



## 蛹虫草线粒体 DNA 与细胞核 DNA 进化关系的比较

张姝<sup>1\*</sup>, 崔宁波<sup>2</sup>, 赵宇翔<sup>2</sup>, 张永杰<sup>2\*</sup>

<sup>1</sup>山西大学应用化学研究所, 山西 太原 030006

<sup>2</sup>山西大学生命科学学院, 山西 太原 030006

**摘要:**【目的】分析蛹虫草是否存在核内线粒体 DNA 片段, 比较蛹虫草线粒体 DNA 与细胞核 DNA 的碱基变异程度及所反映的菌株间的系统发育关系。【方法】通过本地 BLAST 或 LAST 对蛹虫草线粒体基因组和核基因组进行序列相似性搜索; 从 10 个已知线粒体基因组的蛹虫草菌株中分别扩增 7 个细胞核蛋白编码基因片段, 并与其在 14 个线粒体蛋白编码基因上的碱基变异情况进行比较。【结果】蛹虫草核基因组中存在 5 处较短的核内线粒体 DNA 片段, 总长只有 278 bp。蛹虫草核 DNA 的变异频率整体上高于线粒体 DNA。核 DNA 和线粒体 DNA 所反映的蛹虫草菌株间的系统发育关系存在显著差异。【结论】蛹虫草线粒体 DNA 与核 DNA 间不存在长片段的基因交流, 二者变异频率不同, 所反映的蛹虫草菌株间的系统发育关系也有差异。本研究增加了对蛹虫草线粒体与细胞核 DNA 进化关系的认识。

**关键词:** 蛹虫草, 线粒体 DNA, 细胞核 DNA, 基因交流, 变异频率

真核生物除细胞核中含有遗传物质外, 在线粒体、叶绿体等细胞器中也有遗传物质。线粒体的主要功能是合成 ATP, 为细胞生命活动提供能量, 但也参与细胞凋亡、老化、离子平衡的维持等过程<sup>[1-2]</sup>。线粒体 DNA 具有多拷贝、基因组较小、单亲遗传(尤其在动植物中)、较高的碱基变异速率等特性, 已被广泛用作生态学、群体遗传学等研究的分子标记<sup>[3]</sup>。通过对线粒体 DNA 的研究,

大大增加了我们对真核生物起源、多样性及复杂性的认识<sup>[4]</sup>。虽然组成线粒体的许多蛋白为核基因编码的产物, 但线粒体基因组也编码少许蛋白, 并且其遗传独立于核基因组。目前, 对于真核细胞中的核 DNA 和线粒体 DNA 的变异速率存在两种截然相反的观点: 动物线粒体 DNA 通常比核 DNA 变异快, 而植物核 DNA 通常比线粒体 DNA 变异快<sup>[5]</sup>。正因为如此, 动物中广泛使用线粒体

基金项目: 国家自然科学基金(31872162); 山西省自然科学基金(201601D011065); 山西省回国留学人员科研资助项目(2017-015)

\*通信作者。张姝, Tel: +86-351-7017101, E-mail: zhangshu@sxu.edu.cn; 张永杰, E-mail: zhangyj2008@sxu.edu.cn

收稿日期: 2019-01-25; 修回日期: 2019-03-27; 网络出版日期: 2019-07-11

*COI* 基因片段作为物种鉴定的通用 DNA 条形码<sup>[6]</sup>。

真菌属于低等真核生物, 其遗传物质分布于细胞核和线粒体中。近年来, 随着测序成本的降低, 有越来越多真菌的细胞核基因组和线粒体基因组被测序。然而, 关于真菌线粒体 DNA 和核 DNA 进化速率比较的研究还较少。在环孢菌素产生菌膨大弯颈霉 *Tolypocladium inflatum* 中, 基于全基因组序列的比较研究发现, 核基因组的 SNP 频率(0.97%)是线粒体基因组 SNP 频率(0.12%)的 8 倍<sup>[7]</sup>。然而, 在人体致病菌耶氏肺孢子菌 *Pneumocystis jirovecii* 中, 线粒体基因 *cob* 和 *rnl* 显示出比细胞核超氧化物歧化酶基因 *SOD* 更高的多态位点比例(1.57% vs. 0.39%)和更多的基因型(14–22 vs. 4)<sup>[8]</sup>。为了详细了解真菌中线粒体 DNA 与核 DNA 的进化速率, 需对更多真菌开展类似的比较研究。

真核生物体内线粒体基因组与细胞核基因组间可能存在一定程度的基因交流。核内线粒体 DNA 片段(nuclear mitochondrial DNA segments, 简称 Numts), 也叫线粒体假基因(mitochondrial pseudogenes), 是指从线粒体基因组转移并整合到细胞核基因组的 DNA 片段。在线粒体进化的早期, 从线粒体到细胞核的 DNA 转移曾是大规模发生的事件。在现代生物中仍可能检测到这种 DNA 转移的证据, 但在不同生物中, Numts 的数量和长度差别较大<sup>[9–10]</sup>。例如, 小麦颖枯病原颖枯壳针孢 *Phaeosphaeria nodorum* 的 Numts 含量是 77 kb, 而在黑曲霉 *Aspergillus niger* 中只有 298 bp<sup>[10]</sup>。

蛹虫草 *Cordyceps militaris* (L.) Fr. 是虫草属的模式种, 主要寄生阔叶林或混交林地地表土层中鳞翅目昆虫的蛹或幼虫。由于具有较高的医疗保健功效和经济价值, 科研人员近些年从基因组<sup>[11]</sup>、转录组<sup>[12]</sup>、甲基化组<sup>[13]</sup>、蛋白质组<sup>[12]</sup>和代谢网

络<sup>[14]</sup>等多个角度对蛹虫草作了深入研究。近年来, 本研究团队开展了蛹虫草的种内比较线粒体基因组学研究, 发现由于内含子数目的差异(2–8 个), 不同蛹虫草菌株的线粒体基因组大小存在较大变化, 从 26.5 kb 到 33.9 kb, 相差约 7 kb<sup>[15]</sup>。进一步研究发现, 在已知的 8 个内含子位点都具有内含子的基因型为蛹虫草较古老的线粒体基因型; 在蛹虫草种内分化过程中, 总体趋势是丢失线粒体内含子, 但部分内含子可以发生“失而复得”和“得而再失”<sup>[16]</sup>。

自蛹虫草菌株 CM01 的基因组通过二代测序技术完成测序后<sup>[11]</sup>, 科研人员又使用三代测序技术完成了对菌株 01 和 ATCC 34164 的基因组测序, 尤其是菌株 ATCC 34164 的基因组被组装到了染色体水平(7 条染色体)<sup>[17]</sup>。这些核基因组数据为我们研究蛹虫草中的 Numts 提供了便利。另外, 之前在筛选适于蛹虫草遗传多样性分析的线粒体分子标记时, 我们比较了 12 个线粒体 DNA 片段与 3 个核 DNA 片段(ITS、交配型基因 *MATI-1-1* 和 *MATI-2-1*)的序列, 发现这些线粒体 DNA 片段比核 DNA 片段进化快<sup>[18]</sup>。但之前研究中参与比较的核 DNA 片段较少, 且有研究显示核糖体 DNA 信息位点的比例比蛋白编码基因少<sup>[19]</sup>, 因此, 我们尚不能就此认定蛹虫草线粒体 DNA 比核 DNA 变异快。为了比较蛹虫草线粒体 DNA 与核 DNA 的进化关系, 为真菌线粒体与核 DNA 进化关系比较提供更多例证, 本研究首先分析蛹虫草核基因组中的 Numts, 然后利用已知线粒体基因组的 10 个蛹虫草菌株, 比较核 DNA 与线粒体 DNA 在多个蛋白编码基因上的变异速率及所反映的蛹虫草菌株间的系统发育关系。本文的研究结果将丰富我们对蛹虫草线粒体与细胞核 DNA 进化关系的认识。

# 1 材料和方法

## 1.1 菌株培养与 DNA 提取

本研究共使用 10 株蛹虫草菌株(表 1), 均由本实验室保藏。将这些菌株按照张永杰等<sup>[18]</sup>所描述的方法进行培养、菌丝体收集和总 DNA 提取。

## 1.2 线粒体基因组内部重复序列的分析

在目前已知线粒体基因组的蛹虫草菌株中, 菌株 V40-5 因具有 8 个内含子而代表了蛹虫草最大的线粒体基因组(33.9 kb)<sup>[15]</sup>。为了分析蛹虫草线粒体基因组内部不同区域间是否存在较高的序列相似性, 将菌株 V40-5 的线粒体基因组与其自身进行本地 BLASTn 分析<sup>[20]</sup>, 筛选和保留  $E$  值小于  $10^{-5}$  的比对结果。通过 SSRIT (<http://archive.gramene.org/db/markers/ssrtool>) 分析蛹虫草线粒体基因组中的简单序列重复(SSR), 设置最大基序(motif)长度为 10, 最小重复次数为 5。

## 1.3 核内线粒体 DNA 片段的检测

目前, GenBank 数据库中共有 3 个蛹虫草菌株(CM01、01 和 ATCC34164)的基因组组装数据。将这 3 个菌株的基因组组装数据分别下载到本地电脑, 用菌株 V40-5 的线粒体基因组与其分别进行本地 BLASTn<sup>[20]</sup>和 LAST 搜索<sup>[21]</sup>, 筛选和保留  $E$  值小于  $10^{-5}$  的匹配结果。

## 1.4 核 DNA 片段的扩增与测序

从公布的蛹虫草 CM01 菌株的基因组组装数据中随机选择 7 个核基因(分散在染色体 IV 和 VII 上, BLAST 显示全部为单拷贝基因), 利用在线程序 Primer 3 (<http://primer3.ut.ee/>)设计扩增这 7 个基因的引物(表 2)。所设计的引物由金唯智生物科技有限公司(天津, 中国)合成。从表 1 所列的蛹虫草菌株中扩增这些核 DNA 片段。扩增产物由金唯智生物科技有限公司(天津, 中国)利用 Sanger 法进行双向测序, 测序结果使用软件 Vector NTI Suite (Invitrogen, 美国)中的 ContigExpress 进行拼接。

表 1. 蛹虫草菌株核 DNA 片段的序列登录号

Table 1. Accession numbers of nuclear DNA fragments from various *C. militaris* isolates

Isolate	Accession number						
	AD	GD	TA	CP	CS	EP	RF
V26-17	MK299434	MK299474	MK299493	MK299444	MK299454	MK299464	MK299485
V40-4	MK299437	MK299477	MK299496	MK299447	MK299457	MK299467	MK299487
V40-5	MK299438	MK299478	MK299497	MK299448	MK299458	MK299470	MK299489
CM09-9-24	MK299435	MK299475	MK299494	MK299445	MK299455	MK299465	MK299486
CM09-31-28	MK299436	MK299476	MK299495	MK299446	MK299456	MK299466	MK299490
CM552	MK299439	MK299479	MK299498	MK299449	MK299459	MK299468	–
CM01	MK299432	MK299472	MK299491	MK299442	MK299452	MK299462	MK299482
CM06	MK299433	MK299473	MK299492	MK299443	MK299453	MK299463	MK299484
F02	MK299441	MK299481	MK299500	MK299450	MK299461	MK299471	MK299488
CMB	MK299440	MK299480	MK299499	MK299451	MK299460	MK299469	MK299483

Refer to Zhang *et al.*<sup>[18]</sup> for sources of these isolates and their mitogenome accessions. It's noted that we failed to amplify the RF fragment from CM552.

表 2. 扩增核基因片段的引物

Table 2. Primers used for amplifying nuclear DNA fragments

Fragment <sup>a</sup>	Primer name	Primer sequence (5'→3')	Annealing temp./°C	Amplicon/bp	Target locus <sup>b</sup>	Chromosome <sup>c</sup>
AD	AD-F	GCCAGGCCCTCTCTAAATCT	53	1508	CCM_05174	IV
	AD-R	GCTGCAACGGACAAACTCAT				
CP	CP-F	ATCATGCCCTCCACCACTAC	53	1421	CCM_00072	VII
	CP-R	GCTGCGCTTCTTTTCCTTCT				
CS	CS-F	CCAATACCGCACCTACAAGC	52	1386	CCM_05104	IV
	CS-R	AGGCGTTGCTGTGGATAGAT				
EP	EP-F	TAGGAGCAAGCACAGACGTT	53	1606	CCM_05090	IV
	EP-R	ATTCAGCCGTCCCACAAAC				
GD	GD-F	GCACTTCCAGCCTGAGTTG	53	1854	CCM_07752	VII
	GD-R	GAGTCTGCCACCAAACGTTTC				
RF	RF-F	AACCCTGTTTGAAGTTCTGC	53	1570	CCM_00093	VII
	RF-R	TCACCACTCTCTCCACCAAC				
TA	TA-F	CCAACATCTTCCCTCCACAC	53	1862	CCM_00163	VII
	TA-R	GAGTCGGTGCAGTTCTCGAA				

<sup>a</sup>These nuclear genes encoded for ATP-dependent DNA helicase recQ (AD), cutinase palindrome-binding protein (CP), cell surface receptor/MFS transporter (CS), endonuclease/exonuclease/phosphatase family protein (EP), guanosine-diphosphatase (GD), RING finger domain-containing protein (RF), and threonine ammonia-lyase precursor (TA). <sup>b</sup>Target locus means in which gene locus each corresponding nuclear fragment is positioned in the CM01 reference genome by referring to Zheng *et al*<sup>[11]</sup>. <sup>c</sup>Chromosome means in which chromosome each corresponding nuclear fragment is positioned by referring to Kramer & Nodwell<sup>[17]</sup>.

### 1.5 线粒体 DNA 与核 DNA 变异频率的比较

本研究使用的多数核 DNA 片段和少数线粒体 DNA 片段中有内含子。有研究显示内含子与外显子序列的进化速率不同<sup>[15]</sup>。因此,为使线粒体 DNA 与核 DNA 的比较更加客观,本研究只利用外显子序列进行后面的分析。用 MEGA 6.06 软件<sup>[22]</sup>中的 Muscle 程序对获得的不同蛹虫草菌株的各个核 DNA 片段进行序列比对,用 DnaSP 5.10<sup>[23]</sup>软件统计各 DNA 片段上存在的变异碱基数、等位基因数和单元型多样性指数,用 MEGA 6.06<sup>[22]</sup>中的 K2P 模型计算菌株间的平均遗传距离和最大遗传距离。利用同样的方法统计不同蛹虫草菌株在各个线粒体 DNA 片段上的碱基变异情

况。用软件 TOPALI 2.5<sup>[24]</sup>将不同的核 DNA 片段拼接在一起,将不同的线粒体 DNA 片段拼接在一起,用前述同样方法统计和比较线粒体 DNA 与核 DNA 总体的碱基变异情况。

### 1.6 线粒体 DNA 与核 DNA 系统发育关系的比较

首先,利用 DnaSP 5.10 软件<sup>[23]</sup>对各核 DNA 片段和各线粒体 DNA 片段进行中性检验(包括 Tajima's D, Fu and Li's D & F test)。然后,对符合中性假说的各核 DNA 片段间用 PAUP 4.0b10 软件 (Sinauer Associates, 美国)进行分区同质性检验(以  $P < 0.001$  作为判断是否显著冲突的阈值),将不存在冲突的所有核 DNA 片段拼接成核 DNA 数据集。同理,将不存在冲突的所有线粒体 DNA 片段拼接

成线粒体 DNA 数据集。最后,用 PAUP 4.0b10 软件构建最大简约树(MP),用 TreeMap 3b1243 (<https://sites.google.com/site/cophylogeny>) 将线粒体 DNA 树和核 DNA 树关联起来;通过 KH (Kishino-Hasegawa)和 SH (Shimodaira-Hasegawa) 检验(在 PAUP 软件中运算)明确线粒体 DNA 树和核 DNA 树之间是否存在显著的拓扑结构差异。

## 2 结果和分析

### 2.1 线粒体基因组内部的重复序列

将蛹虫草 V40-5 菌株的线粒体基因组与其自身进行本地 BLASTn 搜索,共输出 22 条具有较高相似性的比对结果( $E$  值  $<7E-6$ , 相似性 73.4%–96.4%, 比对长度 28–259 bp; 表 3), 排除冗余后实际只有 13 条比对结果。其中, 比对长度大于 100 bp 的仅有 3 条, 分别对应 *rnl-i1* 非编码区和 *rnl-i2* 非编码区(比对长度 259 bp, 相似性 88.4%)、*rnl-i1* 内的 *orf245* 和 *rnl-i2* 内的 *orf253* (171–174 bp, 76.6%–77.0%)、*rnl-i1* 内的 *orf245* 和 *cox2-i1* 内的 *orf286* (172 bp, 73.8%)。此外, 在蛹虫草线粒体基因组中检测到 4 个 SSR, 分别位于 *rnl*、*rns*、*cox1* 和 *nad1* 的外显子序列中, 其中 TA、AT、TAT 和 TAA 分别重复 5 次。

### 2.2 核内线粒体 DNA 片段的检测

将蛹虫草 V40-5 菌株的线粒体基因组与公布的 3 个蛹虫草菌株(CM01、01 和 ATCC34164)的核基因组序列分别进行 BLASTn 搜索, 只输出 2–4 条比对长度为 38–66 bp 的结果(相似性 87.2%–98.5%; 表 3), 进行 LAST 搜索的结果只比 BLASTn 多找出 1 个比对结果(表 4)。这些结果说明, 蛹虫草核基因组中的 Numts 数量有限(最多 5 条), 总长只有 278 bp。蛹虫草线粒体与细胞核

基因组间已检测不到长片段的基因交流事件。

### 2.3 线粒体与细胞核 DNA 变异频率的比较

除从 CM552 菌株中未扩增到核 DNA 片段 RF 外, 我们从所有蛹虫草菌株中扩增到了本研究涉及的所有 DNA 片段(7 个核 DNA 片段和 14 个线粒体 DNA 片段)。这些片段在外显子序列上均不存在碱基的插入缺失现象。不同核 DNA 片段或线粒体 DNA 片段的碱基变异程度不同(表 5)。供试菌株在 7 个核 DNA 片段上均表现出不同程度的遗传分化, 而在 14 个线粒体 DNA 片段中, *atp8*、*atp9* 和 *nad6* 这 3 个基因非常保守(无碱基变化), 其他线粒体 DNA 片段均存在碱基变化。整体而言, 核 DNA 的变异频率略高于线粒体 DNA (SNP 频率 1.24% vs. 0.71%)。核 DNA 片段在遗传距离、等位基因数及单元型多样性指数等指标上也比多数线粒体 DNA 片段高。

### 2.4 线粒体与细胞核 DNA 系统发育关系的比较

根据中性检验的结果, 线粒体 DNA 片段 *cox3* 和 *nad4L* 不符合中性假说( $P < 0.05$ ), 而其他的线粒体 DNA 片段和所有的核 DNA 片段都符合中性假说。根据分区同质性检验的结果, 核 DNA 片段 EP 与其他核 DNA 片段间存在明显冲突( $P < 0.001$ ), 剩余的核 DNA 片段间及符合中性假说的 12 个线粒体 DNA 片段间均不存在冲突。同时, 又由于有 1 个菌株(即 CM552)缺少核 DNA 片段 RF 的序列, 因此, 我们将除 EP 和 RF 外的 5 个核 DNA 片段组成最终的核 DNA 数据集(总长 6990 bp), 将 12 个线粒体 DNA 片段组成最终的线粒体 DNA 数据集(总长 11946 bp)。用这两个数据集分别构建 MP 系统发育树, 其拓扑结构存在差异(图 1), 进一步通过 KH 和 SH 检验发现这种差异达到极显著水平( $P < 0.0001$ )。

表 3. 蛹虫草 V40-5 线粒体基因组自身及与 3 个核基因组本地 BLAST 搜索的结果

Table 3. Local BLAST analyses of the *C. militaris* V40-5 mitogenome against itself and against three nuclear genomes

Query start	Query end	Sequence ID	Target strain	Subject length/bp	Subject start	Subject end	Subject strand	Alignment length/bp	Percent identity/%	Bit score	<i>E</i> value
V40-5 mitogenome BLAST against itself											
2063	2319	KP722501	V40-5	33967	717	973	+	259	88.42	309.0	8.00E-85
717	973	KP722501	V40-5	33967	2063	2319	+	259	88.42	309.0	8.00E-85
2629	2796	KP722501	V40-5	33967	1259	1429	+	174	77.01	91.6	3.00E-19
1259	1429	KP722501	V40-5	33967	2629	2796	+	171	76.61	91.6	3.00E-19
30227	30290	KP722501	V40-5	33967	29291	29354	+	64	90.62	86.1	1.00E-17
29291	29354	KP722501	V40-5	33967	30227	30290	+	64	90.62	86.1	1.00E-17
12735	12902	KP722501	V40-5	33967	1259	1429	+	172	73.84	63.9	7.00E-11
22558	22622	KP722501	V40-5	33967	19453	19519	+	69	84.06	62.1	2.00E-10
19453	19519	KP722501	V40-5	33967	22558	22622	+	69	84.06	62.1	2.00E-10
32092	32139	KP722501	V40-5	33967	7539	7585	+	48	87.50	54.7	4.00E-08
7539	7585	KP722501	V40-5	33967	32092	32139	+	48	87.50	54.7	4.00E-08
29296	29347	KP722501	V40-5	33967	22495	22546	+	52	84.62	52.8	1.00E-07
22495	22546	KP722501	V40-5	33967	29296	29347	+	52	84.62	52.8	1.00E-07
12808	12861	KP722501	V40-5	33967	2702	2755	+	54	83.33	51.0	5.00E-07
2702	2755	KP722501	V40-5	33967	12808	12861	+	54	83.33	51.0	5.00E-07
22572	22602	KP722501	V40-5	33967	5677	5708	+	32	93.75	47.3	7.00E-06
19468	19499	KP722501	V40-5	33967	5715	5686	-	32	93.75	47.3	7.00E-06
23086	23113	KP722501	V40-5	33967	5720	5693	-	28	96.43	47.3	7.00E-06
30232	30283	KP722501	V40-5	33967	22495	22546	+	52	82.69	47.3	7.00E-06
5693	5720	KP722501	V40-5	33967	23113	23086	-	28	96.43	47.3	7.00E-06
22495	22546	KP722501	V40-5	33967	30232	30283	+	52	82.69	47.3	7.00E-06
21286	21341	KP722501	V40-5	33967	31827	31881	+	57	82.46	47.3	7.00E-06
V40-5 mitogenome BLAST against nuclear genomes											
31680	31726	CP023324	ATCC 34164	8286469	4750747	4750701	-	47	87.23	54.7	1.00E-05
21881	21946	CP023326	ATCC 34164	3849251	2877664	2877729	+	66	98.48	117.0	6.00E-25
9142	9184	MQTM01000005	01	4088212	3848337	3848295	-	43	93.02	63.9	8.00E-09
6431	6468	MQTM01000010	01	1886322	479668	479631	-	38	97.37	65.8	1.00E-09
31680	31726	MQTM01000012	01	1944127	1819799	1819753	-	47	87.23	54.7	2.00E-06
6431	6468	NW_006271974	CM01	3240361	2723649	2723686	+	38	97.37	65.8	2.00E-09
9142	9184	NW_006271970	CM01	5422368	5176892	5176850	-	43	93.02	63.9	1.00E-08
21881	21940	NW_006271975	CM01	2818147	1095496	1095439	-	60	93.33	87.9	3.00E-16
31680	31726	NW_006271973	CM01	3260156	659497	659543	+	47	87.23	54.7	4.00E-06

When performing BLAST analyses against nuclear genomes, three target strains (CM01, 01 and ATCC 34164), which have available nuclear genomes, were employed as listed in the "Target strain" column.

表 4. 蛹虫草 V40-5 线粒体基因组与 3 个核基因组的 LAST 搜索

Table 4. LAST searches of the *C. militaris* V40-5 mitogenome against three nuclear genomes

Query start	Query end	Query position	Subject sequence ID	Subject strain	Subject start	Subject end	Subject position	Alignment length/bp	Percent identity/%	Bit score	<i>E</i> value
21881	21946	<i>cox1</i> -E2	CP023326	ATCC 34164	2877664	2877729	chr III	66	98.48	103	2.60E-19
31748	31680	<i>cox3</i> -i1	CP023324	ATCC 34164	4750679	4750747	chr VII	69	76.81	60.1	1.90E-06
11326	11265	<i>nad3</i>	CP023325	ATCC 34164	3508316	3508377	chr V	62	77.42	55.3	5.00E-05
11265	11326	<i>nad3</i>	MQTM01000013	01	617828	617889	Contig 10	62	82.26	64.8	6.80E-08
31748	31680	<i>cox3</i> -i1	MQTM01000012	01	1819731	1819799	Contig 7	69	78.26	63.2	2.00E-07
9184	9142	<i>trnQ</i>	MQTM01000005	01	3848295	3848337	Contig 2	43	93.02	60.1	1.80E-06
6468	6431	<i>rnl</i> -i4	MQTM01000010	01	479631	479668	Contig 8	38	97.37	58.5	5.40E-06
11265	11326	<i>nad3</i>	NW_006271969	CM01	5079784	5079845	CCM_S00001	62	82.26	64.8	6.70E-08
31680	31748	<i>cox3</i> -i1	NW_006271973	CM01	659497	659565	CCM_S00005	69	78.26	63.2	2.00E-07
6431	6468	<i>rnl</i> -i4	NW_006271974	CM01	2723649	2723686	CCM_S00006	38	97.37	58.5	5.40E-06
21946	21881	<i>cox1</i> -E2	NW_006271975	CM01	1095433	1095496	CCM_S00007	66	90.91	75.9	3.10E-11
9184	9142	<i>trnQ</i>	NW_006271970	CM01	5176850	5176892	CCM_S00002	43	93.02	60.1	1.80E-06

表 5. 不同蛹虫草菌株在各 DNA 片段上的碱基变异情况

Table 5. Nucleotide variations on each DNA fragment among different *C. militaris* isolates

Locus	Length/bp	Variable sites				Overall distance	Max. distance	Allele	Hd
		Pi	S	SNP	%				
Nuclear loci									
AD	1198	5	7	12	1.00	0.0034	0.0076	5 (4)	0.80
CP	1224	5	0	5	0.41	0.0020	0.0041	5 (5)	0.76
CS	1226	19	17	36	2.94	0.0100	0.0150	7 (4)	0.87
EP	1384	14	7	21	1.52	0.0057	0.0096	7 (3)	0.91
GD	1679	9	3	12	0.71	0.0027	0.0042	6 (3)	0.89
RF	1435	2	5	7	0.49	0.0014	0.0028	5 (4)	0.81
TA	1663	11	18	29	1.74	0.0055	0.0092	6 (4)	0.84
Sum (nr)	9809	65	57	122	1.24	–	–	–	–
Mitochondrial loci									
<i>atp6</i>	783	0	4	4	0.51	0.0010	0.0039	4 (7)	0.53
<i>atp8</i>	147	0	0	0	0.00	0.0000	0.0000	1 (10)	0.00
<i>atp9</i>	225	0	0	0	0.00	0.0000	0.0000	1 (10)	0.00
<i>cob</i>	1161	4	4	8	0.69	0.0026	0.0061	5 (4)	0.80
<i>cox1</i>	1593	4	5	9	0.56	0.0018	0.0044	6 (5)	0.78
<i>cox2</i>	750	7	2	9	1.20	0.0055	0.0095	6 (4)	0.84
<i>cox3</i>	810	0	11	11	1.36	0.0028	0.0139	2 (9)	0.20
<i>nad1</i>	1104	3	3	6	0.54	0.0020	0.0036	4 (5)	0.71
<i>nad2</i>	1686	3	7	10	0.59	0.0017	0.0054	5 (4)	0.82
<i>nad3</i>	420	1	0	1	0.24	0.0013	0.0024	2 (5)	0.56
<i>nad4</i>	1449	2	7	9	0.62	0.0017	0.0049	4 (4)	0.73
<i>nad4L</i>	270	3	0	3	1.11	0.0062	0.0114	3 (5)	0.64
<i>nad5</i>	1995	12	10	22	1.10	0.0043	0.0086	4 (5)	0.71
<i>nad6</i>	633	0	0	0	0.00	0.0000	0.0000	1 (10)	0.00
Sum (mt)	13026	39	53	92	0.71	0.0024	0.0046	7 (4)	0.87

Pi: parsimony informative sites; S: singleton sites; SNP: single nucleotide polymorphism sites (=Pi+S); Hd: haplotype diversity. For the allele column, we listed the number of total alleles (outside parentheses) and the number of individuals represented by the dominant allele (within parentheses). For the nuclear locus RF, all data were from 9 isolates; for all other loci, the data were from 10 isolates. In order to make the comparison between mitochondrial DNA fragments and nuclear DNA fragments more objective, this study only used the exon sequences although most nuclear fragments and some mitochondrial fragments contained introns.

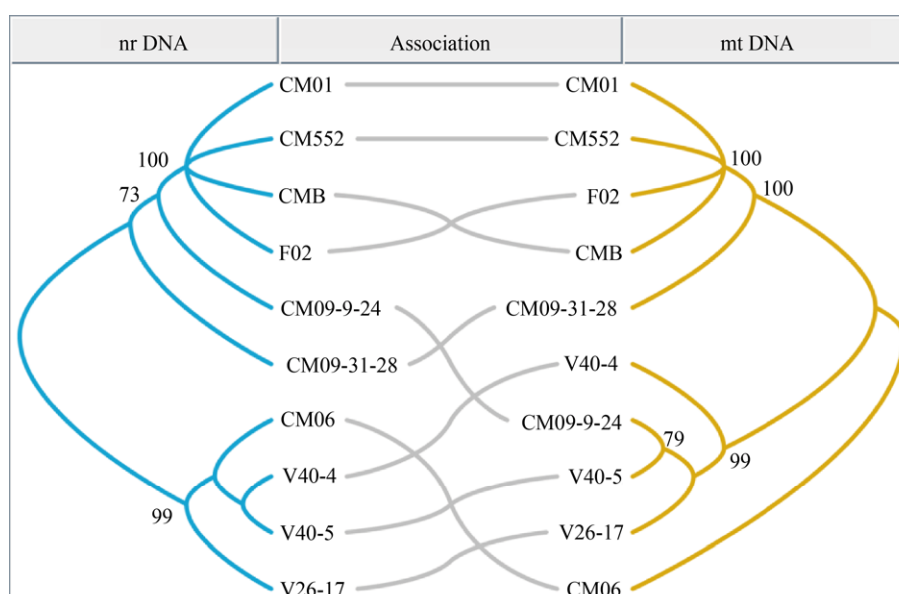


图 1. 核 DNA(左)与线粒体 DNA(右)系统发育树拓扑结构的比较

Figure 1. Comparison between nuclear DNA (nr DNA, left) and mitochondrial DNA (mt DNA, right) phylogenies. Bootstrap values for nodes receiving a bootstrap value larger than 70% were marked.

### 3 讨论

线粒体被普遍认为起源于寄生在真核生物共同祖先细胞内的远古细菌(很可能是某种  $\alpha$ -变形菌)。与自由生活的  $\alpha$ -变形菌相比,线粒体基因组只编码少数功能基因,而多数功能基因可能转移到了细胞核基因组内或被其他具有类似功能的核基因所代替<sup>[25]</sup>。Hazkani-Covo *et al.* (2010) 分析了 85 种真核生物中的核内线粒体 DNA 片段 (Numts), 发现不同生物中 Numts 的数量和长度存在较大差异<sup>[10]</sup>。我们首次检测了蛹虫草中的 Numts, 仅检测到 5 个较短的 Numts (单个 Numt < 70 bp, 总长 278 bp), 没有发现线粒体与核基因组间长片段的基因交流事件。然而, 在另外一些真菌中仍能检测到线粒体与细胞核基因组间长片段的基因交流, 如在膨大弯颈霉中存在 1 条 652 bp 的 Numt<sup>[7]</sup>, 在肺囊康定产生菌 *Glarea*

*lozoyeneis* 中存在 4 条长度超过 388 bp 的 Numts (最长的达到 3.9 kb)<sup>[26]</sup>。同时, 本研究还发现不同蛹虫草菌株中的 Numts 数量存在差异(表 4), 这可能是由于某些 Numts 在插入核基因组后发生了突变或丢失而无法被检测到。

在真核生物线粒体和细胞核内存在的这两套基因组的变异情况如何一直是人们感兴趣的话题。在目前研究过的多数植物中, 线粒体 DNA 的变异速率比核 DNA 慢, 而在多数动物中, 线粒体 DNA 的变异速率比核 DNA 快<sup>[5]</sup>。真菌中开展过类似比较的研究还不多见。最近, 通过系统总结涉及真菌线粒体 DNA 与核 DNA 进化速率比较的文献, 我们发现与植物类似, 多数真菌具有较慢的线粒体 DNA 进化速率和较快的核 DNA 进化速率<sup>[27]</sup>。本研究关于蛹虫草的研究结果与此一致。通过对多个蛋白编码基因外显子区的序列比较, 我们发现蛹虫草核 DNA 的变异速率是线粒



体 DNA 变异速率的 1.7 倍(SNP 频率 1.24% vs. 0.71%)。

之前在筛选适于蛹虫草遗传多样性分析的线粒体分子标记时,我们曾报道线粒体 DNA 比参与比较的核 DNA 进化快<sup>[18]</sup>。造成该结果与本文结论差异的原因可能是由于之前研究中选用的核 DNA 片段(ITS、交配型基因 *MATI-1-1* 和 *MATI-2-1*)恰好较保守。由于蛹虫草为异宗配合真菌,一个菌株通常只有 *MATI-1-1* 或 *MATI-2-1*,因此本研究筛选核 DNA 片段时没有选用交配型基因。本文只对线粒体和核基因组中的蛋白编码基因作了比较,核糖体 RNA 和非编码区是否也是同样的结果仍有待以后深入研究。

本研究发现线粒体 DNA 和核 DNA 反映的不同蛹虫草菌株间的系统发育关系并不完全一致。类似的现象在其他真菌中亦有报道<sup>[28]</sup>。造成这种差异的原因可能主要是由于线粒体 DNA 与核 DNA 的遗传机制不同。细胞核 DNA 遵守典型的孟德尔遗传规律,可通过有性繁殖进行遗传重组,而线粒体 DNA 不遵守典型的孟德尔遗传规律,以克隆繁殖和单亲遗传为主<sup>[29]</sup>。线粒体 DNA 与核 DNA 的差异还表现在一个细胞内通常只有一个或少数几个核基因组拷贝(取决于倍性及细胞核的数目),却有大量的线粒体基因组拷贝。核 DNA 的复制和分裂与细胞有丝分裂同步,经过复制的核基因组受到严密的调控而平均分配给子细胞,然而线粒体的复制和分裂与细胞有丝分裂并不同步,线粒体 DNA 随机分配给子细胞。

## 参 考 文 献

[1] Calderone R, Li DM, Traven A. System-level impact of mitochondria on fungal virulence: to metabolism and beyond.

*FEMS Yeast Research*, 2015, 15(4): fov027.

- [2] Chatre L, Ricchetti M. Are mitochondria the Achilles' heel of the Kingdom Fungi? *Current Opinion in Microbiology*, 2014, 20: 49–54.
- [3] Galtier N, Nabholz B, Glémin S, Hurst GDD. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology*, 2009, 18(22): 4541–4550.
- [4] Gray MW. Mitochondrial evolution. *Cold Spring Harbor Perspectives in Biology*, 2012, 4(9): a011403.
- [5] Lynch M, Koskella B, Schaack S. Mutation pressure and the evolution of organelle genomic architecture. *Science*, 2006, 311(5768): 1727–1730.
- [6] Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 2003, 270(1512): 313–321.
- [7] Zhang YJ, Yang XQ, Zhang S, Humber RA, Xu JP. Genomic analyses reveal low mitochondrial and high nuclear diversity in the cyclosporin-producing fungus *Tolypocladium inflatum*. *Applied Microbiology and Biotechnology*, 2017, 101(23/24): 8517–8531.
- [8] Charpentier E, Garnaud C, Wintenberger C, Bailly S, Murat JB, Rendu J, Pavese P, Drouet T, Augier C, Malvezzi P, Thiébaud-Bertrand A, Mallaret MR, Epaulard O, Cornet M, Larrat S, Maubon D. Added value of next-generation sequencing for multilocus sequence typing analysis of a *Pneumocystis jirovecii* pneumonia outbreak. *Emerging Infectious Diseases*, 2017, 23(8): 1237–1245.
- [9] Richly E, Leister D. NUMTs in sequenced eukaryotic genomes. *Molecular Biology and Evolution*, 2004, 21(6): 1081–1084.
- [10] Hazkani-Covo E, Zeller RM, Martin W. Molecular poltergeists: mitochondrial DNA copies (*numts*) in sequenced nuclear genomes. *PLoS Genetics*, 2010, 6(2): e1000834.
- [11] Zheng P, Xia YL, Xiao GH, Xiong CH, Hu X, Zhang SW, Zheng HJ, Huang Y, Zhou Y, Wang SY, Zhao GP, Liu XZ, St

- Leger RJ, Wang CS. Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional Chinese medicine. *Genome Biology*, 2011, 12(11): R116.
- [12] Yin YL, Yu GJ, Chen YJ, Jiang S, Wang M, Jin YX, Lan XQ, Liang Y, Sun H. Genome-wide transcriptome and proteome analysis on different developmental stages of *Cordyceps militaris*. *PLoS One*, 2012, 7(12): e51853.
- [13] Wang YL, Wang ZX, Liu C, Wang SB, Huang B. Genome-wide analysis of DNA methylation in the sexual stage of the insect pathogenic fungus *Cordyceps militaris*. *Fungal Biology*, 2015, 119(12): 1246–1254.
- [14] Vongsangnak W, Raethong N, Mujchariyakul W, Nguyen NN, Leong HW, Laoteng K. Genome-scale metabolic network of *Cordyceps militaris* useful for comparative analysis of entomopathogenic fungi. *Gene*, 2017, 626: 132–139.
- [15] Zhang S, Hao AJ, Zhao YX, Zhang XY, Zhang YJ. Comparative mitochondrial genomics toward exploring molecular markers in the medicinal fungus *Cordyceps militaris*. *Scientific Reports*, 2017, 7: 40219.
- [16] 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.
- [17] Kramer GJ, Nodwell JR. Chromosome level assembly and secondary metabolite potential of the parasitic fungus *Cordyceps militaris*. *BMC Genomics*, 2017, 18(1): 912.
- [18] Zhang YJ, Guo LH, Zhang S, Liu XZ. Determining mitochondrial molecular markers suitable for genetic diversity analysis of *Cordyceps militaris*. *Acta Microbiologica Sinica*, 2015, 55(7): 826–833. (in Chinese)  
张永杰, 郭丽宏, 张姝, 刘杏忠. 适于蛹虫草遗传多样性研究的线粒体分子标记的筛选. *微生物学报*, 2015, 55(7): 826–833.
- [19] Liu XZ, Wang QM, Theelen B, Groenewald M, Bai FY, Boekhout T. Phylogeny of tremellomycetous yeasts and related dimorphic and filamentous basidiomycetes reconstructed from multiple gene sequence analyses. *Studies in Mycology*, 2015, 81: 1–26.
- [20] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology*, 1990, 215(3): 403–410.
- [21] Kielbasa SM, Wan R, Sato K, Horton P, Frith MC. Adaptive seeds tame genomic sequence comparison. *Genome Research*, 2011, 21(3): 487–493.
- [22] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 2013, 30(12): 2725–2529.
- [23] Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 2009, 25(11): 1451–1452.
- [24] Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, Marshall DF, Wright F. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, 2009, 25(1): 126–127.
- [25] Adams KL, Palmer JD. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Molecular Phylogenetics and Evolution*, 2003, 29(3): 380–395.
- [26] Zhang YJ, Zhao YX, Zhang S, Chen L, Liu XZ. Reanalysis of the mitochondrial genome of the pneumocandin-producing fungus *Glarea lozoyensis*. *Acta Microbiologica Sinica*, 2017, 57(5): 724–736.
- [27] Sandor S, Zhang YJ, Xu JP. Fungal mitochondrial genomes and genetic polymorphisms. *Applied Microbiology and Biotechnology*, 2018, 102(22): 9433–9448.
- [28] Turissini DA, Gomez OM, Teixeira MM, McEwen JG, Matute DR. Species boundaries in the human pathogen *Paracoccidioides*. *Fungal Genetics and Biology*, 2017, 106: 9–25.
- [29] Xu JP, Li H. Current perspectives on mitochondrial inheritance in fungi. *Cell Health and Cytoskeleton*, 2015, 7: 143–154.

# Comparison of evolutionary relationships between mitochondrial and nuclear DNAs in *Cordyceps militaris*

Shu Zhang<sup>1\*</sup>, Ningbo Cui<sup>2</sup>, Yuxiang Zhao<sup>2</sup>, Yongjie Zhang<sup>2\*</sup>

<sup>1</sup>Institute of Applied Chemistry, Shanxi University, Taiyuan 030006, Shanxi Province, China

<sup>2</sup>School of Life Sciences, Shanxi University, Taiyuan 030006, Shanxi Province, China

**Abstract:** [Objective] To detect possible nuclear mitochondrial DNA segments (Numts) within nuclear genomes of *Cordyceps militaris* and to compare the degree of nucleotide variations and phylogenetic relationships between mitochondrial and nuclear DNAs by employing multiple *C. militaris* isolates. [Methods] Nuclear genomes of *C. militaris* were searched for sequence similarities by local BLAST/LAST analyses with its mitochondrial genome as the query. Seven nuclear protein-coding gene fragments were amplified from 10 *C. militaris* isolates that had available mitochondrial genomes. Nucleotide variations at seven nuclear fragments were compared with those at 14 mitochondrial protein-coding genes. [Results] Five short Numts with a total length of 278 bp were identified in the nuclear genomes of *C. militaris*. The overall nucleotide variabilities of *C. militaris* on nuclear fragments were generally higher than that on mitochondrial fragments. The phylogenetic relationships among different *C. militaris* isolates revealed by nuclear and mitochondrial DNAs were significantly different. [Conclusion] There was no detectable gene transfer of long fragments between mitochondrial and nuclear genomes of *C. militaris*. The nucleotide variabilities of mitochondrial and nuclear DNAs were different, and the phylogenetic relationships of different *C. militaris* isolates revealed by them were different either. This study enhanced our understanding of the evolutionary relationship between mitochondrial and nuclear DNAs in *C. militaris*.

**Keywords:** *Cordyceps militaris*, mitochondrial DNA, nuclear DNA, gene transfer, nucleotide variability

(本文责编: 李磊)

Supported by the National Natural Science Foundation of China (31872162), by the Natural Science Foundation of Shanxi Province (201601D011065), and by the Research Project Supported by Shanxi Scholarship Council of China (2017-015)

\*Corresponding author. Shu Zhang, Tel: +86-351-7017101, E-mail: zhangshu@sxu.edu.cn; Yongjie Zhang, E-mail: zhangyj2008@sxu.edu.cn

Received: 25 January 2019; Revised: 27 March 2019; Published online: 11 July 2019