. Indeed, in *Helicosporidium* sp., the *cox1* gene is *trans*-spliced at a site that is 11 nucleotides upstream of that found in *G. margarita* ([Pombert & Keeling, 2010](#)). This position represents a widely recognized ‘hot spot’ for group I introns in various fungi, and is found in *T. adhaerens* and plants ([Cho et al., 1998](#); [Burger et al., 2009](#)). The two closely located *cox1* positions that harbor *trans*-spliced introns map to a highly conserved region of the transmembrane domain VI of the Cox1 protein, and a functionally important histidine codon is located between them ([Shapleigh et al., 1992](#)).

Although it may seem unlikely for *trans*-spliced introns located at identical positions to have independent histories (i.e. in the *cox1* of a fungus and an animal), this site is a known homing site for group I introns, and there is also precedence for the independent conversion of introns from a *cis*- to a *trans*-form. The latter is exemplified by *nad1* intron 4 (group II type) in flowering plants, where the break, in some cases, is upstream of the *matR* intronic ORF and, in other cases, downstream of it ([Qiu & Palmer, 2004](#)). Thus, some group I/II introns may be more susceptible and tolerant to a discontinuous gene structure than others, and this seems to be particularly true for those located in the *cox1* gene. Overall, the notorious invasive nature of group I introns appears to be compatible with an independent and recent acquisition of *trans*-spliced forms at cognate sites in *G. margarita* and *T. adhaerens*.

***Trans*-splicing of group I intron in the small subunit of the rRNA: a recent event?**

Given the hundreds of group I introns that have been identified to date in bacterial and organellar genes, it is surprising that, so far, *trans*-splicing of group I introns has been found only in the *coxI* gene; to our knowledge, the *Gigaspora rns* gene represents the first example of *trans*-splicing of a structural RNA gene in organelles.

The *G. margarita rns* gene harbors all conserved core features found in homologs from other species, but appears to be atypical in other ways. First, the gene is separated into two exons at a very unusual position, the variable region 7, a region of the SSU rRNA structure that is not a known homing site for group I introns. Indeed, very few introns of any type have been found in this region of the SSU rRNA, and only one intron of 950 (790 of which were group I type) was identified nearby, and it was a spliceosomal intron in a nuclear SSU rRNA gene ([Jackson *et al.*, 2002](#)). Accordingly, even in known cases of fragmented mitochondrial *rns* genes, none has been found to be broken within variable region 7 (cf. *Chlamydomonas* species V1, V2, V3, V6, *Scenedesmus* V3, *Tetrahymena* and *Paramecium* V2, *Euglena* V8; [Waller & Jackson, 2009](#); [Barbrook *et al.*, 2010](#); [Spencer & Gray, 2011](#)). Second, the location of the *rns trans*-spliced intron is also unique because introns are typically found in the functionally important regions of the rRNA ([Jackson *et al.*, 2002](#)). The model that is currently favored for the generation of scrambled rRNA molecules is by DNA recombination across repeats located within the variable V regions. In these cases, it has been suggested that the resulting separate RNA molecules can form base pairing within the conserved regions and still function in the ribosome. In *G. margarita*, however, this mechanism does not occur, as this species has used a more classical *trans*-splicing mechanism to link disrupted *rns* exons. So how did this novel *trans*-splicing event arise in the mitochondrial genome of *G. margarita*?

Although *G. intraradices* has a *cis*-splicing counterpart of the *G. margarita coxI trans*-intron, its SSU rRNA gene is intronless. This indicates that the *trans*-splicing event affecting the *rns* gene is unlikely to have derived from their common ancestor, and that it has probably appeared recently in the lineage leading to *Gigaspora*. Whatever the mechanism involved in the origin of this novel *trans*-splicing event, the location of the 3' portion of the *rns* rRNA gene in *G. margarita*, which is close to, but on the opposite strand, relative to the *rnl* gene, indicates that it must have acquired a new promoter on rearrangement in order to be functional (rather than hitch-hiking along in a pre-existing transcriptional unit). Interestingly, the 3' segment of the *trans*-splicing *coxI* gene is located beside the *rns* segment, and therefore it is plausible that the same promoter could have fuelled its expression following a recent disruption of the *rns* gene in *Gigaspora*.

Concluding remarks

The present study further highlights the importance of a broad taxonomic sampling in order to understand the frequency and nature of molecular mechanisms that have long appeared to be very rare, or restricted to a few lineages. Accordingly, the complete mitochondrial genome of *G. margarita* demonstrates that events of group I intron *trans*-splicing may be more common than previously thought. This genome sequence also represents valuable data from a group of organisms for which DNA sequences have been notoriously lacking. This new wealth of AMF sequence data may be used to develop a battery of molecular markers for studies of AMF biodiversity, or to further understand the function of this essential organelle in these ecologically critical organisms, especially during the establishment of the mycorrhizal symbiosis.

Acknowledgements

N.C. is a scholar of the Integrated Microbial Biodiversity Program of the Canadian Institute for Advanced Research (CIFAR-IMB). This work was supported by NSERC-Discovery grants to N.C. and L.B. Work from P.B. and A.S. was funded by a grant from the Regione Piemonte (Converging Technologies Project). We would like to thank three anonymous reviewers for their constructive comments on a previous version of the manuscript, and Mohammed Selman for technical assistance.

References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389–3402.
- Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, Yu YK. 2005. Protein database searches using compositionally adjusted substitution matrices. *FEBS Journal* 272: 5101–5109.
- den Bakker HC, Vankuren NW, Morton JB, Pawlowska TE. 2010. Clonality and recombination in the life history of an asexual arbuscular mycorrhizal fungus. *Molecular Biology and Evolution* 27: 2474–2486.
- Barbrook AC, Howe CJ, Kurniawan DP, Tarr SJ. 2010. Organization and expression of organellar genomes. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences* 365: 785–797.
- Barth D, Berendonk TU. 2011. The mitochondrial genome sequence of the ciliate *Paramecium caudatum* reveals a shift in nucleotide composition and codon usage within the genus *Paramecium*. *BMC Genomics* 12: 272.
- Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27: 573–580.
- Bonfante P, Genre A. 2010. Mechanisms underlying beneficial plant–fungus interactions in mycorrhizal symbiosis. *Nature Communications* 1: 48.
- Borstler B, Raab PA, Thiery O, Morton JB, Redecker D. 2008. Genetic diversity of the arbuscular mycorrhizal fungus *Glomus intraradices* as determined by mitochondrial large subunit rRNA gene sequences is considerably higher than previously expected. *New Phytologist* 180: 452–465.
- Bullerwell CE, Lang BF. 2005. Fungal evolution: the case of the vanishing mitochondrion. *Current Opinion in Microbiology* 8: 362–369.
- Burger G, Yan Y, Javadi P, Lang BF. 2009. Group I-intron trans-splicing and mRNA editing in the mitochondria of placozoan animals. *Trends in Genetics* 25: 381–386.
- Burke JM, RajBhandary UL. 1982. Intron within the large rRNA gene of *N. crassa* mitochondria: a long open reading frame and a consensus sequence possibly important in splicing. *Cell* 31: 509–520.
- Carver T, Berriman M, Tivey A, Patel C, Bohme U, Barrell BG, Parkhill J, Rajandream MA. 2008. Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* 24: 2672–2676.
- Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Muller WE, Wetter T, Suhai S. 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Research* 14: 1147–1159.
- Cho Y, Qiu YL, Kuhlman P, Palmer JD. 1998. Explosive invasion of plant mitochondria by a group I intron. *Proceedings of the National Academy of Sciences, USA* 95: 14244–14249.
- Corradi N, Charest C. 2011. Some like it toxic. *Molecular Ecology* 20: 3289–3290.
- Corradi N, Sanders IR. 2006. Evolution of the P-type II ATPase gene family in the fungi and presence of structural genomic changes among isolates of *Glomus intraradices*. *BMC Evolutionary Biology* 6: 21.
- Croll D, Sanders IR. 2009. Recombination in *Glomus intraradices*, a supposed ancient asexual arbuscular mycorrhizal fungus. *BMC Evolutionary Biology* 9: 13.
- Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27: 1164–1165.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

- Ghignone S, Salvioli A, Anca I, Lumini E, Ortu G, Petiti L, Cruveiller S, Bianciotto V, Piffanelli P, Lanfranco L et al. 2012. The genome of the obligate endobacterium of an AM fungus reveals an interphylum network of nutritional interactions. *ISME Journal* 6: 136–145.
- Glanz S, Kuck U. 2009. Trans-splicing of organelle introns – a detour to continuous RNAs. *Bioessays* 31: 921–934.
- Gordon D. 2003. Viewing and editing assembled sequences using Consed. *Current Protocols in Bioinformatics*. Chapter 11: Unit 11.2.
- Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. *Genome Research* 8: 195–202.
- Gordon D, Desmarais C, Green P. 2001. Automated finishing with autofinish. *Genome Research* 11: 614–625.
- Grewe F, Viehoveer 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 Research* 37: 5093–5104.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307–321.
- Halary S, Malik SB, Lildhar L, Slamovits CH, Hijri M, Corradi N. 2011. Conserved meiotic machinery in *Glomus* spp., a putatively ancient asexual fungal lineage. *Genome Biology and Evolution* 3: 950–958.
- Harrison MJ. 1999. Molecular and cellular aspects of the arbuscular mycorrhizal symbiosis. *Annual Review of Plant Physiology and Plant Molecular Biology* 50: 361–389.
- 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 Biology and Evolution* 3: 344–358.
- Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, Huhndorf S, James T, Kirk PM, Lücking R et al. 2007. A higher-level phylogenetic classification of the Fungi. *Mycological Research* 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.
- Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
- Humphreys CP, Franks PJ, Rees M, Bidartondo MI, Leake JR, Beerling DJ. 2010. Mutualistic mycorrhiza-like symbiosis in the most ancient group of land plants. *Nature Communications* 1: 103.
- Jackson S, Cannone J, Lee J, Gutell R, Woodson S. 2002. Distribution of rRNA introns in the three-dimensional structure of the ribosome. *Journal of Molecular Biology* 323: 35–52.
- James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J et al. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443: 818–822.
- Johansen SD, Vader A, Sjøttem E, Nielsen H. 2006. In vivo expression of a group I intron HEG from the antisense strand of *Didymium* ribosomal DNA. *RNA Biology* 3: 157–162.
- Krüger M, Krüger C, Walker C, Stockinger H, Schüßler A. 2012. Phylogenetic reference data for systematics and phylotaxonomy of arbuscular mycorrhizal fungi from phylum to species level. *New Phytologist* 193: 970–984.
- Kuhn G, Hijri M, Sanders IR. 2001. Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi. *Nature* 414: 745–748.
- Lambowitz AM, Zimmerly S. 2011. Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harbor Perspectives in Biology* 3: a003616.

- 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 Journal* 3: 2129–2136.
- Lang BF, Laforest MJ, Burger G. 2007. Mitochondrial introns: a critical view. *Trends in Genetics* 23: 119–125.
- Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25: 2286–2288.
- 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 Phytologist* 183: 200–211.
- 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 Evolutionary Biology* 9: 272.
- Liu YJ, Hodson MC, Hall BD. 2006. Loss of the flagellum happened only once in the fungal lineage: phylogenetic structure of kingdom Fungi inferred from RNA polymerase II subunit genes. *BMC Evolutionary Biology* 6: 74.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* 25: 955–964.
- Lumini E, Bianciotto V, Jargeat P, Novero M, Salvioli A, Faccio A, Becard G, Bonfante P. 2007. Presymbiotic growth and spore morphology are affected in the arbuscular mycorrhizal fungus *Gigaspora margarita* cured of its endobacteria. *Cellular Microbiology* 9: 1716–1729.
- Pawlowska TE, Taylor JW. 2004. Organization of genetic variation in individuals of arbuscular mycorrhizal fungi. *Nature* 427: 733–737.
- Pombert JF, Keeling PJ. 2010. The mitochondrial genome of the entomoparasitic green alga *Helicosporidium*. *PLoS ONE* 5: e8954.
- Qiu YL, Palmer JD. 2004. Many independent origins of trans-splicing of a plant mitochondrial group II intron. *Journal of Molecular Evolution* 59: 80–89.
- Raab PA, Brennwald A, Redecker D. 2005. Mitochondrial large ribosomal subunit sequences are homogeneous within isolates of *Glomus* (arbuscular mycorrhizal fungi, Glomeromycota). *Mycological Research* 109: 1315–1322.
- Redecker D, Kodner R, Graham LE. 2000. Glomalean fungi from the Ordovician. *Science* 289: 1920–1921.
- Rosendahl S. 2008. Communities, populations and individuals of arbuscular mycorrhizal fungi. *New Phytologist* 178: 253–266.
- Sanders IR. 2002. Ecology and evolution of multigenomic arbuscular mycorrhizal fungi. *American Naturalist* 160: S128–S141.
- Sanders IR. 2003. Preference, specificity and cheating in the arbuscular mycorrhizal symbiosis. *Trends in Plant Science* 8: 143–145.
- Sanders IR, Croll D. 2010. Arbuscular mycorrhiza: the challenge to understand the genetics of the fungal partner. *Annual Review of Genetics* 44: 271–292.
- Schwarzott D, Walker C, Schüßler A. 2001. *Glomus*, the largest genus of the arbuscular mycorrhizal fungi (Glomales), is nonmonophyletic. *Molecular Phylogenetics and Evolution* 21: 190–197.
- Shapleigh JP, Hosler JP, Tecklenburg MM, Kim Y, Babcock GT, Gennis RB, Ferguson-Miller S. 1992. Definition of the catalytic site of cytochrome c oxidase: specific ligands of heme a and the heme a₃-CuB center. *Proceedings of the National Academy of Sciences, USA* 89: 4786–4790.
- Smith DR, Burki F, Yamada T, Grimwood J, Grigoriev IV, Van Etten JL, Keeling PJ. 2011. The GC-rich mitochondrial and plastid genomes of the green alga *Coccomyxa* give insight into the evolution of organelle DNA nucleotide landscape. *PLoS ONE* 6: e23624.

- Spencer DF, Gray MW. 2011. Ribosomal RNA genes in *Euglena gracilis* mitochondrial DNA: fragmented genes in a seemingly fragmented genome. *Molecular Genetics and Genomics* 285: 19–31.
- Stukenbrock EH, Rosendahl S. 2005. Clonal diversity and population genetic structure of arbuscular mycorrhizal fungi (*Glomus* spp.) studied by multilocus genotyping of single spores. *Molecular Ecology* 14: 743–752.
- Tisserant E, Kohler A, Dozolme-Seddas P, Balestrini R, Benabdellah K, Colard A, Croll D, Da Silva C, Gomez SK, Koul R et al. 2012. The transcriptome of the arbuscular mycorrhizal fungus *Glomus intraradices* (DAOM 197198) reveals functional tradeoffs in an obligate symbiont. *New Phytologist* 193: 755–769.
- Vandenkoornhuyse P, Leyval C, Bonnin I. 2001. High genetic diversity in arbuscular mycorrhizal fungi: evidence for recombination events. *Heredity* 87: 243–253.
- Waller RF, Jackson CJ. 2009. Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *Bioessays* 31: 237–245.