

Mitochondrial Genome Rearrangements in *Glomus* Species Triggered by Homologous Recombination between Distinct mtDNA Haplotypes

Denis Beaudet*, Yves Terrat[†], Sébastien Halary[†], Ivan Enrique de la Providencia, and Mohamed Hijri*

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

*Corresponding author: E-mail: denis.beaudet@umontreal.ca; mohamed.hijri@umontreal.ca.

[†]These authors contributed equally to this work.

Accepted: August 1, 2013

Data deposition: The annotated sequences of the three *Glomus* species mitochondrial genomes have been deposited in GenBank under the accessions KC164354–KC164356.

Abstract

Comparative mitochondrial genomics of arbuscular mycorrhizal fungi (AMF) provide new avenues to overcome long-lasting obstacles that have hampered studies aimed at understanding the community structure, diversity, and evolution of these multinucleated and genetically polymorphic organisms. AMF mitochondrial (mt) genomes are homogeneous within isolates, and their intergenic regions harbor numerous mobile elements that have rapidly diverged, including homing endonuclease genes, small inverted repeats, and plasmid-related DNA polymerase genes (*dpo*), making them suitable targets for the development of reliable strain-specific markers. However, these elements may also lead to genome rearrangements through homologous recombination, although this has never previously been reported in this group of obligate symbiotic fungi. To investigate whether such rearrangements are present and caused by mobile elements in AMF, the mitochondrial genomes from two *Glomeraceae* members (i.e., *Glomus cerebriforme* and *Glomus sp.*) with substantial mtDNA synteny divergence, were sequenced and compared with available glomeromycotan mitochondrial genomes. We used an extensive nucleotide/protein similarity network-based approach to investigate *dpo* diversity in AMF as well as in other organisms for which sequences are publicly available. We provide strong evidence of *dpo*-induced inter-haplotype recombination, leading to a reshuffled mitochondrial genome in *Glomus sp.* These findings raise questions as to whether AMF single spore cultivations artificially underestimate mtDNA genetic diversity. We assessed potential *dpo* dispersal mechanisms in AMF and inferred a robust phylogenetic relationship with plant mitochondrial plasmids. Along with other indirect evidence, our analyses indicate that members of the Glomeromycota phylum are potential donors of mitochondrial plasmids to plants.

Key words: arbuscular mycorrhizal fungi, mitochondrial genome rearrangements, homologous recombination, plasmid-related DNA polymerase genes, short inverted repeats, nucleotide/protein similarity network.

Introduction

The association between arbuscular mycorrhizal fungi (AMF) and plant roots is one of the most widespread symbioses involving plants, and thus has an important role in terrestrial ecosystems (Smith and Read 2008). In exchange for carbohydrates, AMF improve plant fitness by enhancing mineral nutrient and water uptake (Ruiz-Lozano et al. 1995). AMF have also been shown to alleviate salt stress (Evelin et al. 2009) and to provide protection against root pathogens (St-Arnaud and Vujanovic 2007; Ismail and Hijri 2012). Despite the fact that these symbioses contribute to important services in natural, agricultural and reclaimed ecosystems (Gianinazzi et al. 2010),

the species richness, community structure and functional diversity of AMF is not well understood (Wehner et al. 2010) due to a lack of reliable molecular tools. The intra-isolate genetic diversity of nuclear DNA and RNA sequences observed in AMF (Sanders et al. 1995; Kuhn et al. 2001; Pawlowska and Taylor 2004; Mathimaran et al. 2008; Stockinger et al. 2009; Boon et al. 2010), has made it difficult to determine evolutionary relatedness at high levels of resolution (i.e., between genetically similar species and/or isolates).

In contrast, the first mitochondrial (mt) genome sequenced from *Glomus irregulare* (synonym *Rhizophagus irregularis*; formerly *G. intraradices*) was shown to be homogeneous

(Lee and Young 2009), a state often referred to as homoplasmic. Since then, the mitochondrial large subunit rRNA gene has been explored for its usefulness as a marker (Raab et al. 2005; Börstler et al. 2008; Thiéry et al. 2010), but determination of its specificity at the isolate level is challenging due to a lack of genetic variability. Recently, the mitochondrial genomes of six *G. irregulare* isolates (Formey et al. 2012) and a closely related species (Beaudet et al. 2013) were compared. Striking features of these data sets were the presence and variability among isolates of numerous mobile elements such as plasmid-related DNA polymerase genes (*dpo*), short inverted repeats (SIRs), and homing endonuclease genes, which have given rise to mitochondrial gene hybrids through horizontal gene transfer (HGT). Intra-specific divergence in the eroded portions of these elements revealed the high value of designing much-needed isolate-specific molecular markers (Corradi and Bonen 2012).

Mitochondrial DNA has long been considered a valuable tool in evolutionary biology because of its high mutation rate and the assumption that individuals only possessed one mtDNA haplotype, directly maternally inherited, without recombination (Birky et al. 1978; Birky 2001). However, reports of mitochondrial recombination events and heteroplasmy in animals, but especially in plants and fungi, led to re-evaluate this assertion (Barr et al. 2005; White et al. 2008). Molecular recombination is an important evolutionary mechanism that allows deleterious mutations removal and generates allelic diversity in a population. The main causes of heteroplasmy are bi-parental inheritance and/or mutations, which result in the coexistence of mitochondrial genomes that differ in either nucleotide composition (site heteroplasmy) or length (length heteroplasmy). The latter can involve large-scale rearrangements or insertion/loss of coding regions (Boursot et al. 1987; Volz-Lingenhöhl et al. 1992). In AMF, transient mitochondrial length heteroplasmy through anastomosis has recently been demonstrated for geographically distant *G. irregulare* in vitro isolates (de la Providencia et al. 2013). The divergence in length between the two mtDNA haplotypes was caused by a variation in plasmid-related *dpo* insertions. The integration of elements such as *dpo* and SIRs has been shown to produce genomic rearrangements through recombination in other organisms (Schofield et al. 1992; Bi and Liu 1996; Paquin et al. 2000; Ferandon et al. 2008; Tanaka et al. 2012).

Integrated plasmid-related *dpo* genes have been reported in all AMF mitochondrial genomes that have been sequenced thus far. Plasmids are self-replicating extra-chromosomal DNA molecules that were originally identified in bacteria. Analogous molecules have now been reported in eukaryotes, notably plants and numerous fungal species. The majority of plasmids described in filamentous fungi are strictly mitochondrial, but cases of nuclear localization have been reported in yeasts (reviewed in Griffiths [1995]) and a *Mucor*-like fungus (Hänfler et al. 1992). Plasmid structure varies widely, from

circular molecules to linear invertrons with terminal inverted repeats (TIRs). Both types harbor one or two open reading frames (ORFs) encoding DNA and/or RNA polymerase genes (Nargang et al. 1984; Akins et al. 1988; Pande et al. 1989; Sakaguchi 1990; Shiu-Shing Chan et al. 1991; Court and Bertrand 1992; Li and Nargang 1993). Plasmids have also been shown to integrate into fungal mtDNA through recombination and can either be cryptic or elicit changes in phenotypic expression (Akins et al. 1986; Robison et al. 1991; Griffiths 1992; Hänfler et al. 1992; Oeser et al. 1993; Hermanns et al. 1994). Such insertions have been reported to be acquired vertically from a common ancestor in plants (Robison and Wolyn 2005) and also in fungi (Robison et al. 1991). However, numerous cases of HGT of those elements have been reported in fungi, offering an explanation for their ubiquitous distribution in this group (Collins and Saville 1990; Griffiths et al. 1990; Kempken et al. 1992; Arganoza et al. 1994; Debets et al. 1994). Furthermore, plant mitochondrial plasmids are suspected to have been acquired from a fungal donor based on their similar structure and low GC content, which is a common feature of fungal mitochondrial genomes (reviewed in Rosewitch and Kistler [2000]).

Another type of mobile element found in the mtDNA of Glomeromycota, SIRs (Formey et al. 2012), were observed in the mitochondrial genome of the green algae *Chlamydomonas reinhardtii* (Boer and Gray 1991), but their structure is reminiscent of GC clusters in yeast (de Zamaroczy and Bernardi 1986) and the *Pst* I palindromes of *Neurospora crassa* (Yin et al. 1981). SIRs can be folded into hairpin secondary structures, and have been suspected to have created genome rearrangements in *Chlamydomonas* spp. by intramolecular recombination (Denovan-Wright and Lee 1994; Nedelcu and Lee 1998) through a model that is analogous to what has been observed in fungi (Almasan and Mishra 1991; Weiller et al. 1991), plants (André et al. 1992) and bacteria (Schofield et al. 1992; Bi and Liu 1996). They are putatively mobile, although their transposition mechanism is not well understood (Grindley and Reed 1985; Nakazono et al. 1994). Nevertheless, mobility has been demonstrated in similar repeats that fold into double hairpin structures in *Allomyces* spp. mtDNA (Paquin et al. 2000).

No information about the underlying mechanisms and consequences of such integrations into the mtDNA of Glomeromycota is currently available. Understanding whether mitochondrial recombination occurs in AMF is important because it can modify the interpretation of several aspects of research, including phylogenetic and population genetic analyses (Ballard and Whitlock 2004). To investigate this issue, we compared two novel mitochondrial genomes from the Glomeraceae family with a *G. irregulare* isolate. As all AMF mitochondrial genomes from the Glomeraceae family that have been sequenced so far share the same mitochondrial synteny, we used this as a comparative basis for our study. We assessed mitochondrial genome evolution with regards to

two other species of this ubiquitous AMF family, which showed divergence in mitochondrial gene organization. We wanted to know whether the new mtDNA reported here, from the species *Glomus sp.* DAOM-240422 and *G. cerebriforme* DAOM-227022, could shed light on the impact of mobile element insertions with regards to genome organization, and thus assessed the diversity, distribution, and propagation of mitochondrial *dpo* sequences within the Glomeromycota phylum. We also explored their evolutionary relatedness to other taxa, with respect to their global positioning within all known plasmid-related *dpo* sequences in the public databases.

Materials and Methods

Fungal Material and DNA Extraction

Spores and mycelium of *G. irregulare* (DAOM-234179), *Glomus sp.* (DAOM-240422), and *G. cerebriforme* (DAOM-229022) were cultivated in vitro on a minimal (M) medium solidified with 0.4% (w/v) gellan gum (Gelzantm, Gelrite) in association with Ri-T transformed carrot roots following the protocol described by Bécard and Fortin (1988). Plates were incubated in the dark in an inverted position at 25 °C, and after several weeks, abundant mycelia and spores were produced. Spores and mycelia were extracted from M medium by solubilization of the gellan gum in 10 mM sodium acetate-citrate buffer (pH 6) (Doner and Becard 1991) and washed in sterile water. The resulting fungal material was further purified by hand under a binocular microscope to remove root fragments.

Spores and mycelium were suspended in 400 μ L of the DNeasy Plant Mini Kit AP1 buffer (Qiagen) and crushed with a pestle in 1.5 ml micro tubes, and the DNA was purified according to the manufacturer's recommendations. Purified DNA in a final elution volume of 40 μ L was stored at –20 °C until use.

Long Polymerase Chain Reactions

The long polymerase chain reaction (PCR) protocol was performed using Takara LA Taq (Takara Bio, Canada) following the manufacturer's recommendations in a volume of 50 μ L containing 0.2 mM deoxynucleotide (dNTP) solution mix, 1.5 mM MgCl₂, and 0.5 μ M of each primer and 1 μ L of template DNA. The primer pairs summarized in table 1 were used to validate the synteny in *Glomus sp.* DAOM-240422. Cycling parameters were 94 °C/3 min, followed by 38 cycles of: 94 °C/30 s, 54 °C/25 s, 68 °C/12 min, and a final elongation at 72 °C for 8 min. PCR products were separated by electrophoresis in a 1% (w/v) agarose gel and visualized with GelRed under ultraviolet light.

Sequencing, Assembly, and Gene Annotation

Glomus irregulare DAOM-234179, *Glomus sp.* DAOM-240422, and *G. cerebriforme* total DNA was sequenced

using 454 Titanium FLX shotgun technology and the resulting reads were assembled in one mitochondrial contig using Newbler (Genome Quebec Innovation Center, McGill University, Montreal, Canada). Gene annotation was performed with the automated organellar annotation software MFannot (<http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl>, last accessed August 23, 2013) (for more information on MFannot see Burger et al. [2013]), followed by manual inspection for the presence of frame shifts and the introduction of missing gene features. SIRs were also manually annotated using NCBI BLAST, and their secondary structure was predicted using the RNAfold web server (Gruber et al. 2008). The annotated sequences of the three new AMF mitochondrial genomes were deposited in GenBank under the accession numbers KC164354, KC164355, and KC164356 for *G. irregulare* DAOM-234179, *Glomus sp.* DAOM-240422, and *G. cerebriforme*, respectively. Circular and linear maps of the mitochondrial genomes (figs. 1 and 2) were created using OGDRAW v.1.1 (Lohse et al. 2007).

Network Analyses

Two similarity network analyses, where each node represents a *dpo* sequence and an edge represents a sequence similarity/identity were performed using EGN (Halary et al. 2013). The first network was constructed from a data set of 91 Glomeromycotan *dpo* sequences (from our collection and retrieved from GenBank). All of the sequences were compared with each other using BLASTn (Altschul et al. 1990) according to the following parameters: reciprocal best hit with a minimum of 1E–10 e-value, and 30% minimum identity covering at least 30% of the smallest sequence. The second and larger network includes an extensive sampling of all DNA polymerases from a broad range of taxa. This data set was built using translated Glomeromycota *dpo* sequences as queries for BLASTp searches on the Genbank nr database. After retrieving hits that showed a minimal e-value of 1E–05, we performed a second BLASTp search to be confident that we had covered the largest possible diversity of DNA polymerase gene sequences. The protein similarity network was calculated using BLASTp with a minimal e-value threshold of 1E–40, 20% minimal similarity covering at least 20% of the smallest sequence. The network layouts were further produced by Cytoscape software, using an edge-weighted force-directed model, meaning that genes sharing more DNA identity/protein similarity appear closer in the display.

Phylogenetic Analyses

Maximum likelihood phylogenetic analyses of AMF mitochondrial genes were performed with a data set of seven mitochondrial gene nucleotide sequences (*atp6*, *atp9*, *cox2*, *nad1*, *nad2*, *nad4*, and *nad6*) that were concatenated using DAMBE software version 5.2.13 (Xia and Xie 2001). The alignment was done using MUSCLE version 3.8.31. The analyses were

Table 1

Long PCR Primers Used to Validate the *Glomus sp.* DAOM-240422 Mitochondrial Synteny

mtDNA region	Primers	Sequences (5'–3')	Size (bp)
<i>atp9-cox1</i>	F	CTTGGCTCTATTGCGCTTAATGA	5,029
	R	ACCAGGAAGAAGATCATAACGA	
<i>nad4L-rrl</i>	F	TGATAGGATTGATGGGTTTCATAG	13,094
	R	GAACATACTTAGCTTTGATGATGGT	
<i>cox3-rns</i>	F	TACGGGTGTACAGCTCTATGAGT	13,610
	R	TGGACTACGAGGGTATCTAATCCT	
<i>cox1 control</i>	F	TGCTAAAGATATTGGGACTCTCT	8,785
	R	GTAATCTGGTATTCTTCGAGGCA	

performed using the GTR + G model (with five distinct gamma categories), with 1,000 bootstrap replicates. Bayesian analysis was performed using MrBayes version 3.2 with the GTR + G model (with five distinct gamma categories), four independent chains, one million cycles, tree sampling every 10 generations and a burn-in value of 40% (supplementary fig. S1, Supplementary Material online), which was adequate for Markov chain convergence.

The global *dpo* protein phylogeny (fig. 4B) was performed with a data set extracted from the connected component of the protein similarity network containing fungal and plant sequences. The MCODE Cytoscape plugin (Bader and Hogue 2003) was used to extract the most connected cluster, that is, the cluster of nodes which connect the most sequences to each other. This process allows a sampling in which sequences will share a maximum similarity, improving the downstream phylogenetic inference. The selected protein sequences were locally aligned with COBALT version 2.01 (Papadopoulos and Agarwala 2007). The maximum likelihood phylogenetic analysis was performed with RaxML using the rtREV model with the CAT option (PROTRREVCAT) and 1,000 bootstrap replicates. This phylogenetic model is an amino acid substitution matrix that has been shown to be appropriate for the analyses of polymerase genes (Dimmic et al. 2002). Further, Bayesian inference was performed with MrBayes using the rtREV + CAT model, four independent chains, one million cycles, tree sampling every 10 generations, and a burn-in value of 40%, which was adequate for Markov chain convergence.

The Glomeromycota *dpo* phylogeny (supplementary fig. S6, Supplementary Material online) was performed on a conserved domain of 75 amino acids that was present in 60 out of the 91 AMF *dpo* sequences. The sequence alignment was done with COBALT version 2.01. The maximum likelihood analysis was conducted with the rtREV + G model (with five distinct gamma categories), with 1,000 bootstrap replicates. Bayesian inference was performed with MrBayes using the rtREV + G model, four independent chains, one million cycles, tree sampling every 10 generations, and a burn-in value of 40%, which was adequate for Markov chain convergence.

All three phylogenetic analyses were done using the integrative software TOPALi version 2.5 (Milne et al. 2004). The best phylogenetic model was determined using TOPALi version 2.5 model selection test based on the Akaike information criterion. The tree figures were completed using TreeGraph version 2.0.47 (Stover and Muller 2010).

Recombination Analyses

The recombination analyses were performed on the Glomeromycota nucleotide identity network group 1, 2, and 8 (where the *Glomus sp.* 240422 *dpo* inserted in the reshuffled intergenic regions clustered). To easily detect the presence of distinct recombination signatures, a distance-based recombination analyses (supplementary fig. S7, Supplementary Material online) was achieved with the recombination analyses tool (RAT) (Etherington et al. 2005), with the low threshold set to 70% nucleotide identity and the high threshold set to over 90% identity. To further confirm the occurrence of the putative recombination events, we tested their statistical significance with a phylogenetic method using the Hidden Markov Model (HMM) with the F84 + gaps nucleotide substitution model (Husmeier and Wright 2001), performed with TOPALi v2.5 (default parameters).

Results

Mitochondrial Genome Description and Comparison

The complete mtDNA sequences of *G. irregulare* (DAOM-234179), *Glomus sp.* (DAOM-240422), and *G. cerebriforme* (DAOM-229022) are double-stranded circular DNA molecules, are homogeneous within isolate with no DNA sequence polymorphisms, that is, genetic segregation of mtDNA in these strains appears to be as effective as in other published AMF genome. These mtDNAs harbor the typical set of 41 mitochondrial genes reported so far in AMF (Lee and Young 2009): two rRNAs, 14 protein-coding genes (PCGs) and 25 tRNAs. The PCGs include three ATP synthetase (*atp*), one cytochrome *b* (*cob*), three cytochrome *c* oxidase (*cox*), and seven NADH dehydrogenase (*nad*) genes (fig. 1). The *G. irregulare* isolate DAOM-234179 mtDNA has a genome size of 75,075 bp with a GC content of 36.7%, which is within range of the sizes of other sequences reported for this species (68,995–87,754 bp with a \approx 37% average GC content) (Lee and Young 2009; Formey et al. 2012; Nadimi et al. 2012). The only true outlying *G. irregulare* isolate with respect to its mitochondrial genome was reported by Formey et al. (2012) (i.e., MUCL43204) and has a size of 87,754 bp, whereas its conserved mitochondrial genes show only 65.8% identity (including introns and ORFs) with *G. irregulare* isolate 494. The *Glomus sp.* DAOM-240422 and *G. cerebriforme* mtDNAs were 86,170 and 59,633 bp, respectively, with GC contents of 36.7% and 46.7%. The latter is the smallest AMF mitochondrial genome yet reported, and is highly divergent

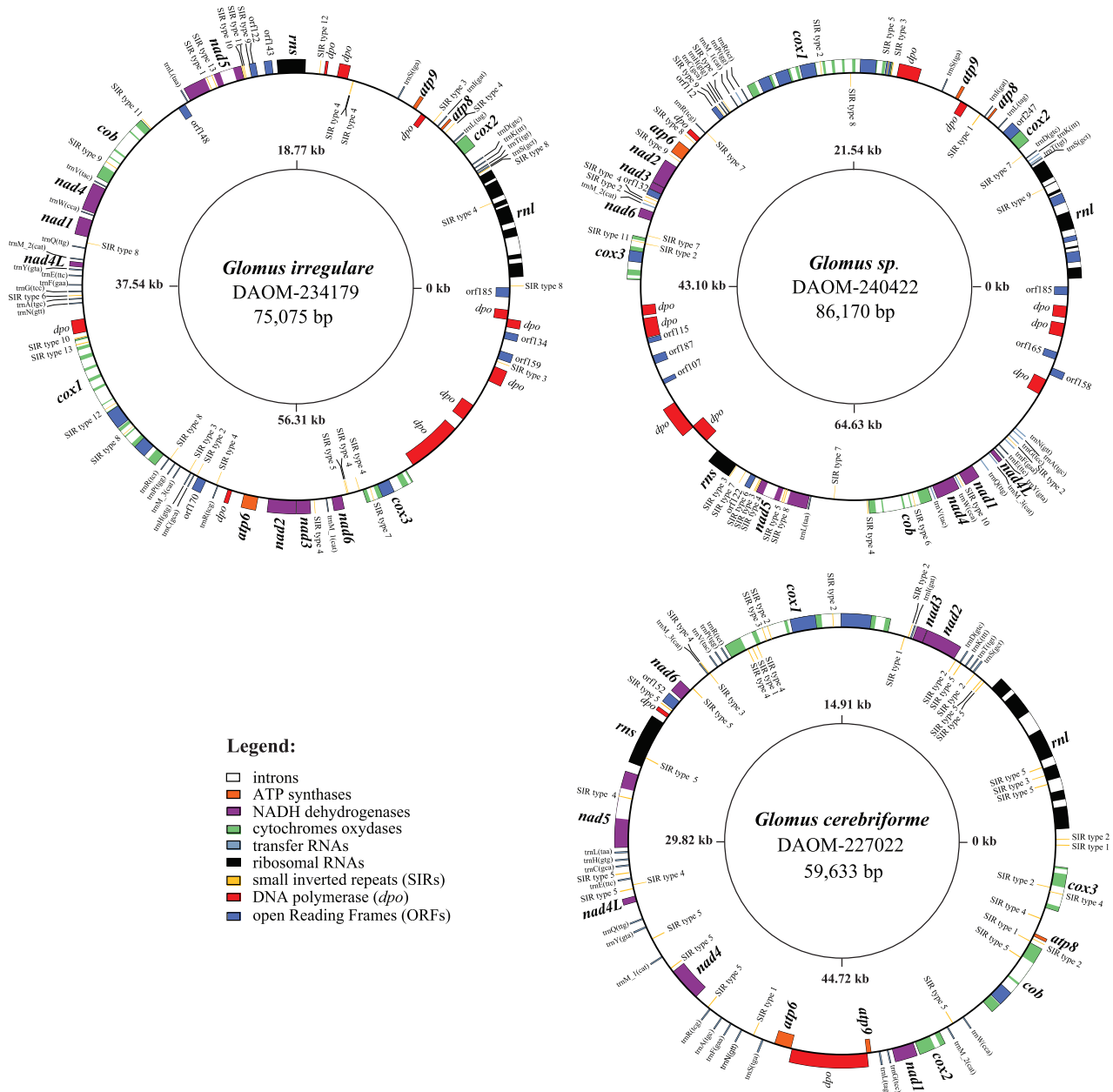


Fig. 1.—Comparison of *Glomus irregulare* DAOM-234179, *Glomus sp.* DAOM-240422, and *G. cerebriforme* mitochondrial genomes. The circular-mapping genomes were opened upstream of *rnl* to allow for easier comparisons. Genes on the outer and inner circumference are transcribed in a clockwise and counterclockwise direction, respectively. Gene and corresponding product names are *atp6*, 8, 9, ATP synthase subunit 6; *cob*, apocytochrome *b*; *cox1*–3, cytochrome *c* oxidase subunits; *nad1*–4, 4*L*, 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 followed by their anticodon. ORFs smaller than 100 amino acids are not shown.

when compared with the available mtDNAs of *Glomus* spp. and *Gigaspora* spp. The inflated GC% observed in *G. cerebriforme* is mostly due to reduced intergenic region size and the lack of mobile elements, which are mostly AT-rich regions, compared with the other strains. The difference in the mtDNA sequences observed between these taxa is caused by variation in intron content, *dpo* insertions and the

presence of ORFs (table 2). *Glomus cerebriforme* genes are encoded on both strands, their order is completely reshuffled compared with the other *Glomus* spp., and they are also phylogenetically distant from those of *G. irregulare* (supplementary fig. S1, Supplementary Material online). On the other hand, *Glomus sp.* DAOM-240422 showed high similarity to *G. irregulare* with regards to coding sequence identity, but

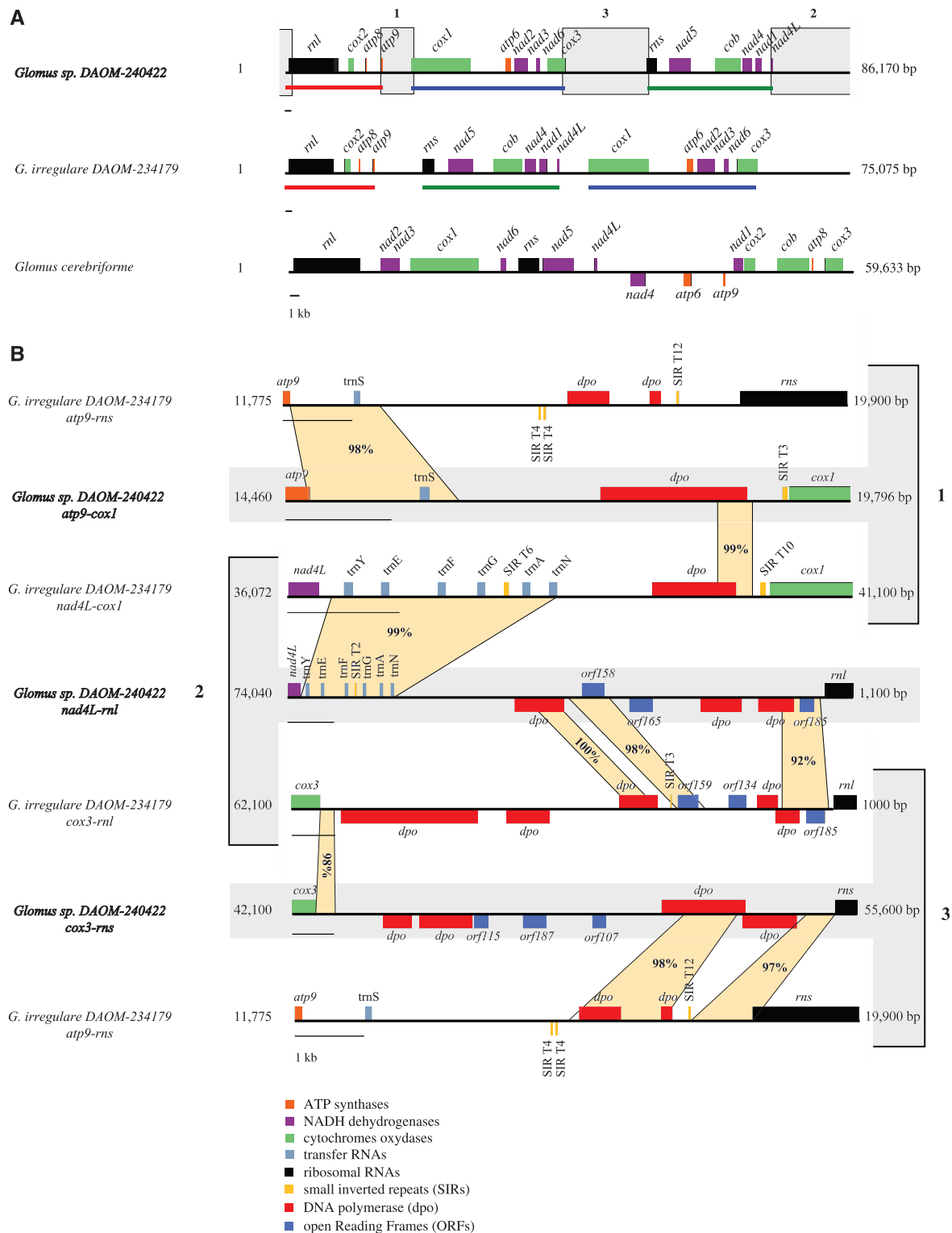


Fig. 2.—(A) Linear genome representation to compare the mitochondrial synteny between *Glomus irregulare* DAOM-234179, *Glomus sp.* DAOM-240422, and *G. cerebriforme*. The linear-mapping genomes were opened upstream of *rnl* to allow for easier comparisons. Corresponding gene clusters between *G. irregulare* DAOM-234179 and *Glomus sp.* DAOM-240422 are underlined in red, green, and blue. The newly formed intergenic regions of *Glomus sp.* DAOM-240422 are boxed in gray and tagged with a number. (B) Nucleotide identity comparison with BLASTn between the reshuffled *Glomus sp.* DAOM-240422 intergenic regions (*atp9-cox1* [box 1], *nad4L-rnl* [box 2], and *cox3-rns* [box 3]) with their putative homologous sequence in *G. irregulare* DAOM-234179. Homologous regions are indicated by the projections with their corresponding percent identity.

Table 2

AMF Mitochondrial Genome Features

AMF Strain	Genome Size (kb)	Introns ^a	ORFs ^b	<i>dpo</i>	SIRs ^c	Molecular Features
<i>G. irregulare</i> ^d FACE494	70,606	26	10	5	26	—
<i>G. irregulare</i> ^d DAOM197198	70,783	26	10	6	27	—
<i>G. irregulare</i> ^d MUCL46239	70,818	26	10	6	28	—
<i>G. irregulare</i> ^d MUCL46240	74,797	26	15	5	32	—
<i>G. irregulare</i> ^d MUCL46241	74,797	26	15	5	32	—
<i>G. irregulare</i> DAOM240415	70,781	26	9	8	?	—
<i>G. irregulare</i> DAOM234328	68,995	26	9	8	?	—
<i>G. irregulare</i> DAOM234179	75,075	26	10	10	31	—
<i>G. irregulare</i> ^d MUCL43204	87,754	32	23	4	50	—
<i>Glomus</i> sp. DAOM229456	87,763	32	19	4	?	Gene hybrids/HGT
<i>Glomus</i> sp. DAOM240422	86,170	25	19	10	29	Genome rearrangements
<i>G. cerebriforme</i>	59,633	16	4	2	40	Synteny shuffling
<i>Gigaspora margarita</i>	96,998	14	4	0	?	Transplicing/synteny shuffling
<i>Gig. rosea</i>	97,349	14	4	0	?	Transplicing/synteny shuffling

^aAll types including group I and II introns.^bORFs greater than 100 amino acids encoding unknown protein, LAGLIDADG and GIYYIG endonuclease are listed, including intronic ORFs but not *dpo* fragments.^cSIRs reported so far in AMF mitochondrial genomes.^dBased on data reported in Formey et al. (2012), but eroded ORFs were excluded.

surprisingly, the synteny of these two isolates differed substantially.

Plasmid-Related DNA Polymerase and Sequence Diversity of Small Inverted Repeats

We found 10 *dpo* insertions in *G. irregulare* DAOM-234179 and *Glomus* sp. DAOM-240422, making them the most *dpo*-rich AMF isolate reported so far, but only two copies were observed in *G. cerebriforme*. The three isolates harbor a huge bona fide *dpo* with ORF extensions, rendering a structure similar to the mitochondrial plasmid found in other fungi (Kim et al. 2000). However, we did not find TIRs, which are usually a characteristic of recent plasmid integration. We also found numerous SIRs in the three genomes. The 31 SIRs in *G. irregulare* DAOM-234179, 29 in *Glomus* sp. DAOM-240422, and 40 in *G. cerebriforme* were divided into 13, 11, and 5 distinct types, respectively, based on their hairpin secondary structure (supplementary figs. S3–S5, Supplementary Material online). The high number of SIR types, compared with the five previously described by Formey et al., result from the characterization of SIRs found in single copy, homologous to those previously described that could not have been detected otherwise. Some types are present in all *G. irregulare* isolates and closely related species such as *Glomus* sp. DAOM-240422, whereas others are endogenous to each strain, giving rise to a high SIRs sequence diversity. These elements were always present in intergenic regions, introns, or at the edge of endonuclease-encoding ORFs (supplementary tables S1–S3, Supplementary Material online), and suggest a close relationship with the *homing* mechanism for endonuclease integrations, as proposed by Formey et al. (2012). However, the

diversity of possible cleaving sites seems to be greater than previously expected with the here-reported novel mtDNA sequences.

Divergent Synteny in Two Novel Glomeraceae Species mtDNA

The *G. cerebriforme* mtDNA shows a completely different gene order compared with the other AMF mitochondrial genomes sequenced so far, with many type 5 SIR insertions in intergenic regions that could potentially be involved in synteny rearrangements. Nevertheless, as mentioned earlier, only two *dpo* insertions are present, in the *cox3-ml* and *nad6-ms* intergenic regions. As *G. cerebriforme* is phylogenetically distant from the publicly available mtDNAs of *Glomus* spp. and *Gigaspora* spp., no reference scaffold is available to confirm any gene order reorganization. In contrast, the comparison of *Glomus* sp. DAOM-240422 mitochondrial genome with its close relative *G. irregulare* DAOM-234179, revealed a reshuffling of two large gene clusters (fig. 2A). One gene cluster encompasses *cox1*, *atp6*, *nad2*, *nad3*, *nad6*, and *cox3*, whereas the other contains *ms*, *nad5*, *cob*, *nad4*, *nad1*, and *nad4L*. This rearrangement creates three novel intergenic regions in *Glomus* sp. DAOM-240422, which are *atp9-cox1*, *cox3-ms*, and *nad4L-ml*. This result was confirmed by long PCR performed with primers spanning those regions (supplementary fig. S2, Supplementary Material online). We then assessed the sequence identity between those intergenes and their corresponding regions in *G. irregulare* DAOM-234179 (fig. 2B). We compared the *Glomus* sp. DAOM-240422 *atp9-cox1* with *G. irregulare* DAOM-234179 *atp9-ms* and *nad4L-cox1* intergene; the *cox3-ms* with

G. irregulare DAOM-234179 *cox3-ml* and *atp9-rns* regions; and finally the *nad4L-ml* was compared with *nad4L-cox1* and *cox3-ml*. Each *Glomus sp.* DAOM-240422 intergenic region showed high sequence identity to its putative former counterpart in *G. irregulare* DAOM-234179, but covered only a small fraction of the intergene length, because many more ORFs (mostly *dpo* insertions) are present in *Glomus sp.* DAOM-240422. Local alignment using BLASTn, between the newly formed intergenic regions, did not reveal high similarity matches. We also constructed a sequence identity matrix of all the *dpo* sequences in its mtDNA and the highest match was only 53.1%.

Glomeromycota *dpo* Nucleotide Identity Network and Phylogenetic Analyses

To test whether the *dpo* sequences in *Glomus sp.* DAOM-240422 were orthologous with the sequences found in other AMF and to assess evolution of *dpo* within the Glomeromycota, we constructed a nucleotide similarity network of all *dpo* genes reported so far in AMF mitochondrial genomes (fig. 3A). In total, 12 similarity groups were formed and seven singletons were left aside. We observed that five *Glomus sp.* DAOM-240422 *dpo* sequences (black circles) clustered in distinct identity groups (G1, G2, G3, G6, and G8), whereas the rest were singletons inserted in reshuffled intergenic regions (*cox3-ml* and *nad4L-ml*), with low sequence similarity to each other or to known AMF *dpo* sequences. The *Gig. rosea* putative mitochondrial plasmid that was previously described (Nadimi et al. 2012) was also one of the singletons. The two *G. cerebriforme* *dpo* copies (light green circles) were the sole representatives of similarity group 12. Also, the mtDNA-integrated *dpo* showed a clustering tendency toward their intergenic region localization (fig. 3B). Interestingly, three *Glomus sp.* DAOM-240422 *dpo* sequences which are inserted in the reshuffled intergenic regions *atp9-cox1*, *cox3-rns*, and *nad4L-ml* (fig. 2) clustered, respectively, with the AMF *dpo* group 8 (*cox3-ml*), group 1 (*atp9-rns*), and group 2 (*cox3-ml* and the *Glomus fasciculatum atp9-rns_3* sequence).

To compare the identity network and the phylogenetic approach, and to assess the relationship of the *dpo* present within each respective mtDNA (i.e., paralogy vs. orthology), we constructed a phylogeny based on a conserved protein domain. The alignment of all 91 Glomeromycota *dpo* protein sequences pointed out to a 75 amino acid region that was common in 60 of the sequences (supplementary fig. S6A, Supplementary Material online). Interestingly, most of the *dpo* were truncated either upstream and/or downstream of that conserved domain. The resulting maximum likelihood phylogenetic tree showed robust bootstrap values (>75%) in the distal branches that support clades corresponding to the similarity groups observed in the gene network, whereas the basal branches displayed were weakly supported

(supplementary fig. S6B, Supplementary Material online). All similarity network groups were represented in the tree with strong statistical support in bootstrap and Bayesian inference values, except for the identity group 11 that was not represented, because the sequences did not harbor the conserved gene core. Surprisingly, this tree revealed only one occurrence of paralogy, in the group 12, for the two *dpo* sequences inserted in *G. cerebriforme* mtDNA (like it was also observed in the nucleotide network). Every other *dpo* phylogenetic clusters represent orthology groups. Finally, only one of the three *Glomus sp.* DAOM-240422 *dpo* sequences found in reshuffled intergenic regions was present in the phylogenetic analyses, in the group 8 cluster.

Glomus sp. DAOM-240422 *dpo* Recombination Analyses

Three *Glomus sp.* *dpo* genes present in the reshuffled intergenic regions clustered with *dpo* of other AMF taxa inserted in different mtDNA localization, rather than with paralogous mtDNA sequences (within DAOM-240422 isolate) in the nucleotide network. To easily screen for homologous recombination, we performed a distance-based recombination analyses in which these *Glomus sp.* *dpo* sequences were compared with their three closest relatives in the nucleotide identity network. For the *Glomus sp. atp9-cox1* and *cox3-rns* (supplementary fig. S7A and B, Supplementary Material online) comparison, we observed a disruption in the identity of the *dpo* sequences going from 70% to 80% identity and spiking abruptly to a near perfect 100% nucleotide identity. When compared together, the *Glomus sp. nad4L-ml_1 dpo* sequence started off with less than 80% identity to the *G. fasciculatum atp9-rns_3 dpo* and then rose sharply to almost 95% identity. Conversely, the *G. irregulare* MUCL43204 *cox3-ml_1 dpo* started at more than 99% similar to the *Glomus sp. nad4L-ml_1 dpo* sequence, but then dropped to less than 80% identity (supplementary fig. S7C, Supplementary Material online). These *dpo* sequences showed typical signatures of homologous recombinants regarding their sequence identity, highly reminiscent to what was found for a recombinant HIV-1 strain (Liitsola et al. 2000).

To further confirm these findings, we performed a statistical phylogenetic recombination analysis using the HMM. The *Glomus sp.* DAOM-240422 *atp9-cox1* showed a perfect Bayesian inference value of 1.0 for each three topologies in a given portion of the alignment (supplementary fig. S7A, Supplementary Material online), which allowed the finding of putative recombination break points (between two given topologies where one reach a Bayesian inference value >0.95 and the other <0.05) at nucleotide position 357 and 613 bp of the sequence, respectively. The *Glomus sp.* DAOM-240422 *nad4L-ml* also showed a perfect Bayesian inference value of 1.0, but for only two topologies (I and II), thus allowing for the determination of one putative break point at position 161 bp

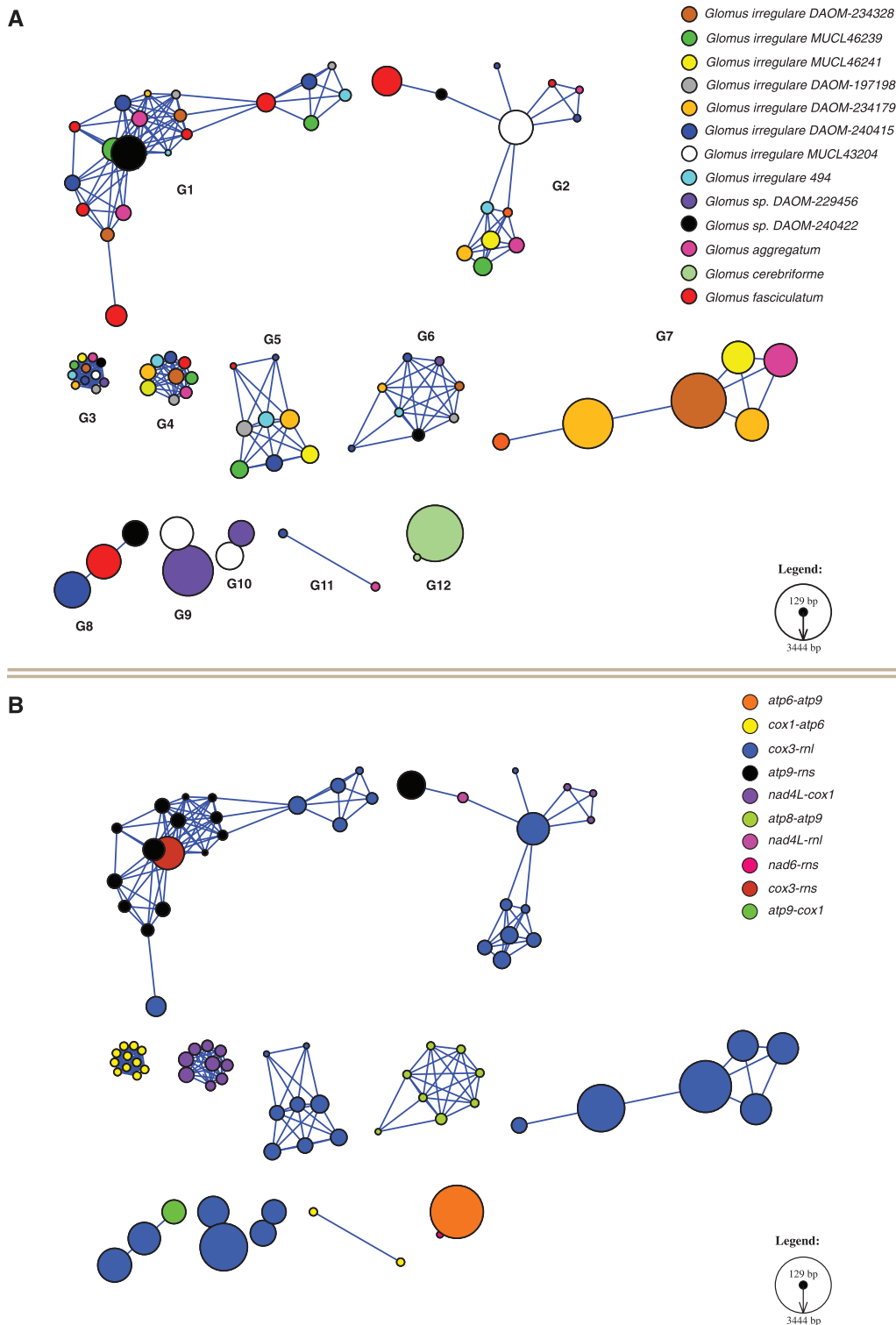


Fig. 3.—Similarity network of *dpo* sequences inserted in AMF mitochondrial genomes. (A) Each node represents a *dpo* insert colored by mitochondrial genome. The node size is proportional to the sequence length. Two nodes are connected by an edge when they share significant similarity (reciprocal best BLAST hit with a minimum of $1E-10$ score, and 30% minimum identity covering at least 30% of the smallest sequence). The layout was produced by Cytoscape, using an edge-weighted force-directed model, which brings closer sequences sharing more similarity. Twelve homology groups were formed (labeled G1–G12). There are 97 nodes on that network with 336 edges. (B) Same network but nodes are colored by their mitochondrial intergenic region localization. The legend shows a linear size relationship between the smallest *dpo* sequence (*Glomus irregulare* DAOM-240415_cox3-ntl_4, 129 bp) and the largest one (*G. cerebriforme atp6-atp9*, 3444 bp).

(supplementary fig. 7C, Supplementary Material online). Conversely, the *Glomus sp.* DAOM-240422 *cox3-ms* analyses did not provide statistical support for any given topologies tested with the HMM phylogenetic method (supplementary fig. 7B, Supplementary Material online).

Global Protein Similarity Network with All Known DNA Polymerase Genes in the Public Database

To assess a possible origin and transmission mode of AMF mitochondrial *dpo* sequences, we constructed a protein similarity network using a broad spectrum of *dpo* sequences. Interestingly, most of the glomeromycotan plasmid-related DNA polymerase proteins, 60 in total, were found clustered with 35 other fungi, 28 plants (including two red algae of the genus *Porphyra*), five other eukaryotes (*Physarum polycephalum*, *Ochromonas danica*, two *Placozoa sp.* and a *Cnidaria* species), four bacteria (*Bacillus thuringiensis*, two *Bacillus cereus*, and *Brevibacillus sp.*), and three bacteriophages (fig. 4A). All *dpo* sequences are present within their respective mtDNA or are located in a mitochondrial plasmid. The patchy distribution observed among phylogenetically distant taxa such as plants and fungi is similar to that of a *cox1* intron that invaded plant mitochondria and was thought to have been acquired horizontally from a fungal donor (Cho et al. 1998). To assess the evolutionary relationship of AMF DPO proteins with the other taxa present in the cluster, we selected the group gathering the most connected nodes, that is, the sequences sharing the most similarity all together. The corresponding sequences (including plant, Glomeromycotan and other fungal sequences), were used to perform a maximum likelihood phylogenetic analyses (fig. 4B). We used the viral and bacterial proteins present in the sub-network to root the tree, because they present a distal position in the network. The resulting tree showed robust support (100% bootstrap and 1.0 Bayesian inference value) for the viral-bacterial group, and surprisingly, the plant DPO proteins clustered with those of the Glomeromycota as a sister group, supported by a 76% bootstrap and 1.0 Bayesian inference value. Except for the other fungi group clade, the phylogeny was robustly supported either for the basal and the distal branches.

Discussion

Interhaplotype Homologous Recombination in *Glomus sp.* DAOM-240422

The numerous type 5 SIRs present in *G. cerebriforme* intergenic regions may have had the potential to trigger intra-molecular gene reshuffling. The two *G. cerebriforme* *dpo* sequences were shown to be paralogs, with a nucleotide identity of 81%. However, we cannot be sure that intra-molecular recombination occurred in *G. cerebriforme*, because we do not have closely related taxa as comparators. In the *Glomus sp.* DAOM-240422 mitochondrial genome, the two most

similar *dpo* sequences share a low nucleotide identity percentage of 53.1%. However, we observe a near perfect sequence identity for PCGs and some intergenic regions with close relative *G. irregulare*. Taken together, these observations suggest that the rearrangements observed in *Glomus sp.* DAOM-240422 are recent and do not result from intramolecular recombination from an ancient *dpo* duplication. Further, no recent *dpo* duplications were observed in any given AMF strains surveyed. The *Glomus sp.* DAOM-240422 *dpo* sequences that are inserted in the reshuffled intergenic regions clustered in the nucleotide network with *dpo* sequences of other AMF species that were inserted in the putatively former intergenic arrangement. Comparison of the *Glomus sp.* DAOM-240422 *atp9-cox1* and *nad4L-rrl_1 dpo*, with their three closest relatives showed a typical signature of homologous recombination (supplementary fig. S7A and C, Supplementary Material online), with genes sharing a high sequence identity with different portions of the alignment, where intersections represent putative recombination break points. These recombination events were strongly statistically supported with the HMM method. Even if the *Glomus sp.* 240422 *cox3-ms dpo* recombination was not supported by HMM, it still clustered with *dpo* present in other intergenic regions, and the lack of statistical support might be explained by the absence of its recombination partner in the cluster. These observations strongly suggest that the mitochondrial gene cluster reorganization in *Glomus sp.* DAOM-240422 is the result of *dpo*-mediated homologous recombination between different mitochondrial haplotypes harboring highly similar *dpo* orthologs present in distinct genome localization.

Two types of mobile elements reported in AMF mitochondrial genomes have been previously shown to be involved in genome rearrangements through homologous recombination in other organisms: the plasmid-related *dpo* genes in Basidiomycota (Ferandon et al. 2008) and SIRs in bacteria and plants (Schofield et al. 1992; Bi and Liu 1996; Tanaka et al. 2012). Furthermore, the fact that the number of *dpo* sequences varies in different intergenic regions among isolates of the same species and closely related ones (Formey et al. 2012; Beaudet et al. 2013), combined with our findings suggest that 1) there is no evidence of intra-molecular recombination in *Glomus sp.* DAOM-240422; 2) recombination between distinct mtDNA haplotypes harboring different *dpo* insertions could occur; 3) as there is no intra-isolate mtDNA polymorphism, the reshuffled mtDNA haplotype in *Glomus sp.* DAOM-240422 is fixed in the population; and 4) those rearrangements are triggered by *dpo* insertions through homologous recombination.

Such recombination events could take place following hyphal fusion (i.e., anastomosis). This last feature not only facilitates the distribution of nutrients and signaling molecules through the entire AMF mycelium, but also plays a key role in genetic exchange (Giovannetti et al. 1999) and segregation (Angelard and Sanders 2011). Those anastomoses can occur

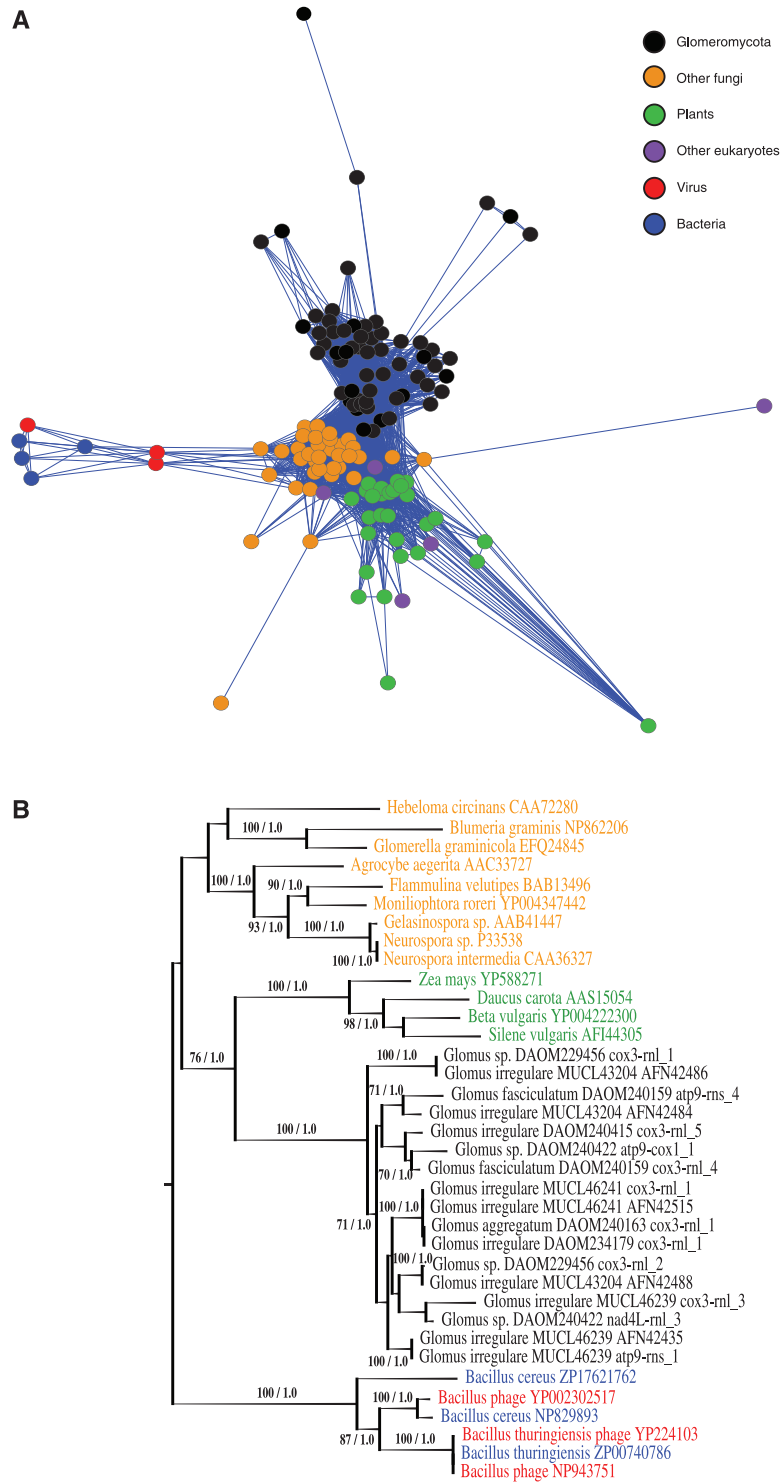


FIG. 4.—(A) Sub-cluster of the global network of shared amino acid similarities between the Glomeromycota DPO proteins and all homologous sequences found on GenBank. Building this data set was performed using translated Glomeromycota sequences as queries for a BLASTp search on the GenBank nr database (minimal e-value threshold of $1E-40$, 20% minimal similarity covering at least 20% of the smallest sequence). The network layouts were further produced by Cytoscape software, using an edge-weighted force-directed model, meaning that genes sharing more protein similarity appear closer in the display. There are 135 nodes in that network with 2,520 edges. (B) DPO protein maximum likelihood tree obtained with the rtREV + CAT phylogenetic model. The bacterial/viral cluster was used to root the tree. Number on branches indicates bootstrap support values (<60% cut-off) and Bayesian inference values, respectively. The Bayesian analyses gave a similar topology (data not shown). The tree includes AMF (black), other fungi (orange), plants (green), virus (red), and bacteria (blue).

either between genetically different AMF isolates originating from the same experimental field (Croll et al. 2009; Angelard et al. 2010) or between strains isolated from distant locations (e.g., from the same [Purin and Morton 2012], or even different continents [de la Providencia et al. 2013]). Both of these studies suggested that mitochondria might also be exchanged via anastomosis, thus creating a heteroplasmic state.

Evidence of Heteroplasmy Challenges the Concept of an AMF Individual

All published AMF mitochondrial genomes are homoplasmic, that is, homogeneous within isolates. In other fungi, mechanisms contributing to homoplasmy are often related to the recombination of parental mtDNAs or by the selection of one of the parental haplotypes due to the presence of segregation mechanisms (reviewed by Birky [2001]). In the absence of external factors, heteroplasmy should be the default state for mtDNA under a simple mutation-drift scenario (White et al. 2008). In yeasts, the presence of a heteroplasmic state was shown to be transient due to the presence of segregation mechanisms, such as nucleoid formation, and mtDNA molecular repair pathways which maintain mitochondrial genome integrity (Zimmer et al. 1991; Hu et al. 1995; Yasuhira and Yasui 2000). The presence of nucleoids, similar mtDNAs linked by Holliday junctions, induces a genetic bottleneck responsible for a fast segregation (Lockshon et al. 1995; White and Lilley 1997; MacAlpine et al. 1998). The mitochondrial genome homogeneity observed for in vitro and in vivo AMF cultures could be the result of such effective mtDNA segregation and repair mechanisms that would take place during the subcultivation processes.

The AMF colonies are characterized by the interconnectedness of its different parts by means of hyphal fusion (i.e., anastomosis) (Giovannetti et al. 1999; de la Providencia et al. 2005; Voets et al. 2006; Purin and Morton 2012), which can lead to genetic exchange, even between genetically distinct individuals. This hyphal fusion dynamic in natural populations can also prevent the loss of genetic diversity (Bever and Wang 2005). Recently, anastomoses between geographically distant *G. irregulare* isolates cultivated in vitro were shown to induce transient mitochondrial length heteroplasmy (with variation in *dpo* insertions between the two mitochondrial haplotypes) (de la Providencia et al. 2013). Although in situ heteroplasmy have never been demonstrated, these observations, in addition to the interhaplotype *dpo* homologous recombination evidences presented here, suggest that transient heteroplasmy could occur in the field.

The coexistence of numerous AMF mitochondrial haplotypes in the same cytoplasm and the occurrence of homologous mitochondrial recombination might be commonplace in nature, as it was shown in natural populations of the basidiomycete fungus *Armillaria gallica* (Saville et al. 1998). The heteroplasmy stability in natural populations, may provide

useful additional information for defining haplotypes, and resolving further the relationships among individuals at a population level (White et al. 2008). Given the increasing use of mitochondrial markers in AMF research, the demonstration that genetically distinct mtDNAs can be exchanged through anastomosis, coexist in a common cytoplasm and give rise to a recombinant haplotype, should lead to a thorough discussion on the AMF individuality concept. If heteroplasmy is confirmed in natural populations, the use of mtDNA as a criterion to define a reliable AMF taxonomic unit would certainly facilitate our understanding of the processes occurring in natural populations, although caution should still be applied because it would be artificial, and would not reflect the genotypic diversity of AMF nuclear DNA, because mitochondrial and nuclear DNA segregation are mostly stochastic mechanisms.

Vertical Versus Horizontal Inheritance of the AMF Mitochondrial *dpo*

Given the high genetic diversity of *dpo* and their variable truncation patterns, a classical phylogenetic approach requiring multiple alignment of highly similar sequences, is not suitable to analyze their overall evolutionary relationships, especially for recombinant ones (Baptiste et al. 2012). Indeed, the classical approach only allowed us to analyze 60 of 91 Glomeromycota *dpo* sequences, on a limited 75 amino acids conserved domain, leading to a tree with poorly supported basal branches (supplementary fig. 6B, Supplementary Material online). As an alternative, gene similarity network analyses have already been proven to be useful in the study of highly diverse gene families and introgressive descent events, such as recombination and HGT (Baptiste et al. 2012).

The AMF *dpo* nucleotide network highlights the existence of a clustering tendency for *dpo* sequences located in the same intergenic regions (fig. 3B). The high similarity (short edges) observed between the sequences of each group and the phylogenetic topology (supplementary fig. S6B, Supplementary Material online) suggest that they are orthologs that could have recently been acquired vertically from a common ancestor. This is not surprising for the *G. irregulare* isolates, but implies that the other species that shared those sequences have also diverged recently (i.e., *G. fasciculatum* and *G. aggregatum*) given the high within-group sequence similarity. However, the seven singletons left aside by the network nucleotide similarity clustering (five in the closely related species *Glomus sp.* DAOM-240422) suggest that numerous independent plasmid-mediated *dpo* insertions, from unknown origin, and/or followed by a fast divergence might also have occurred. The putative *Gig. rosea* plasmid had low sequence similarity with the other *dpo* insertions in AMF mtDNA and no extra chromosomal *dpo* were found within any of the surveyed AMF genomes. This *dpo* sequence diversity could indicate that AMF mitochondria harbor a

broad range of plasmid types, as described in *Neurospora* species, where seven circular and four linear homology groups of mitochondrial plasmids have been reported (Schulte and Lambowitz 1991; Yang and Griffiths 1993a, 1993b; Arganoza et al. 1994; Marcinko-Kuehn et al. 1994).

It has been found that mitochondrial plasmids in plants could be horizontally acquired through pollen (Handa 2007). HGT has also been hypothesized to be an important factor in the dispersal of mitochondrial plasmids between fungal strains and species. That assumption is based on 1) discordant phylogenetic relationships between plasmids and host genomes (Kempken et al. 1992), 2) the geographic distribution of plasmids within and among host species (Arganoza et al. 1994), and 3) direct experimental demonstration of HGT (Collins and Saville 1990; Griffiths et al. 1990; Debets et al. 1994). Until now, no direct experimental evidence of HGT of *dpo* sequences was provided, and no mitochondrial plasmid has yet been found in AMF (except a putative linear plasmid in *Gig. rosea*). However, the distribution of the plasmid-related *dpo* sequences is uneven between isolates of the same AMF species (Formey et al. 2012) and we showed that there is a large diversity of *dpo* sequences within AMF (12 orthology groups, along with seven singletons). Another intriguing observation is that the *dpo* sequences are also present in all AMF family surveyed including the Glomeraceae, Gigasporaceae, Ambisporaceae, Diversisporaceae, and Archaeosporaceae (data not shown, based on a preliminary 454 assembly). A mechanism known as post-fusion incompatibility might be the most plausible explanation for the ubiquitous distribution of AMF mtDNA *dpo*. This mechanism implies a brief hyphal fusion between incompatible AMF species and the formation of septa. This process has already been shown to allow exchange of nuclear genetic material (Croll et al. 2009), and also offers an explanation for the mobile endonuclease-mediated HGT, leading to the formation of gene hybrids in the mitochondrial genome of *Glomus diaphanum*-like species (Beaudet et al. 2013). Keeping in perspective that divergent AMF taxa can simultaneously colonize roots of the same plant species, it is likely that a mitochondrial plasmid could spread rapidly through the Glomeromycota phylum even between phylogenetically distant taxa.

Glomeromycota *dpo* Are Closely Related to Plant Mitochondrial Plasmids

More than 50 linear plasmids have been reported in the mitochondria of 20 fungal species, in contrast to the 14 found in eight plant species. These have only been observed twice in animal mtDNA, in the placozoan *Trichoplax adhaerens* and in the moon jelly *Aurelia aurita* (Shao et al. 2006; Signorovitch et al. 2007). This uneven distribution of mitochondrial plasmids within the Eukaryota raises questions about their origin and their mode of transmission.

The structure of the plant mitochondrial plasmids consists of an invertron structure, long inverted repeats with proteins covalently bound to their 5'-ends. This structure is remarkably similar to fungal plasmids and to some DNA viruses (Rosewich and Kistler 2000). Additional evidence, such as low GC content in plant plasmids (30% in the *Brassica* 11.6 kb plasmid, 38.9% in the 10.4 kb sugar beet plasmid, and 37–39% in the maize S plasmids) compared with their respective mtDNA (*Brassica napus* 45.2% [Handa 2003], *Beta vulgaris* 43.9% [Kubo et al. 2000]) and the maize 44.0% (GenBank accession DQ490951), supports HGT from a fungal donor as the source of linear mitochondrial plasmids in plants, because low GC content is typically found in fungal mtDNA (Handa 2008) (with an average of 37% in *G. irregulare* isolates [Handa 2008; Formey et al. 2012]). Moreover, our phylogenetic analyses robustly support the DPO proteins of plants and Glomeromycota as a sister group. Taken together, these evidences suggest that AMF could be potential donors of plasmid-related *dpo* to plants. However, species of Glomeromycota colonize the same ecological niche and often co-occur with many pathogenic or endophytic fungi in plant roots, and one of these co-occurring fungi could be the common donor to both plants and AMF, also plant plasmids could have different fungal origins. If this is the case, the organism in question has yet to be sequenced.

To further study the AMF mitochondrial plasmid-related *dpo* genes, the mitochondrial genome of *Geosiphon pyriformis*, an ancestral member of Glomeromycota that lives in endocytobiotic association with the cyanobacterium *Nostoc punctiforme* (Gehrig et al. 1996), could provide insights into the origin and evolution of those elements in that phylum. Also, some DNA polymerase genes were present in the mtDNA of the ascomycete fungal partner of the lichen *Peltigera malacea* (Xavier et al. 2012). Those sequences clustered within the other fungi group in the global DPO protein network. Another interesting avenue would be to assess the horizontal transmission of mitochondrial plasmids in different fungi-plant symbiotic systems.

Conclusions and Outlook

The mitochondrial genome comparison of the three *Glomus* species showed the extent of mtDNA plasticity with regards to their synteny. It also offers a basis to develop molecular tool kits aimed specifically at identifying and quantifying taxa of the ecologically and agriculturally important Glomeraceae family. However, finding evidence of interhaplotype homologous recombination suggests the occurrence of mtDNA heteroplasmy in natural populations. Although this has yet to be demonstrated, it raises questions about the artificiality of single spore cultivations of these fungi, which probably induces an underestimation of the mtDNA diversity in situ. Approaches like genotyping-by-sequencing and ecotilling could be useful to assess this particular hypothesis.

Mitochondrial genome comparisons can also provide insights into molecular processes related to the biology and community structure of AMF. The close phylogenetic relationship between the plasmid-related *dpo* of plants and Glomeromycota suggests that HGT events could have occurred between plants and their fungal symbionts. The intimate relationship between plants and AMF in nature may unravel other transfer events that have taken place through evolution. Also, the role of viruses and bacteria and their interactions with fungi and plants represents an interesting research avenue to generate knowledge about the evolution and origin of mitochondrial plasmids in those groups.

Supplementary Material

Supplementary tables S1–S3 and figures S1–S7 are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

Acknowledgments

The authors thank Biopterre centre du développement des bioproduits and CRBM. They thank Dr B.F. Lang for bioinformatics assistance and access to an automated organelle genome annotation software, Dr Simon Joly for recommendations on the phylogenetic analyses, Stéphanie Berthiaume for her assistance in long PCRs, and Dr Terrence Bell and Karen Fisher Favret for English editing and comments on the manuscript. They also thank two anonymous reviewers for helpful comments. This work is a part of a research project organized and coordinated by Premier Tech. This work was supported by the NSERC Cooperative Research and Development grant (RDGPJ 395241-09), Premier Tech, and CRIBIQ.

Literature Cited

- Akins RA, et al. 1988. Nucleotide sequence of the Varkud mitochondrial plasmid of *Neurospora* and synthesis of a hybrid transcript with a 5' leader derived from mitochondrial RNA. *J Mol Biol.* 204:1–25.
- Akins RA, Kelley RL, Lambowitz AM. 1986. Mitochondrial plasmids of *Neurospora*: integration into mitochondrial DNA and evidence for reverse transcription in mitochondria. *Cell* 47:505–516.
- Almasan A, Mishra NC. 1991. Recombination by sequence repeats with formation of suppressive or residual mitochondrial DNA in *Neurospora*. *Proc Natl Acad Sci U S A.* 88:7684–7688.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403–410.
- André C, Levy A, Walbot V. 1992. Small repeated sequences and the structure of plant mitochondrial genomes. *Trends Genet.* 8: 128–132.
- Angelard C, Colard A, Niculita-Hirzel H, Croll D, Sanders IR. 2010. Segregation in a mycorrhizal fungus alters rice growth and symbiosis-specific gene transcription. *Curr Biol.* 20:1216–1221.
- Angelard C, Sanders IR. 2011. Effect of segregation and genetic exchange on arbuscular mycorrhizal fungi in colonization of roots. *New Phytol.* 189:652–657.
- Arganoza MT, Min J, Hu Z, Akins RA. 1994. Distribution of seven homolog groups of mitochondrial plasmids in *Neurospora*: evidence for widespread mobility between species in nature. *Curr Genet.* 26: 62–73.
- Bader G, Hogue C. 2003. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4:2.
- Ballard JWO, Whitlock MC. 2004. The incomplete natural history of mitochondria. *Mol Ecol.* 13:729–744.
- Baptiste E, et al. 2012. Evolutionary analyses of non-genealogical bonds produced by introgressive descent. *Proc Natl Acad Sci U S A.* 109: 18266–18272.
- Barr CM, Neiman M, Taylor DR. 2005. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. *New Phytol.* 168: 39–50.
- Beaudet D, Nadimi M, Iffis B, Hijri M. 2013. Rapid mitochondrial genome evolution through invasion of mobile elements in two closely related species of arbuscular mycorrhizal fungi. *PLoS One* 8:e60768.
- Bécard G, Fortin JA. 1988. Early events of vesicular-arbuscular mycorrhiza formation on Ri T-DNA transformed roots. *New Phytol.* 108:211–218.
- Bever JD, Wang M. 2005. Arbuscular mycorrhizal fungi: hyphal fusion and multigenomic structure. *Nature* 433:E3–E4.
- Bi X, Liu LF. 1996. DNA rearrangement mediated by inverted repeats. *Proc Natl Acad Sci U S A.* 93:819–823.
- Birky CW. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu Rev Genet.* 35: 125–148.
- Birky CW, Strausberg RL, Forster JL, Perlman PS. 1978. Vegetative segregation of mitochondria in yeast: estimating parameters using a random model. *Mol Gen Genet.* 158:251–261.
- Boer PH, Gray MW. 1991. Short dispersed repeats localized in spacer regions of *Chlamydomonas reinhardtii* mitochondrial DNA. *Curr Genet.* 19:309–312.
- 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.
- Börstler B, Raab PA, Thiéry 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 Phytol.* 180: 452–465.
- Boursot P, Yonekawa H, Bonhomme F. 1987. Heteroplasmy in mice with deletion of a large coding region of mitochondrial DNA. *Mol Biol Evol.* 4:46–55.
- Burger G, Gray MW, Forget L, Lang BF. 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout Jakobid protists. *Genome Biol Evol.* 5:418–438.
- Cho Y, Qiu Y-L, Kuhlman P, Palmer JD. 1998. Explosive invasion of plant mitochondria by a group I intron. *Proc Natl Acad Sci U S A.* 95: 14244–14249.
- Collins RA, Saville BJ. 1990. Independent transfer of mitochondrial chromosomes and plasmids during unstable vegetative fusion in *Neurospora*. *Nature* 345:177–179.
- Corradi N, Bonen L. 2012. Mitochondrial genome invaders: an unselfish role as molecular markers. *New Phytol.* 196:963–965.
- Court DA, Bertrand H. 1992. Genetic organization and structural features of maranhar; a senescence-inducing linear mitochondrial plasmid of *Neurospora crassa*. *Curr Genet.* 22:385–397.
- Croll D, et al. 2009. Nonself vegetative fusion and genetic exchange in the arbuscular mycorrhizal fungus *Glomus intraradices*. *New Phytol.* 181: 924–937.
- de la Providencia IE, Nadimi M, Beaudet D, Rodriguez Morales G, Hijri M. 2013. Detection of a transient mitochondrial DNA heteroplasmy in the progeny of crossed genetically divergent isolates of arbuscular mycorrhizal fungi. *New Phytol.* Advance Access published June 24, 2013, doi: 10.1111/nph.12372.

- de Zamaroczy M, Bernardi G. 1986. The GC clusters of the mitochondrial genome of yeast and their evolutionary origin. *Gene* 41:1–22.
- Debets F, Yang X, Griffiths AF. 1994. Vegetative incompatibility in *Neurospora*: its effect on horizontal transfer of mitochondrial plasmids and senescence in natural populations. *Curr Genet*. 26: 113–119.
- Denovan-Wright EM, Lee RW. 1994. Comparative structure and genomic organization of the discontinuous mitochondrial ribosomal RNA genes of *Chlamydomonas eugametos* and *Chlamydomonas reinhardtii*. *J Mol Biol*. 241:298–311.
- Dimmic MW, Rest JS, Mindell DP, Goldstein RA. 2002. rtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny. *J Mol Evol*. 55:65–73.
- Doner L, Becard G. 1991. Solubilization of gellan gels by chelation of cations. *Biotechnol Tech*. 5:25–28.
- Etherington GJ, Dicks J, Roberts IN. 2005. Recombination analysis tool (RAT): a program for the high-throughput detection of recombination. *Bioinformatics* 21:278–281.
- Evelin H, Kapoor R, Giri B. 2009. Arbuscular mycorrhizal fungi in alleviation of salt stress: a review. *Ann Bot*. 104:1263–1280.
- Ferandon C, Chatel SEK, Castandet B, Castroviejo M, Barroso G. 2008. The *Agrocybe aegerita* mitochondrial genome contains two inverted repeats of the *nad4* gene arisen by duplication on both sides of a linear plasmid integration site. *Fungal Genet Biol*. 45:292–301.
- Formey D, et al. 2012. Comparative analysis of mitochondrial genomes of *Rhizophagus irregularis*—syn. *Glomus irregulare*—reveals a polymorphism induced by variability generating elements. *New Phytol*. 196: 1217–1227.
- Gehrig H, Schübler A, Kluge M. 1996. *Geosiphon pyriforme*, a fungus forming endocytobiosis with *Nostoc* (Cyanobacteria), is an ancestral member of the glomales: evidence by SSU rRNA analysis. *J Mol Evol*. 43:71–81.
- Gianinazzi S, et al. 2010. Agroecology: the key role of arbuscular mycorrhizas in ecosystem services. *Mycorrhiza* 20:519–530.
- Giovannetti M, Azzolini D, Citeresi AS. 1999. Anastomosis formation and nuclear and protoplasmic exchange in arbuscular mycorrhizal fungi. *Appl Environ Microbiol*. 65:5571–5575.
- Griffiths AF, et al. 1990. Heterokaryotic transmission of senescence plasmid DNA in *Neurospora*. *Curr Genet*. 17:139–145.
- Griffiths AJ. 1995. Natural plasmids of filamentous fungi. *Microbiol Rev*. 59:673–685.
- Griffiths AJF. 1992. Fungal senescence. *Annu Rev Genet*. 26:351–372.
- Grindley NDF, Reed RR. 1985. Transpositional recombination in prokaryotes. *Annu Rev Biochem*. 54:863–896.
- Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. 2008. The Vienna RNAsuite A. *Nucleic Acids Res*. 36:W70–W74.
- Halary S, McInerney JO, Lopez P, Baptiste E. 2013. EGN: a wizard for construction of gene and genome similarity networks. *BMC Evol Biol*. 13:146.
- Handa H. 2003. The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res*. 31:5907–5916.
- Handa H. 2007. Investigation of the origin and transmission of linear mitochondrial plasmid based on phylogenetic analysis in Japanese rapeseed varieties. *Genome* 50:234–240.
- Handa H. 2008. Linear plasmids in plant mitochondria: peaceful coexistences or malicious invasions? *Mitochondrion* 8:15–25.
- Hänfler J, et al. 1992. Circular extrachromosomal DNA codes for a surface protein in the (+) mating type of the zygomycete *Absidia glauca*. *Curr Genet*. 22:319–325.
- Hermanns J, Asseburg A, Osiewacz HD. 1994. Evidence for a life span-prolonging effect of a linear plasmid in a longevity mutant of *Podospora anserina*. *Mol Gen Genet*. 243:297–307.
- Hu J, Vanderstraeten S, Foury F. 1995. Isolation and characterization of ten mutator alleles of the mitochondrial DNA polymerase-encoding MIP1 gene from *Saccharomyces cerevisiae*. *Gene* 160:105–110.
- Husmeier D, Wright F. 2001. Detection of recombination in DNA multiple alignments with hidden Markov models. *J Comput Biol*. 8:401–427.
- Ismail Y, Hijri M. 2012. Arbuscular mycorrhisation with *Glomus irregulare* induces expression of potato PR homologues genes in response to infection by *Fusarium sambucinum*. *Funct Plant Biol*. 39:236–245.
- Kempken F, Hermanns J, Osiewacz HD. 1992. Evolution of linear plasmids. *J Mol Evol*. 35:502–513.
- Kim E-K, Jeong J-H, Youn HS, Koo YB, Roe J-H. 2000. The terminal protein of a linear mitochondrial plasmid is encoded in the N-terminus of the DNA polymerase gene in white-rot fungus *Pleurotus ostreatus*. *Curr Genet*. 38:283–290.
- Kubo T, et al. 2000. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNACys(GCA). *Nucleic Acids Res*. 28:2571–2576.
- Kuhn G, Hijri M, Sanders IR. 2001. Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi. *Nature* 414:745–748.
- 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.
- Li Q, Nargang FE. 1993. Two *Neurospora* mitochondrial plasmids encode DNA polymerases containing motifs characteristic of family B DNA polymerases but lack the sequence Asp-Thr-Asp. *Proc Natl Acad Sci U S A*. 90:4299–4303.
- Liitsola K, et al. 2000. An AB recombinant and its parental HIV type 1 strains in the area of the former Soviet Union: low requirements for sequence identity in recombination. *AIDS Res Hum Retroviruses*. 16: 1047–1053.
- Lockshon D, et al. 1995. A role for recombination junctions in the segregation of mitochondrial DNA in yeast. *Cell* 81:947–955.
- Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet*. 52:267–274.
- MacAlpine DM, Perlman PS, Butow RA. 1998. The high mobility group protein Abf2p influences the level of yeast mitochondrial DNA recombination intermediates in vivo. *Proc Natl Acad Sci U S A*. 95: 6739–6743.
- Marcinko-Kuehn M, Yang X, Debets F, Jacobson DJ, Griffiths AJF. 1994. A kailo-like linear plasmid in Louisiana field isolates of the pseudohomothallic fungus *Neurospora tetrasperma*. *Curr Genet*. 26: 336–343.
- Mathimaran N, et al. 2008. Microsatellites for disentangling underground networks: Strain-specific identification of *Glomus intraradices*, an arbuscular mycorrhizal fungus. *Fungal Genet Biol*. 45:812–817.
- Milne I, et al. 2004. TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics* 20:1806–1807.
- Nadimi M, Beaudet D, Forget L, Hijri M, Lang BF. 2012. Group I intron-mediated trans-splicing in mitochondria of *Gigaspora rosea*, and a robust phylogenetic affiliation of arbuscular mycorrhizal fungi with Mortierellales. *Mol Biol Evol*. 29:2199–2210.
- Nakazono M, Kanno A, Tsutsumi N, Hirai A. 1994. Palindromic repeated sequences (PRSS) in the mitochondrial genome of rice: evidence for their insertion after divergence of the genus *Oryza* from the other Gramineae. *Plant Mol Biol*. 24:273–281.
- Nargang FE, Bell JB, Stohl LL, Lambowitz AM. 1984. The DNA sequence and genetic organization of a *Neurospora* mitochondrial plasmid suggest a relationship to introns and mobile elements. *Cell* 38: 441–453.

- Nedelcu AM, Lee RW. 1998. Short repetitive sequences in green algal mitochondrial genomes: potential roles in mitochondrial genome evolution. *Mol Biol Evol.* 15:690–701.
- Oeser B, Rogmann-Backwinkel P, Tudzynski P. 1993. Interaction between mitochondrial DNA and mitochondrial plasmids in *Claviceps purpurea*: analysis of plasmid-homologous sequences upstream of the rRNA-gene. *Curr Genet.* 23:315–322.
- Pande S, Lemire EG, Nargang FE. 1989. The mitochondrial plasmid from *Neurospora intermedia* strain Labelle-1b contains a long open reading frame with blocks of amino acids characteristic of reverse transcriptases and related proteins. *Nucleic Acids Res.* 17:2023–2042.
- Papadopoulos JS, Agarwala R. 2007. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* 23:1073–1079.
- Paquin B, Laforest M-J, Lang BF. 2000. Double-hairpin elements in the mitochondrial DNA of *Allomyces*: evidence for mobility. *Mol Biol Evol.* 17:1760–1768.
- Pawlowska TE, Taylor JW. 2004. Organization of genetic variation in individuals of arbuscular mycorrhizal fungi. *Nature* 427:733–737.
- Purin S, Morton J. 2012. Anastomosis behavior differs between asymbiotic and symbiotic hyphae of *Rhizophagus clarus*. *Mycologia* 105: 589–602.
- Raab PA, Brennwald A, Redecker D. 2005. Mitochondrial large ribosomal subunit sequences are homogeneous within isolates of *Glomus* (arbuscular mycorrhizal fungi, Glomeromycota). *Mycol Res.* 109:1315–1322.
- Robison MM, Royer JC, Horgen PA. 1991. Homology between mitochondrial DNA of *Agaricus bisporus* and an internal portion of a linear mitochondrial plasmid of *Agaricus bitorquus*. *Curr Genet.* 19:495–502.
- Robison MM, Wolyn DJ. 2005. A mitochondrial plasmid and plasmid-like RNA and DNA polymerases encoded within the mitochondrial genome of carrot (*Daucus carota*). *Curr Genet.* 47:57–66.
- Rosewich UL, Kistler HC. 2000. Role of horizontal gene transfer in the evolution of fungi. *Annu Rev Phytopathol.* 38:325–363.
- Ruiz-Lozano JM, Azcon R, Gomez M. 1995. Effects of arbuscular-mycorrhizal *Glomus* species on drought tolerance: physiological and nutritional plant responses. *Appl Environ Microbiol.* 61:456–460.
- Sakaguchi K. 1990. Invertrons, a class of structurally and functionally related genetic elements that includes linear DNA plasmids, transposable elements, and genomes of adeno-type viruses. *Microbiol Rev.* 54: 66–74.
- Sanders IR, Alt M, Gropp K, Boller T, Wiemken A. 1995. Identification of ribosomal DNA polymorphisms in spores of the Glomales: application to studies on the genetic diversity of arbuscular mycorrhizal fungal communities. *New Phytol.* 130:419–427.
- Saville BJ, Kohli Y, Anderson JB. 1998. mtDNA recombination in a natural population. *Proc Natl Acad Sci U S A.* 95:1331–1335.
- Schofield MA, Agbunag R, Miller JH. 1992. DNA inversions between short inverted repeats in *Escherichia coli*. *Genetics* 132:295–302.
- Schulte U, Lambowitz AM. 1991. The LaBelle mitochondrial plasmid of *Neurospora intermedia* encodes a novel DNA polymerase that may be derived from a reverse transcriptase. *Mol Cell Biol.* 11:1696–1706.
- Shao Z, Graf S, Chaga OY, Lavrov DV. 2006. Mitochondrial genome of the moon jelly *Aurelia aurita* (Cnidaria, Scyphozoa): a linear DNA molecule encoding a putative DNA-dependent DNA polymerase. *Gene* 381: 92–101.
- Shiu-Shing Chan B, Court DA, Vierula PJ, Bertrand H. 1991. The kalilo linear senescence-inducing plasmid of *Neurospora* is an invertron and encodes DNA and RNA polymerases. *Curr Genet.* 20: 225–237.
- Signorovitch AY, Buss LW, Dellaporta SL. 2007. Comparative genomics of large mitochondria in Placozoans. *PLoS Genet.* 3:e13.
- Smith S, Read D, editors. 2008. *Mycorrhizal symbiosis*. Cambridge: Academic Press.
- St-Arnaud M, Vujanovic V. 2007. Effects of arbuscular mycorrhizal fungi on plant diseases and pests mycorrhizae in crop production: applying knowledge. Binghamton (NY): Haworth Press.
- Stockinger H, Walker C, Schussler A. 2009. '*Glomus intradices* DAOM197198', a model fungus in arbuscular mycorrhiza research, is not *Glomus intradices*. *New Phytol.* 183:1176–1187.
- Stover B, Muller K. 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics* 11:7.
- Tanaka Y, Tsuda M, Yasumoto K, Yamagishi H, Terachi T. 2012. A complete mitochondrial genome sequence of Ogura-type male-sterile cytoplasm and its comparative analysis with that of normal cytoplasm in radish (*Raphanus sativus* L.). *BMC Genomics* 13:352.
- Thiéry O, Börstler B, Ineichen K, Redecker D. 2010. Evolutionary dynamics of introns and homing endonuclease ORFs in a region of the large subunit of the mitochondrial rRNA in *Glomus* species (arbuscular mycorrhizal fungi, Glomeromycota). *Mol Phylogenet Evol.* 55: 599–610.
- Voets L, De La Providencia IE, Declerck S. 2006. *Glomeraceae* and *Gigasporaceae* differ in their ability to form hyphal networks. *New Phytologist* 172:185–188.
- Volz-Lingenhöhl A, Salignac M, Sperlich D. 1992. Stable heteroplasmy for a large-scale deletion in the coding region of *Drosophila subobscura* mitochondrial DNA. *Proc Natl Acad Sci U S A.* 89:11528–11532.
- Wehner J, Antunes PM, Powell JR, Mazukatow J, Rillig MC. 2010. Plant pathogen protection by arbuscular mycorrhizas: a role for fungal diversity? *Pedobiologia* 53:197–201.
- Weiller GF, Bruckner H, Kim SH, Pratič E, Schweyen RJ. 1991. A GC cluster repeat is a hotspot for mit—macro-deletions in yeast mitochondrial DNA. *Mol Gen Genet.* 226:233–240.
- White DJ, Wolff JN, Pierson M, Gemmill NJ. 2008. Revealing the hidden complexities of mtDNA inheritance. *Mol Ecol.* 17:4925–4942.
- White MF, Lilley DM. 1997. Characterization of a Holliday junction—resolving enzyme from *Schizosaccharomyces pombe*. *Mol Cell Biol.* 17:6465–6471.
- Xavier BB, Miao VPW, Jónsson ZO, Andrésón ÓS. 2012. Mitochondrial genomes from the lichenized fungi *Peltigera membranacea* and *Peltigera malacea*: features and phylogeny. *Fungal Biol.* 116:802–814.
- Xia X, Xie Z. 2001. DAMBE: software package for data analysis in molecular biology and evolution. *J Hered.* 92:371–373.
- Yang X, Griffiths AJ. 1993a. Male transmission of linear plasmids and mitochondrial DNA in the fungus *Neurospora*. *Genetics* 134: 1055–1062.
- Yang X, Griffiths AJ. 1993b. Plasmid diversity in senescent and non-senescent strains of *Neurospora*. *Mol Gen Genet.* 237:177–186.
- Yasuhira S, Yasui A. 2000. Alternative excision repair pathway of UV-damaged DNA in *Schizosaccharomyces pombe* operates both in nucleus and in mitochondria. *J Biol Chem.* 275:11824–11828.
- Yin S, Heckman J, RajBhandary UL. 1981. Highly conserved GC-rich palindromic DNA sequences flank tRNA genes in *Neurospora crassa* mitochondria. *Cell* 26:325–332.
- Zimmer M, Krabusch M, Wolf K. 1991. Characterization of a novel open reading frame, urf a, in the mitochondrial genome of fission yeast: correlation of urf a; mutations with a mitochondrial mutator phenotype and a possible role of frameshifting in urf a; expression. *Curr Genet.* 19:95–102.

Associate editor: John Archibald