微生物学报
Acta Microbiologica Sinica
2017, 57(5):724-736
http://journals.im.ac.cn/actamicrocn
DOI: 10.13343/j.cnki.wsxb.20160416

Research Article 研究报告

Reanalysis of the mitochondrial genome of the pneumocandin-producing fungus *Glarea lozoyensis*

Yongjie Zhang[1*], Yuxiang Zhao[1], Shu Zhang[1], Li Chen[2], Xingzhong Liu[3]

[1] School of Life Sciences, Shanxi University, Taiyuan 030006, Shanxi Province, China
[2] Department of Cell and Developmental Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, U.S.A.
[3] State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

Abstract: [Objective] *Glarea lozoyensis* is a filamentous fungus used for industrial production of the antifungal drug caspofungin. Previously, the mitochondrial genome (mitogenome) of a mutant strain ATCC 74030 was reported. The purpose of the current study is to test if mutagen treatments have caused changes on the mitogenome of the fungus. **[Methods]** The mitogenome of the wild strain ATCC 20868 was assembled and compared with the published mitogenome of ATCC 74030. PCR assays were done for both strains. Additional analyses were done using correct mitogenome sequences. **[Results]** We successfully assembled the mitogenome of the wild strain ATCC 20868. Initial comparison of the mitogenomes of the wild and mutant strains indicated six variable nucleotide sites and two regions with length variations. PCR assays and subsequent sequencing, however, showed no difference between the two strains. The differences observed from initial comparison were due to sequence errors present in the published mitogenome of ATCC 74030. Interestingly, three intron-containing tRNAs and a *rnpB* gene were detected in the mitogenome of the fungus. Obvious repetitive elements were identified within the *G. lozoyensis* mitogenome, and duplication events were identified between its mitochondrial and nuclear genomes. **[Conclusion]** We verified that there existed erroneous sequences in the published mitogenome of ATCC 74030; mutagens did not cause variations on the mitogenome of *G. lozoyensis*. We reported the authentic mitogenome sequence of *G. lozoyensis* and found frequent gene transfer between mitochondrial and nuclear genomes in the fungus.

Keywords: gene transfer, *Glarea lozoyensis*, intron-containing tRNA, mitochondrial genome, mutagen

Glarea lozoyensis Bills & F. Peláez, formerly classified as *Zalerion arboricola* Buczacki, is an anamorphic fungus belonging to the Leotiales order of ascomycetes[1]. It produces lipopeptides with antifungal activities called pneumocandins which belong to the group of the echinocandin antibiotics[2]. The fungus was originally discovered by plating filtrates of pond water near Madrid, Spain in 1985[1]. The wild-type

Supported by the National Natural Science Foundation of China (81102759), by the Natural Science Foundation of Shanxi Province (2014021030-2, 201601D011065), by the Fund Program for the Scientific Activities of Selected Returned Overseas Professionals in Shanxi Province and by the Special Fund for Large Scientific Instruments and Equipments in Shanxi Province
*Corresponding author. E-mail: zhangyj2008@sxu.edu.cn
Received: 15 October 2016; Revised: 2 December 2016; Published online: 21 December 2016

strain ATCC 20868 produces pneumocandin A_0 predominantly[3]. From the parent wild strain, several mutant strains were obtained during development of pneumocandin B_0 as a product candidate[4]. Pneumocandin B_0 is chemically converted into caspofungin (CANCISAS^{TM}), a potent antibiotic against clinically important fungal pathogens[5–6]. Among those mutant strains, ATCC 74030, which was obtained after two cycles of mutagenesis, is an overproducer of pneumocandin B_0[4]. During two dozen years after the description of the medicinally important fungus, only one wild strain (i.e., ATCC 20868 = CBS 492.88) is known. Until very recently, new *G. lozoyensis* strains were reported from soil or leaf litter in Argentina and the USA, indicating a wide distribution of this fungus[7].

As an antibiotic-producing fungus, researches regarding *G. lozoyensis* have traditionally focused on new potent metabolite exploration[2], strategies for increasing target metabolite yield[8–10], and understanding molecular bases underlying its metabolite production[11–12]. To elucidate the biosynthetic pathway to pneumocandins, genome sequences of the wild strain ATCC 20868 and the mutant strain ATCC 74030 were recently sequenced[13–14]. About 50 secondary metabolite biosynthesis gene clusters were predicted from the genome[13]. The gene cluster responsible for pneumocandin production was predicted *in silico* and identified by deletion of the core genes *glpks4* and *glnrps4* and bioassay experiments[13]. Genetic manipulations of the biosynthetic pathway changed it metabolic profile[15–16].

In the scope of a whole genome sequencing of ATCC 20868, the complete mitochondrial genome (mitogenome) of *G. lozoyensis* ATCC 20868 was assembled by our research group. During the implementation of our project, the mitogenome of the mutant strain ATCC 74030, which is generated from ATCC 20868, was published[17]. Comparison of the two mitogenomes showed several nucleotide differences and the translocation of a large fragment. In one of our recent investigations, we found that chemical mutagenesis disrupted the *GLOXY4* gene function of *G. lozoyensis* by introducing two amino acid mutations in GLOXY4[15]. We suspected whether the variations observed between the two mitogenomes also resulted from mutagen treatments used for generating ATCC 74030. To test our hypothesis, we performed PCR assays with genomic DNAs extracted from the two *G. lozoyensis* strains as templates. Different from our expectations, however, no difference was detected mitogenomes between the parental and mutant strains. In this study, we provided authentic mitogenome sequences of *G. lozoyensis*, re-annotated the mitogenome, and performed additional analysis.

1 Materials and methods

1.1 Fungal material, cultivation and DNA extraction

Two *G. lozoyensis* strains ATCC 20868 (wild-type) and ATCC 74030 (mutant strain) were grown for 20 days at 25 °C on potato dextrose agar plates. They were then subcultured for additional 20 days on the same medium but with a piece of sterile cellophane paper covering the medium surface. Fresh mycelia were scraped off the plate and then used for extracting total genomic DNA using the cetyltrimethylammonium bromide method[18].

1.2 Assembly of the complete mitogenome of *G. lozoyensis* ATCC 20868

When we initiated the current project, the mitogenome of ATCC 74030 has not been released. To assemble the mitogenome of ATCC 20868, we performed local BLAST searches against the raw genome data of *G. lozoyensis* ATCC 20868 using the *Phialocephala subalpina* mitogenome sequence (NC_015789) as the query. Two scaffolds harboring mitochondrial genes were identified, but they were incomplete due to existence of 25 gap regions. To fill the gaps and to join the two scaffolds, multiple primer pairs were designed based on known sequences using the online program Primer3 (http://bioinfo.ut.ee/primer3/). PCRs were performed

using a high-fidelity DNA polymerase KOD FX (TOYOBO Bio-Technology Co. LTD, Japan), and sequences of amplicons were determined by Sanger sequencing at SinoGenoMax Co. LTD. (Beijing, China).

1.3 Comparison between the mitogenomes of ATCC 74030 and 20868

The mitogenome of *G. lozoyensis* ATCC 20868 assembled in this study was aligned with the published mitogenome of ATCC 74030 using the online alignment program MAFFT (http://mafft.cbrc.jp/alignment/server/). Nucleotide differences from the alignment were illustrated by Mauve 2.3.1[19]. For regions showing difference between the two mitogenomes, PCR primers were designed according to flanking sequences in order to verify if the observed difference was true (Table 1). PCR amplification and sequence determination referred to those described above.

1.4 Annotation of the mitogenome of ATCC 20868

Since there are sequence errors in the mitogenome of ATCC 74030 published by Youssar *et al*. (2013), we provided a new annotation of the mitogenome of *G. lozoyensis* ATCC 20868. The mitogenome of *G. lozoyensis* ATCC 20868 was first annotated automatically using the MFannot tool (http://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl) based on the Mold/Protozoan/Coelenterate mitochondrial genetic code (i.e., genetic

Table 1. Primers used in this study

Locus	Primer names	Primer sequences (5′→3′)	Direction	Annealing temp./°C	Expected size/bp
VG1 [a]	VG1-F1	CAATATACTCAGATTCATCAAGCGT	F	51	940
	VG1-R1	TGTGTGAAGGCTGAAGAGATATAA	R		
VG2 [a]	VG2-F1	GCTATCCCTAGTACGCAGCT	F	52	1700
	VG2-R1	ACCCGTTGTAATTCCTAGTATACCT	R		
R3937_mt [b]	3937mt-F1	GGCCACCTAAGAACTACCCG	F	51	4402
	3937mt-R1	TCTTGAAAATAGAGTCCCCAGG	R		
R3937_nr [b]	3937nr-F2	AATAGGCACAAGAGATATAGGACC	F	51	4212
	3937nr-R2	ATCCGCGCTATAAATACTAGAATTC	R		
R585_mt [b]	585mt-F1	ACCCAAAACTAACCCAGGTC	F	54–51	1053
	585mt-R1	ACACACATCGCTTGGAAGGC	R		
R585_nr [b]	585nr-F1	TTACCGGAATGTTGCTGCTG	F	52	1551
	585nr-R1	TTTCATGGGCACTGTCTGAC	R		
R418_mt [b]	418mt-F1	CCAGCTATGTCAGTATTAGCCG	F	52	958
	418mt-R1	TCAACCTGTACCTGCACCAT	R		
R418_nr [b]	418nr-F1	GTGCCTGAGAAGTGAACGATT	F	52	731
	418nr-R1	ACTAGCCAACCTCTCTGAGTG	R		
R388_mt [b]	388mt-F1	GACAAGTTTAACCGTTCGCCT	F	51	845
	388mt-R1	ACTTACCGATTAGTCCACAACAT	R		
R388_nr [b]	388nr-F1	AACTTCAACCTCGGGAGCTT	F	52	1052
	388nr-R1	ATCGCTGGTGTATTTGGAGC	R		

[a] These fragments were amplified to verify if there is indeed length variation between ATCC 20868 and ATCC 74030. Refer to Figure 1 for positions of VG1 and VG2. [b] These fragments were amplified to confirm sequence authenticity of corresponding mitochondrial (mt) and nuclear (nr) fragments generated by blasting the mitogenome against the nuclear genome. Primers anchored at regions flanking the target fragment. For the mitochondrial and nuclear partners of each alignment, different expected sizes were considered when designing primers to facilitate the judging of correct amplification.

code 4), and then manual corrections were performed. First, the boundaries of rRNA genes (especially for the 5′ and 3′ termini of *rnl*) were determined by aligning with rRNAs from other Leotiomycetes whose mitogenomes were reported. Second, tRNA genes were predicted by RNAweasel (http://megasun.bch.umontreal.ca/cgi-bin/RNAweasel/RNAweaselInterface.pl), tRNAscan-SE 1.3.1[20], ARAGORN 1.2.36[21], and ARWIN 1.2.3[22] using genetic code 4. A potential tRNA was considered as proven when it was found by at least two of these tools. Third, introns within rRNA genes and protein-coding genes were identified by aligning with corresponding intronless genes from a related species and for protein-coding genes also by BLASTX searches of remaining sequences after removing putative introns. Fourth, open reading frames (ORFs) within introns and intergenic regions were identified using ORF Finder (http://www.ncbi.nlm.nih.gov), BLASTX searches against NCBI databases, HMM searches against the protein families in the Pfam 28.0 database[23], and ExPASy translation tool (http://web.expasy.org/translate/). Those having no significant similarity to known genes were annotated as hypothetical proteins. Only ORFs > 300 bp were considered.

1.5 Identification of repetitive elements

To know if there is intra-genomic duplication of large fragments and to detect interspersed repeats, a local BLASTn search of the whole mitogenome against itself was performed; matches with E value < 10^{-5} were taken into account. These hits were clustered as a function of similarity in CD-HIT Suite[24], and the h-cd-hit-est algorithm was run with three consecutive cut-off values of 1.00, 0.90 and 0.80, where repetitive sequences were first clustered at a higher identity and then non-redundant sequences were further clustered at a lower identity.

In addition, tandem repeats were analyzed by the online program Tandem Repeat Finder (http://tandem.bu.edu/trf/trf.basic.submit.html) using default parameters. Simple sequence repeats (SSRs) were detected by SSR Finder (http://www.csufresno.edu/ssrfinder/) under default parameters. Repeated sequences were also searched by REPuter (https://bibiserv2.cebitec.uni-bielefeld.de/reputer), which identifies and locates forward (direct), reverse, complement, and palindromic (revere complement) repeats. Default settings of REPuter were used, but we filtered to use only repeats with *E* value < 10^{-5}.

1.6 Gene transfer between the mitochondrial and nuclear genomes

To know if there is gene transfer between mitochondrial and nuclear genomes of the fungus, we performed BLAST searches of the *G. lozoyensis* ATCC 20868 mitogenome against its nuclear genome (GenBank accession ALVE00000000) with the same strategy as described above. Sequence authenticity of corresponding mitochondrial and nuclear fragments related to some large alignments was assayed by PCR amplification using designed primers (Table 1). Origins of these fragments were deduced by performing online BLAST searches.

1.7 GenBank accession number

The mitogenome sequence of *G. lozoyensis* ATCC 20868 was deposited in GenBank under the accession number KX450332.

2 Results and analyses

2.1 Confirmation of the mitogenome sequence of *G. lozoyensis*

The mitogenome of the wild-type strain ATCC 20868 was successfully assembled as a circular molecule of 45501 bp in length. It was 463 bp longer than the published mitogenome of the mutant strain ATCC 74030, which is 45038 bp in length (GenBank accession no. KF169905). Sequence alignment indicated six variable nucleotide sites and two insertion/deletion (indel) regions between the two mitogenomes (Figure 1). At the first indel region (VG1), ATCC 74030 was 354 bp longer than ATCC 20868; at the second indel region (VG2), ATCC

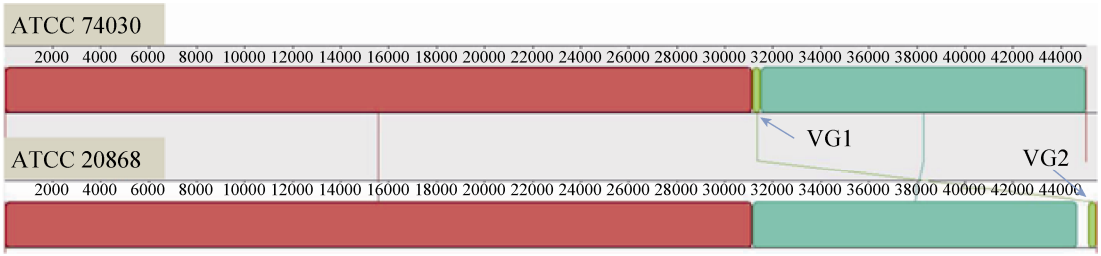

Figure 1. Alignment between the published mitogenome of ATCC 74030 and that of ATCC 20868 assembled in this study. Each locally collinear block was indicated by a unique color. The alignment revealed two major varying regions (VG1 and VG2) plus six variable nucleotide sites (all adjacent to 5′ VG1). The two varying regions showed length variations between the two mitogenomes and involved the jumping of a fragment from one region to another.

20868 was 817 bp longer than ATCC 74030. A fragment seemed to jump between the two regions of both strains (Figure 1). PCR assays performed in this study, however, did not find length variation at any of the two indel regions between the two strains (Figure 2). Neither other nucleotide differences were confirmed in this study. Actually, the mitogenomes of ATCC 20868 and ATCC 74030 are quite identical without any nucleotide difference.

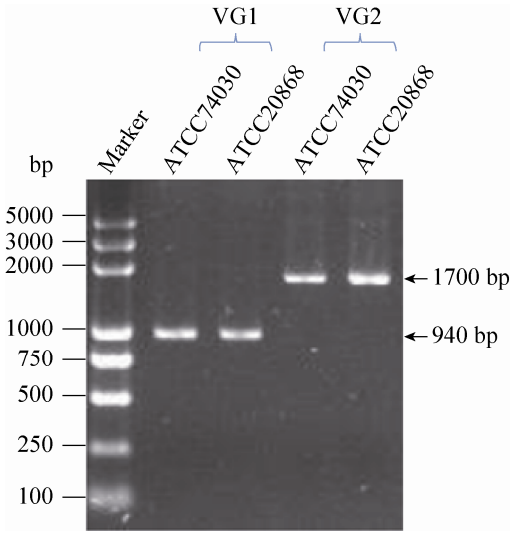

Figure 2. PCR assays of the two varying regions. Primers were designed to amplify the two varying regions observed in Figure 1. Inconsistent to the alignment shown in Figure 1, mitogenomes of ATCC 74030 and 20868 are actually identical without any nucleotide variation.

2.2 Annotation of the mitogenome of *G. lozoyensis*

The mitogenome of *G. lozoyensis* contained two ribosomal RNAs (*rnl* and *rns*), 33 tRNAs, 14 standard proteins of the oxidative phosphorylation system, and five free-standing ORFs (Table 2, Figure 3). The 33 tRNAs coded for all 20 standard amino acids, and three (*trnN*_2, *trnI*_2 and *trnI*_3) contained an intron of 6–23 bp. For the five free-standing ORFs, four (*orf291*, *orf254*, *orf309* and *orf132*) are in the forward strand, and one (*orf104*) is in the reverse strand. Two ORFs, *orf132* and *orf104*, located on opposite strands but overlapped by 248 bp. An *rnpB* gene (RNase P RNA) was identified at the *trnC*/*nad1* intergenic region.

Five introns were found in ribosomal RNA/protein-encoding genes, including one each in *rnl* and *cox3*, and three in *cox1* (Table 2). Three others introns (one each in *nad2* and *nad5*, and an additional intron in *cox1*) identified by Youssar *et al*. (2013) were not recognized by us. Within the five introns that we identified, the *cox3* intron and *cox1*-i3 harbored no intronic ORFs, but other three introns harbored an intronic ORF encoding for RPS3 (in *rnl*), a LAGLIDADG homing endonuclease (in *cox1*-i1), and a GIY-YIG homing endonuclease (in *cox1*-i2). Degeneration of the intronic ORF in *cox1*-i2 was obvious because of the detection of frame shifts and multiple stop codon mutations.

Table 2. General features of the mitogenome of *G. lozoyensis*

Gene [a]	Start	End	Length/bp	Start codon	Stop codon	Anti-codon	Notes
nad4	299	1783	1485	ATG	TAA		
rns	2048	3844	1797				
trnY	4728	4812	85			GTA	
trnN_1	4894	4964	71			GTT	
trnR_1	4973	5044	72			TCG	
nad6	5332	5988	657	ATG	TAA		
trnV	6063	6134	72			TAC	
cox3	6586	8072	1487	GTG	TAA		With an intron
trnK_1	8329	8400	72			TTT	
trnG_1	8483	8553	71			TCC	
atp9	8790	9014	225	ATG	TAA		
trnD	9065	9136	72			GTC	
trnS	9210	9289	80			GCT	
trnW	9332	9402	71			TCA	
trnI_1	9558	9629	72			GAT	
trnS	9743	9828	86			TGA	
trnP	9915	9987	73			TGG	
rnl	10095	15989	5895				With an intron
trnT	16151	16221	71			TGT	
trnE	16225	16296	72			TTC	
trnM_1	16299	16369	71			CAT	
trnM_2	16503	16575	73			CAT	
trnL	16617	16699	83			TAA	
trnA	17546	17618	73			TGC	
trnF	17674	17746	73			GAA	
trnL	17751	17835	85			TAG	
trnQ	17866	17938	73			TTG	
trnH	18017	18089	73			GTG	
trnM_3	18168	18240	73			CAT	
orf291	18809	19684	876	ATG	TAA		
trnR	20018	20088	71			TCT	
trnK_2	20163	20234	72			TTT	
trnG_2	20317	20387	71			TCC	
trnN_2	20928	21021	94			GTT	With an intron
trnR_2	21030	21101	72			TCG	
orf254	21424	22188	765	ATG	TAA		
cob	22532	23698	1167	ATG	TAG		
orf309	23781	24710	930	ATG	TAG		
orf104	26011	25697	315	ATG	TAG		At reverse strand
orf132	25764	26162	399	ATG	TAG		
trnF_1	26274	26354	81			AAA	
nad4L	26748	27017	270	ATG	TAA		
nad5	27017	29551	2535	ATG	TAG		
trnF_2	29903	29981	79			AAA	
cox2_1	30683	31165	483	TTA	TAA		Incomplete *cox2* copy
cox1	31455	36881	5427	ATG	TAA		With 3 introns

(To be continued)

Table 2 continued

trnI_2	37115	37193	79		GAT	With an intron
nad2	37195	39066	1872	GTG	TAA	
nad3	39067	39549	483	ATG	TAG	
trnI_3	39806	39884	79		GAT	With an intron
atp8	40546	40692	147	ATG	TAA	
atp6	40868	41644	777	ATG	TAG	
trnC	41657	41728	72		GCA	
rnpB	41835	42071	237			
nad1	43041	44123	1083	ATG	TAA	
cox2_2	44661	45416	756	ATG	TAA	Complete *cox2* copy

[a] A total 33 tRNA genes are identified and they code for all 20 standard amino acids. Most of them are clustered at regions flanking *rnl* (6 before *rnl* and 16 after *rnl*) and *rns* (3 after *rns*). For methionine, three tRNA genes are found with the same anticodon. Three tRNAs are identified for phenylalanine, isoleucine and arginine, but each with two different anticodons. For glycine, lysine and asparagine, there are two tRNA genes with identical anticodons. For leucine and serine, there are two tRNA genes with different anticodons. Three intron-containing tRNAs (trnN_2, trnI_2, and trnI_3) were also identified and they harbored an intron of 6–23 bp.

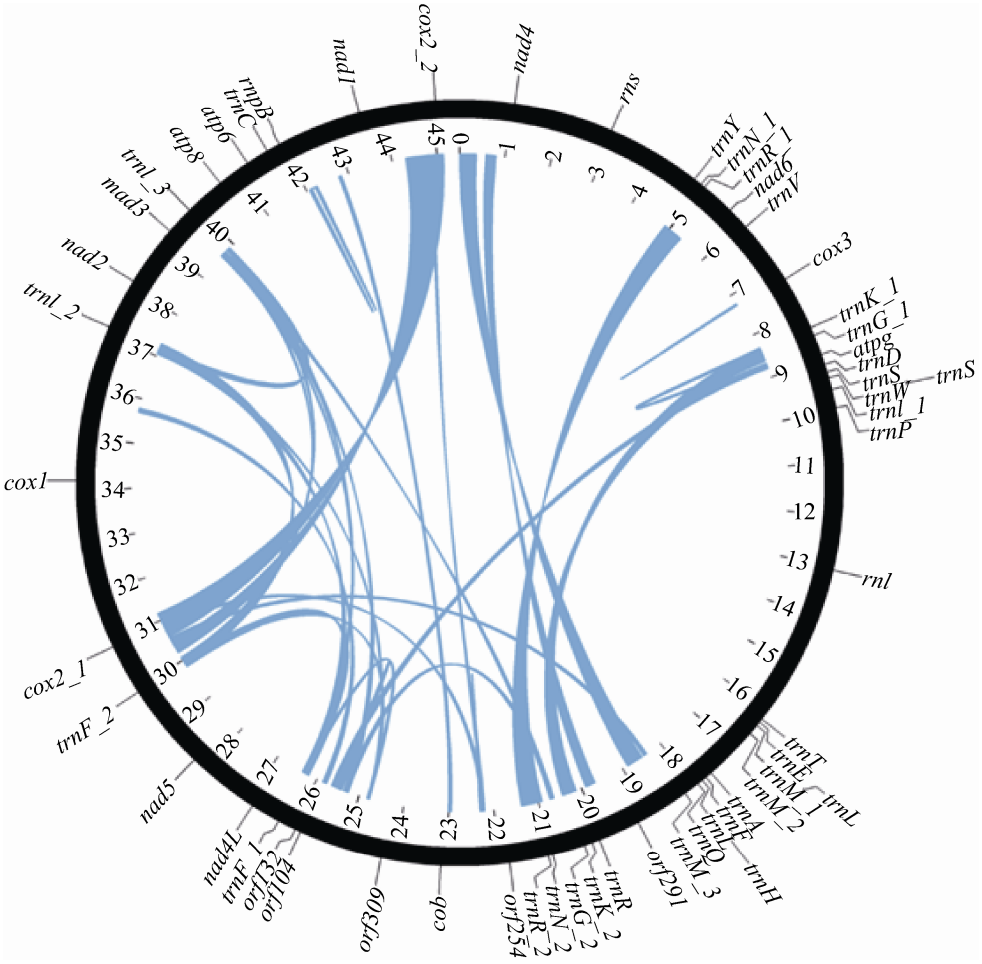

Figure 3. Circular map of the *G. lozoyensis* mitogenome. The mitogenome of *G. lozoyensis* contained two ribosomal RNAs (*rnl* and *rns*), 33 tRNAs, 14 standard proteins of the oxidative phosphorylation system, and five free-standing ORFs. An *rnpB* gene and an additional incomplete copy of *cox2* (*cox2_1*) were also identified. The precise positions of genes are listed in Table 2. Blue ribbons indicate repetitive elements occurring between different mitochondrial regions.

Intra-genomic duplication events were detected. For example, partial sequences of *orf132*/*orf104*, which located at around 5.7 kb upstream of *cox1*, showed some similarity (85%, 81/95) to the *cox1* gene. Two copies of *cox2*, one being incomplete (*cox2*_1, 483 bp) and one being complete (*cox2*_2, 756 bp) were identified, and they showed a similarity of 98% (467/476). Locations of *cox2*_1 and *cox2*_2 corresponded to the two above-mentioned indel regions (VG1 and VG2), and are separated by about 13.5 kb in the mitogenome. In the published mitogenome of ATCC 74030, *cox2*_1 and *cox2*_2 erroneously hybridized as one gene. This is also the main sequence difference between the mitogenome reported by Youssar *et al*. (2013) and that reported in the present study.

2.3 Repetitive sequences in the mitogenome of *G. lozoyensis*

Since we identified intra-genomic duplication events in the above analyses, it is interesting to know if there are other repetitive elements in the *G. lozoyensis* mitogenome. BLASTn analysis of the whole mitogenome against itself revealed 69 repeat alignments (Figure 3). These repeat alignments showed 76.1%–100% identities at an alignment length of 30–559 bp. Of these repeats, 60 were direct repeats (plus/plus matches), and 9 were inverted repeats (plus/minus matches). The largest direct repeat (a 559 bp alignment) occurred between *cox2_1* and *cox2_2*; the largest inverted repeat (a 111 bp alignment) occurred between the *trnG*_1/*atp9* intergenic region and the *orf309*/*orf104* intergenic region. After clustering analysis (21 repeats left), repeated sequences accounted for about 7.3% (3323 bp) of the *G. lozoyensis* mitogenome.

In addition, we found 12 tandem repeats in a total of 726 bp, accounting for 1.6% of the mitogenome, ranging from 14 to 43 bp in the period size. Total length of SSRs identified by SSR Finder was 804 bp (< 1.8% of the mitogenome) with only dimers and trimers detected. REPuter was used to identify 40 forward (in total 2621 bp), one reverse (23 bp) and three palindromic (76 bp) repeats, accounting for 5.9% of the mitogenome and targeting either intergenic or coding regions. No complement repeat was identified by REPuter.

2.4 Gene transfer between the mitochondrial and nuclear genomes of *G. lozoyensis*

BLASTN analyses of the mitogenome of *G. lozoyensis* ATCC 20868 against its nuclear genome revealed 27 alignments with 75%–100% identities at the lengths of 33–3937 bp (Table 3). These mitochondrial fragments mainly located at either the coding or intergenic regions, and their nuclear partners distributed in 13 of the total 22 scaffolds of the nuclear genome of *G. lozoyensis* ATCC 20868 (Table 3; Figure 4-A). After clustering analysis, 6399 bp (14.1%) of the mitogenome sequences (21 alignments after de-replication) showed similarities with nuclear genome sequences of *G. lozoyensis*. Mitochondrial and nuclear fragments related to the four largest alignments were all confirmed by PCR assays (Figure 4-B), excluding the possibility of sequence errors/contaminations. Corresponding nuclear/mitochondrial fragments all had best hits with mitochondrial sequences from fungi as revealed by online BLASTn searches, suggesting their gene transfer from mitochondrion to the nucleus.

Discussion

In this study, we assembled the mitogenome of the pneumocandin-producing fungus *G. lozoyensis* using a wild-type strain ATCC 20868. We initially suspected that mutagen treatments might have caused mutations in the mitogenome of the mutant strain ATCC 74030. PCR assays and subsequent Sanger sequencing, however, revealed that the mitogenomes of ATCC 20868 and ATCC 74030 were exactly identical without any nucleotide difference. The two chemical mutagens (*N*-nitroso-*N*-methylurethane and *N*-methyl-*N*-

Table 3. Local BLAST analysis of the ATCC 20868 mitochondrial genome against its nuclear genome

Query start	Query end	Subject sequence ID	Subject scaffolds	Subject sequence length/bp	Subject start	Subject end	Subject strand	Alignment length/bp	Percent identity /%	Bit score	*E* value
44843	44899	KE145353	GLAREA10	2366044	500534	500478	minus	57	84.21	56.5	1.00E-06
30865	30921	KE145353	GLAREA10	2366044	500534	500478	minus	57	82.46	51.0	5.00E-05
31220	**31631**	**KE145354**	**GLAREA11**	**1492152**	**926036**	**926447**	**plus**	**418**	**77.99**	**252.0**	**8.00E-66**
43704	43742	KE145355	GLAREA12	1502965	1484370	1484408	plus	39	100.00	73.1	7.00E-12
35098	35157	KE145357	GLAREA14	2105285	1475927	1475987	plus	61	90.16	78.7	2.00E-13
37291	37340	KE145360	GLAREA17	1175165	481697	481648	minus	50	92.00	71.3	2.00E-11
15056	15105	KE145361	GLAREA18	842895	774071	774120	plus	50	86.00	54.7	1.00E-06
3266	3359	KE145362	GLAREA19	791833	63474	63567	plus	94	93.62	141.0	9.00E-33
40524	**44458**	**KE145365**	**GLAREA21**	**575289**	**245429**	**249364**	**plus**	**3937**	**91.90**	**5500.0**	**0**
30199	30348	KE145365	GLAREA21	575289	249212	249361	plus	150	94.00	228.0	5.00E-59
42141	42208	KE145365	GLAREA21	575289	247090	247155	plus	69	86.96	75.0	7.00E-13
42182	42251	KE145365	GLAREA21	575289	247041	247112	plus	72	80.56	54.7	9.00E-07
42092	42124	KE145365	GLAREA21	575289	247027	246995	minus	33	93.94	51.0	1.00E-05
19398	**19970**	**KE145367**	**GLAREA3**	**3594336**	**3038249**	**3037670**	**minus**	**585**	**85.47**	**593.0**	**3.00E-168**
580	798	KE145367	GLAREA3	3594336	3037962	3037741	minus	222	90.99	296.0	9.00E-79
26760	26868	KE145367	GLAREA3	3594336	2890209	2890313	plus	110	86.36	115.0	3.00E-24
36109	36162	KE145367	GLAREA3	3594336	1410134	1410187	plus	54	92.59	78.7	3.00E-13
19971	20013	KE145367	GLAREA3	3594336	3037663	3037705	plus	43	93.02	63.9	9.00E-09
36432	36567	KE145368	GLAREA4	1824436	1206355	1206218	minus	139	79.86	99.0	1.00E-19
36608	36719	KE145368	GLAREA4	1824436	1206163	1206053	minus	112	75.00	51.0	4.00E-05
10829	10902	KE145369	GLAREA5	2448679	1459303	1459379	plus	78	84.62	73.1	1.00E-11
392	461	KE145369	GLAREA5	2448679	1459455	1459385	minus	71	84.51	69.4	1.00E-10
17278	**17647**	**KE145372**	**GLAREA8**	**1378442**	**472599**	**472986**	**plus**	**388**	**92.78**	**545.0**	**3.00E-154**
36749	36789	KE145372	GLAREA8	1378442	1263988	1263948	minus	41	95.12	65.8	1.00E-09
21218	21272	KE145373	GLAREA9	2086076	1540686	1540632	minus	55	100.00	102.0	1.00E-20
5170	5215	KE145373	GLAREA9	2086076	1540677	1540632	minus	46	93.48	69.4	1.00E-10
35202	35305	KE145373	GLAREA9	2086076	1340852	1340750	minus	107	78.50	63.9	5.00E-09

The four alignments that are chosen to be confirmed by PCR assays are shown in bold (see Figure 4).

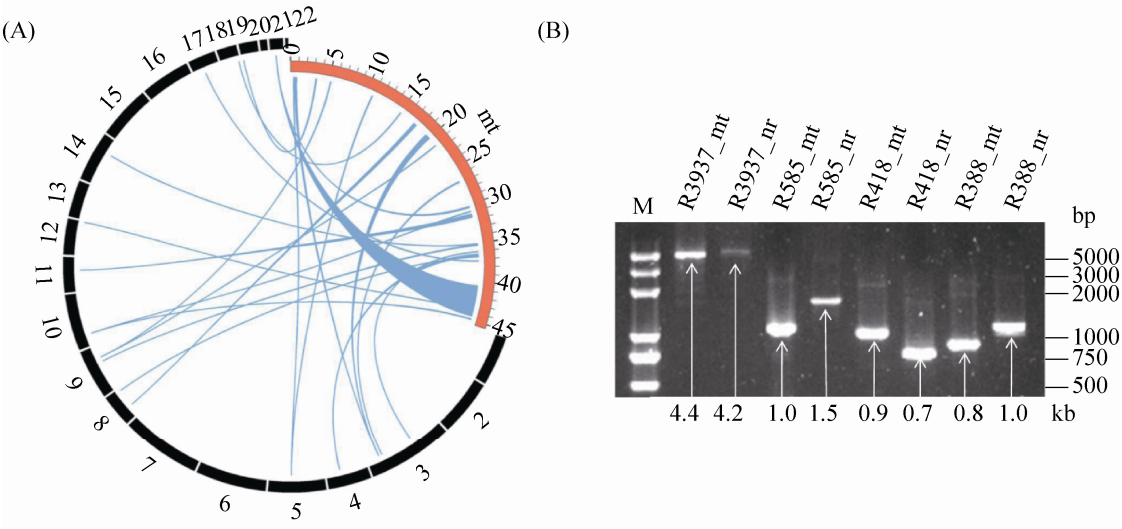

Figure 4. Gene transfer between mitochondrial and nuclear genomes. A: blue ribbons linked regions of the mitogenome (shown in orange) with regions of the nuclear genome (shown in black) with significant similarities. Numbers in the mitogenome indicate positions in kb. The nuclear genome is composed of 22 scaffolds as indicated in numbers (1–22). B: corresponding mitochondrial and nuclear fragments related to the four largest alignments were all confirmed by PCR assays. The DNA ladder is as in Figure 2.

nitro-*N*-nitrosoguanidine) used for generating ATCC 74030 obviously caused mutations on nuclear DNA, as reported previously[15]. Since only one mutant strain was tested in this study, we cannot tell that the two chemical mutagens cannot mutate the mitogenome of *G. lozoyensis* until additional mutant strains are examined. Variation of mitochondrial DNA of an organism has been reported on exposure to factors such as arsenic, smoking and chemical mutagens[25–27].

The presence of erroneous sequence in the published mitogenome of ATCC 74030 is reasonable considering that it was assembled directly by combination of data from Illumina mate-pair and pair-end sequencing[17]. The assembly errors mainly resulted from the presence of two copies of *cox2*. The two *cox2* copies erroneously hybridized together in the published mitogenome of ATCC 74030 although they are actually separated by about 13.5 kb (Table 2). In this study, we provided the correct mitogenome of *G. lozoyensis* using the wild-type strain ATCC 20868.

Due to the presence of sequence errors in the published mitogenome of ATCC 74030, we re-annotated the *G. lozoyensis* mitogenome using ATCC 20868. One interesting finding is the presence of three tRNAs with introns (i.e., *trnN*_2, *trnI*_2, and *trnI*_3), where the introns all located 1 base 3′ to the anticodon. As much as we know, intron-containing tRNAs have not been reported from mitogenomes of eukaryotes but were seen from archaeal, bacterial and eukaryotic nuclear genomes as well as plant plastid genomes[28–29]. Whether the three intron-containing tRNAs are functional and how their introns are removed during tRNA maturation remain unknown. The second interesting finding is the presence of a mitochondrial *rnpB* gene, which encodes the RNA subunit of RNase P (P-RNA), the enzymatically active part of an endonuclease (ribonucleoprotein) responsible for tRNA maturation[30]. The *rnpB* gene has so far only been recognized in the mitogenomes of a few fungal species[30–32], including the fungus *Sclerotinia borealis*[33], who belongs to the same order (i.e.,

Helotiales) as *G. lozoyensis*. It is interesting to know whether *rnpB* accounts for the maturation of intron-containing tRNAs found in the mitogenome of *G. lozoyensis*.

Another major finding is the detection of gene duplication events. On the one hand, some DNA fragment duplications are detected within the mitogenome of *G. lozoyensis*, such as partial sequences of *orf132* or *orf104* showing similarity to the *cox1* gene, an additional incomplete copy of the *cox2* gene, and more than two copies of some tRNA genes (Table 2). The phenomena of DNA duplication within mitogenomes are frequently reported from fungi, especially those species with a large mitogenome[33–34]. On the other hand, evidence for DNA fragment duplications between mitochondrial and nuclear genomes of *G. lozoyensis* is also obvious (Table 3, Figure 4-A). The natural transfer of DNA from mitochondria to the nucleus that generates nuclear copies of mitochondrial DNA (numts) is an ongoing evolutionary process and is estimated in a wide variety of eukaryotes[35]. Exceptions, however, do exist, such as in the ascomycetous fungus *Cordyceps militaris*, where no nuclear scaffolds with significant sequence similarities to mitochondrial sequences were found[36].

参 考 文 献

[1] Bills GF, Platas G, Peláez F, Masurekar P. Reclassification of a pneumocandin-producing anamorph, *Glarea lozoyensis* gen. et sp. nov., previously identified as *Zalerion arboricola*. *Mycological Research*, 1999, 103(2): 179–192.

[2] Schwartz RE, Sesin DF, Joshua H, Wilson KE, Kempf AJ, Goklen KA, Kuehner D, Gailliot P, Gleason C, White R, Inamine E, Bills G, Salmon P, Zitano L. Pneumocandins from *Zalerion arboricola*. I. Discovery and isolation. *The Journal of Antibiotics*, 1992, 45(12): 1853–1866.

[3] Schwartz RE, Giacobbe RA, Bland JA, Monaghan RL. L-671, 329, a new antifungal agent. I. Fermentation and isolation. *The Journal of Antibiotics*, 1989, 42(2): 163–167.

[4] Masurekar PS, Fountoulakis JM, Hallada TC, Sosa MS, Kaplan L. Pneumocandins from *Zalerion arboricola*. II. Modification of product spectrum by mutation and medium manipulation. *The Journal of Antibiotics*, 1992, 45(12): 1867–1874.

[5] McCormack PL, Perry CM. Caspofungin: a review of its use in the treatment of fungal infections. *Drugs*, 2005, 65(14): 2049–2068.

[6] Leonard Jr WR, Belyk KM, Conlon DA, Bender DR, DiMichele LM, Liu J, Hughes DL. Synthesis of the antifungal β-1,3-glucan synthase inhibitor CANCIDAS (caspofungin acetate) from pneumocandin B_0. *Journal of Organic Chemistry*, 2007, 72(7): 2335–2343.

[7] Peláez F, Collado J, Platas G, Overy DP, Martín J, Vicente F, del Val AG, Basilio A, De la Cruz M, Tormo JR, Fillola A, Arenal F, Villareal M, Rubio V, Baral HO, Galán R, Bills GF. Phylogeny and intercontinental distribution of the pneumocandin-producing anamorphic fungus *Glarea lozoyensis*. *Mycology*, 2011, 2(1): 1–17.

[8] Connors N, Petersen L, Hughes R, Saini K, Olewinski R, Salmon P. Residual fructose and osmolality affect the levels of pneumocandins B_0 and C_0 produced by *Glarea lozoyensis*. *Applied Microbiology and Biotechnology*, 2000, 54(6): 814–818.

[9] Connors N, Pollard D. Pneumocandin B_0 production by fermentation of the fungus *Glarea lozoyensis*//An ZQ. Handbook of Industrial Mycology. New York: Marcel Dekker, 2004: 515–538.

[10] Petersen LA, Hughes DL, Hughes R, DiMichele L, Salmon P, Connors N. Effects of amino acid and trace element supplementation on pneumocandin production by *Glarea lozoyensis*: impact on titer, analogue levels, and the identification of new analogues of pneumocandin B_0. *Journal of Industrial Microbiology and Biotechnology*, 2001, 26(4): 216–221.

[11] Lu P, Zhang A, Dennis LM, Dahl-Roshak AM, Xia YQ, Arison B, An Z, Tkacz JS. A gene (*pks2*) encoding a putative 6-methylsalicylic acid synthase from *Glarea lozoyensis*. *Molecular Genetics and Genomics*, 2005, 273(2): 207–216.

[12] Zhang A, Lu P, Dahl-Roshak AM, Paress PS, Kennedy S, Tkacz JS, An Z. Efficient disruption of a polyketide synthase gene (*pks1*) required for melanin synthesis through *Agrobacterium*-mediated transformation of *Glarea lozoyensis*. *Molecular Genetics and Genomics*, 2003, 268(5): 645–655.

[13] Chen L, Yue Q, Zhang XY, Xiang MC, Wang CS, Li SJ, Che YS, Ortiz-López FJ, Bills GF, Liu XZ, An ZQ. Genomics-driven discovery of the pneumocandin biosynthetic gene cluster in the fungus *Glarea lozoyensis*. *BMC Genomics*,

2013, 14: 339.

[14] Youssar L, Grüning BA, Erxleben A, Günther S, Hüttel W. Genome sequence of the fungus *Glarea lozoyensis*: the first genome sequence of a species from the *Helotiaceae* family. *Eukaryotic Cell*, 2012, 11(2): 250.

[15] Chen L, Yue Q, Li Y, Niu XM, Xiang MC, Wang WZ, Bills GF, Liu XZ, An ZQ. Engineering of *Glarea lozoyensis* for exclusive production of the pneumocandin B_0 precursor of the antifungal drug caspofungin acetate. *Applied and Environmental Microbiology*, 2015, 81(5): 1550–1558.

[16] Li Y, Chen L, Yue Q, Liu XZ, An ZQ, Bills GF. Genetic manipulation of the pneumocandin biosynthetic pathway for generation of analogues and evaluation of their antifungal activity. *ACS Chemical Biology*, 2015, 10(7): 1702–1710.

[17] Youssar L, Grüning BA, Günther S, Hüttel W. Characterization and phylogenetic analysis of the mitochondrial genome of *Glarea lozoyensis* indicates high diversity within the order Helotiales. *PLoS One*, 2013, 8(9): e74792.

[18] Zhang YJ, Zhang S, Liu XZ, Wen HA, Wang M. A simple method of genomic DNA extraction suitable for analysis of bulk fungal strains. *Letters in Applied Microbiology*, 2010, 51(1): 114–118.

[19] Darling AE, Mau B, Perna NT. Progressive-mauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*, 2010, 5(6): e11147.

[20] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, 1997, 25(5): 955–964.

[21] Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research*, 2004, 32(1): 11–16.

[22] Laslett D, Canbäck B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics*, 2008, 24(2): 172–175.

[23] Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M. Pfam: the protein families database. *Nucleic Acids Research*, 2014, 42(D1): D222–230.

[24] Huang Y, Niu B, Gao Y, Fu L, Li W. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics*, 2010, 26(5): 680–682.

[25] Bodhicharla R, Ryde IT, Prasad GL, Meyer JN. The tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) induces mitochondrial and nuclear DNA damage in *Caenorhabditis elegans*. *Environmental and Molecular Mutagenesis*, 2014, 55(1): 43–50.

[26] Liu CS, Chen HW, Lii CK, Tsai CS, Kuo CL, Wei YH. Alterations of plasma antioxidants and mitochondrial DNA mutation in hair follicles of smokers. *Environmental and Molecular Mutagenesis*, 2002, 40(3): 168–174.

[27] Sinha S, Giri AK, Chowdhury R, Ray K. Mitochondrial genome variations among arsenic exposed individuals and potential correlation with apoptotic parameters. *Environmental and Molecular Mutagenesis*, 2014, 55(1): 70–76.

[28] Fujishima K, Kanai A. tRNA gene diversity in the three domains of life. *Frontiers in Genetics*, 2014, 5: 142.

[29] Yoshihisa T. Handling tRNA introns, archaeal way and eukaryotic way. *Frontiers in Genetics*, 2014, 5: 213.

[30] Seif E, Leigh J, Liu Y, Roewer I, Forget L, Lang BF. Comparative mitochondrial genomics in zygomycetes: bacteria-like RNase P RNAs, mobile elements and a close source of the group I intron invasion in angiosperms. *Nucleic Acids Research*, 2005, 33(2): 734–744.

[31] Aguileta G, de Vienne DM, Ross ON, Hood ME, Giraud T, Petit E, Gabaldon T. High variability of mitochondrial gene order among fungi. *Genome Biology and Evolution*, 2014, 6(2): 451–465.

[32] Seif ER, Forget L, Martin NC, Lang BF. Mitochondrial RNase P RNAs in ascomycete fungi: lineage-specific variations in RNA secondary structure. *RNA*, 2003, 9(9): 1073–1083.

[33] Mardanov AV, Beletsky AV, Kadnikov VV, Ignatov AN, Ravin NV. The 203 kb mitochondrial genome of the phytopathogenic fungus *Sclerotinia borealis* reveals multiple invasions of introns and genomic duplications. *PLoS One*, 2014, 9(9): e107536.

[34] Férandon C, Xu JP, Barroso G. The 135 kb mitochondrial genome of *Agaricus bisporus* is the largest known eukaryotic reservoir of group I introns and plasmid-related sequences. *Fungal Genetics and Biology*, 2013, 55: 85–91.

[35] Hazkani-Covo E, Zeller RM, Martin W. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genetics*, 2010, 6(2): e1000834.

[36] Zhang YJ, Zhang S, Zhang GZ, Liu XZ, Wang CS, Xu JP. Comparison of mitochondrial genomes provides insights into intron dynamics and evolution in the caterpillar fungus *Cordyceps militaris*. *Fungal Genetics and Biology*, 2015, 77: 95–107.

肺囊康定产生菌 *Glarea lozoyensis* 线粒体基因组序列的再分析

张永杰[1*]，赵宇翔[1]，张姝[1]，陈里[2]，刘杏忠[3]

[1] 山西大学生命科学学院，山西 太原 030006

[2] 美国宾夕法尼亚大学细胞与发育生物学系，宾夕法尼亚州 费城 19104

[3] 中国科学院微生物研究所真菌学国家重点实验室，北京 100101

摘要：【目的】*Glarea lozoyensis* 是抗真菌药物卡泊芬净的产生菌，其突变菌株 ATCC 74030 的线粒体基因组已被报道。我们此前的研究发现诱变剂能引起该菌某些细胞核基因的突变，但诱变剂是否也能引起线粒体 DNA 序列的改变并不清楚。【方法】组装野生型菌株 ATCC 20868 的线粒体基因组，并与发表的突变型菌株 ATCC 74030 的线粒体基因组进行比较。通过 PCR 验证野生和突变菌株线粒体基因组间表现差异之处，并利用正确的线粒体基因组序列进行新的分析。【结果】我们成功组装出野生型菌株 ATCC 20868 的线粒体基因组，通过比较其与发表的 ATCC 74030 的线粒体基因组序列，发现存在 6 处单核苷酸变异位点和 2 处具有长度差异的区域。然而，随后的 PCR 验证和序列比较并没有发现 2 个菌株间存在这些差异。最初观察到的碱基差异是因为发表的 ATCC 74030 线粒体基因组存在序列错误。有趣的是，在 *Glarea lozoyensis* 的线粒体基因组中，我们发现存在 3 个具有内含子的 tRNA 基因和 1 个 *rnpB* 基因。同时，该菌线粒体基因组中存在多种重复序列，在其线粒体和细胞核基因组间也存在明显的 DNA 片段重复事件。【结论】诱变剂没有引起 *G. lozoyensis* 线粒体 DNA 的任何改变；发表的 ATCC 74030 的线粒体基因组存在序列错误。我们报道 *G. lozoyensis* 正确的线粒体基因组序列，并且发现该菌线粒体和细胞核基因组间频繁的基因交流。

关键词：基因转移，*Glarea lozoyensis*，内含子 tRNA，线粒体基因组，诱变剂

(本文责编：张晓丽)

基金项目：国家自然科学基金(81102759)；山西省自然科学基金(2014021030-2，201601D011065)；山西省留学回国人员科技活动择优资助项目；山西省大型科学仪器设备专项

[*]通信作者。E-mail：zhangyj2008@sxu.edu.cn

收稿日期：2016-10-15；修回日期：2016-12-02；网络出版日期：2016-12-21