

枸杞属植物 nrDNA ITS 致同进化不完全

张 娇¹, 聂颖兰¹, 郭 威², 钱 丹^{1*}, 黄璐琦^{3*}

(1. 中国中医科学院医学实验中心, 中医药防治重大疾病基础研究北京市重点实验室, 北京 100700; 2. 山东大学附属山东省妇幼保健院, 山东 济南 250014; 3. 中国中医科学院, 北京 100700)

摘要: 为探讨枸杞属 (*Lycium*) nrDNA ITS 序列致同进化不完全现象及假基因问题, 本研究以 cDNA ITS 为参照, 对枸杞属栽培品种的 ITS 进行测序。综合 GenBank 数据库获取的 44 条枸杞属 ITS 序列进行分析。考察碱基变异、GC 含量、5.8S 保守基序、二级结构最小自由能、进化模式及系统发育关系等对枸杞属的 ITS 多态性和假基因进行鉴定。结果表明, 枸杞属植物内存在明显的 ITS 多态性, 植物种间及个体内存在致同进化不完全现象, 144 条序列中共 83 条序列被鉴定为假基因序列。本研究首次对枸杞属 ITS 多态性和假基因进行了较为系统地检验, 阐述了枸杞属 ITS 假基因特性, 评估了假基因对枸杞属系统进化的影响。为 ITS 序列用于枸杞属种质资源研究和 DNA 条形码鉴别研究等提供了参考。

关键词: 枸杞属; nrDNA ITS; 多态性; 致同进化不完全; 假基因

中图分类号: R931 文献标识码: A 文章编号: 0513-4870(2022)02-0525-08

Incomplete concerted evolution of nrDNA ITS in *Lycium*

ZHANG Jiao¹, NIE Ying-lan¹, GUO Wei², QIAN Dan^{1*}, HUANG Lu-qi^{3*}

(1. Beijing Key Laboratory of Research of Chinese Medicine on Prevention and Treatment for Major Diseases, Experimental Research Center, China Academy of Chinese Medical Sciences, Beijing 100700, China;

2. Shandong Provincial Maternal and Child Health Care Hospital, Shandong University, Jinan 250014,

China; 3. China Academy of Chinese Medical Sciences, Beijing 100700, China)

Abstract: In this study, nrDNA ITS sequences of *Lycium* cultivars were sequenced and used to test the existence of incomplete concerted evolution and pseudogenes. Together with 44 ITS sequences retrieved from GenBank, the pattern of base substitutions, GC content, 5.8S conserved motifs, the minimum free energy of secondary structures, nucleotide diversity and phylogenetic relationship of the samples were analyzed. While 83 of the 144 sequences were identified as pseudogenes, the results suggested a high degree of polymorphism and putative pseudogenes in *Lycium*, suggesting an incomplete concerted evolution of the ITS region. ITS polymorphism and pseudogene of *Lycium* were systematically tested for the first time. This research provides a references for ITS sequence to be used in the study of *Lycium* germplasm resources and DNA barcode identification.

Key words: *Lycium*; nrDNA ITS; polymorphism; incomplete concerted evolution; pseudogenes

枸杞是我国历史悠久的药食同源品种, 具有广泛

的药用和经济价值。2020 版《中国药典》收录枸杞来源为茄科植物宁夏枸杞 *Lycium barbarum* L. 的干燥成熟果实, 具有滋补肝肾、益精明目的功效^[1]。现代药理研究表明枸杞具有抗氧化、抗衰老、神经保护、保肝明目等多种活性^[2], 药用价值广泛。此外, 枸杞常作为保健品、食品的优质天然原料及普通群众的滋补茶蔬, 具有显著的经济价值。枸杞原植物来源于枸杞属, 枸杞

收稿日期: 2021-09-06; 修回日期: 2021-10-11.

基金项目: 国家自然科学基金资助项目 (81803671); 中央级公益性科研院所基本科研业务费专项资金 (ZZ13-YQ-079); 中国中医科学院科技创新工程 (CI2021A04104).

*通讯作者 Tel: 86-10-64089527,

E-mail: qiandan2104@163.com; huangluqi01@126.com

DOI: 10.16438/j.0513-4870.2021-1303

属在世界范围内广泛分布, 仅在我国作为经济植物广泛栽培于西北和华北地区^[3]。

核糖体 DNA (nuclear ribosomal DNA, nrDNA), 是由一些高度重复序列组成的多基因家族, 在高等植物中有成千上万的拷贝, 这些拷贝在基因组内串联重复, 分布于一对或多对染色体上^[4]。每个重复单元由 3 个 rRNA (18S rRNA、5.8S rRNA、26S rRNA) 共同构成, 基因间由转录间隔区 (internal transcribed spacer, ITS1 和 ITS2) 隔开^[5-7]。nrDNA (包括 ITS) 的进化模式通常为致同进化 (concerted evolution), 拷贝之间趋于均一化, 几乎没有差异, 为 ITS 序列的 PCR 扩增产物直接测序奠定了基础。此外, ITS 序列进化速率快, 片段长度适中, 能够提供较多的变异信息, 已经成为现今植物系统发育和 DNA 条形码鉴定研究中常用的分子标记之一^[4,8,9]。

随着研究的不断深入和测序技术的突飞猛进, 植物类群中 ITS 多态性逐渐受到学者们的关注。有研究发现, ITS 序列在植物基因组中存在致同进化不完全的现象, 拷贝之间存在明显的差异。此外, 一些 ITS 拷贝可能会由于功能退化而变成假基因 (pseudogenes)。已有研究证实, 在许多植物类群中存在 ITS 假基因现象, 如苏铁属 (*Cycas*)^[10]、梨属 (*Pyrus*)^[11]、栎属 (*Quercus*)^[12]、落叶松属 (*Larix*)^[13]、山茶属 (*Camellia*)^[14-16]、樟属 (*Cinnamomum*)^[17]、胡枝子属 (*Lespedeza*)^[18]、乳突球属 (*Mammillaria*)^[19,20]、茜草科 (*Rubiaceae*)^[21,22] 等。药用植物中 ITS 多态性也逐渐引起重视, 如紫堇属 (*Corydalis*)^[23]、人参属 (*Panax*)^[24] 等植物 ITS 多态性已被证明, 但是, 对药用植物 ITS 假基因的现象尚未展开深入研究。

中国枸杞属植物有 7 种 3 变种, 目前枸杞属 ITS 的研究主要集中于物种鉴定和遗传关系分析。2017 年首次有报道提及枸杞品种及个体内的 ITS 多态性和假基因现象^[25], 但尚未展开系统而深入的研究。本研究以枸杞属植物为研究对象, 通过直接测序、克隆测序 DNA 和 cDNA 的 ITS 序列, 评估枸杞种内和种间的 ITS 多态性, 在此基础上进行 ITS 假基因的鉴定, 同时推测 ITS 致同进化不完全和假基因的成因, 以期对枸杞属系统进化和物种形成研究提供一定的参考。

材料与方法

材料 本研究以枸杞属栽培品种为研究对象, 采集

新鲜叶片, 通过变色硅胶干燥和液氮低温保存两种方式储存, 具体情况见表 1。所有样品由中国医学科学院药用植物研究所姚入宇博士和中国中医科学院医学实验中心钱丹副研究员鉴定。同时, 从 GenBank 下载 44 条枸杞属 ITS 序列 (登录号: JQ320139~JQ320172; KJ189761~KJ189770), 具体情况见表 2。

植物总 DNA 提取和 PCR 扩增 采用天根生化科技 (北京) 有限公司的 Plant Genomics DNA Kit (Cat No. DP305-03) 提取植物总 DNA。使用通用引物 ITS1 (5'-TCCGTAGGTGAACCTGCGG-3') 和 ITS4 (5'-TCCTCGCTTATTGATATGC-3')^[26] 进行 ITS 序列扩增。

PCR 扩增使用全式金生物技术 (北京) 有限公司的 EasyTaq 酶。PCR 扩增体系为 50 μ L: 2 \times SuperMix 25 μ L、上下游引物各 2 μ L、模板 DNA 2 μ L、ddH₂O 19 μ L。扩增程序如下: 94 $^{\circ}$ C, 预变性 5 min; 94 $^{\circ}$ C 变性 30 s, 56 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C 延伸 45 s, 38 个循环; 72 $^{\circ}$ C 延伸 10 min。

植物总 RNA 提取和 RT-PCR 为了更好地判断基因组 5.8S 序列是否具有功能, 将相关参数与 RNA 反转录合成的 cDNA ITS 进行比较^[19]。使用天根生化科技 (北京) 有限公司的 RNA Easy Fast Plant Tissue Kit (Cat No. DP452) 提取植物总 RNA。用 FastKing RT Kit (With gDNase) 试剂盒合成 cDNA 第一链, 再以与 PCR 扩增相同的引物和条件进行 RT-PCR 获得 cDNA ITS。

克隆和测序 PCR 和 RT-PCR 产物经 1.0% 琼脂糖凝胶电泳后, 切下目的片段并用 Thermo Scientific GeneJET Gel Extraction Kit 凝胶回收试剂盒回收、纯化产物, 使用天根生化科技 (北京) 有限公司 Lethal Based Fast Cloning Kit 试剂盒进行克隆。然后每个样品挑选 11 个克隆产物, 经 PCR 鉴定阳性克隆后送往北京诺赛基因组研究中心有限公司测序, 同时, 对 PCR 产物也进行了直接测序。

序列处理与分析 使用 Seqman (DNASar, Madison, WI) 软件, 通过将测序序列与相应峰图对比完成单向序列的校正, 然后将正向和反向互补序列进行对比拼接, 以保证所得序列的准确性。根据 GenBank 中已有的枸杞属植物 ITS 序列, 确定 ITS1、5.8S 和 ITS2 的边界。

为了确定枸杞属植物是否存在 ITS 假基因, 进行如下分析: ① 使用 MEGA 7 计算 ITS 序列长度和 GC 含量, 并基于 K2P 双参数模型 (Kimura 2-parameter model)

Table 1 The information of the *Lycium* materials

Code	Chinese name	Speices	Number	Collection locality
N1	Ningqi NO.1	<i>L. barbarum</i>	1	Institute of Medicinal Plant Development
N5	Ningqi NO.5	<i>L. barbarum</i>	1	Institute of Medicinal Plant Development
N7	Ningqi NO.7	<i>L. barbarum</i>	1	Institute of Medicinal Plant Development
QH-N5	Ningqi NO.5	<i>L. barbarum</i>	10	Delingha, Qinghai
ZH	Chinese Gouqi	<i>L. chinense</i>	1	China Academy of Chinese Medical Sciences

Table 2 The information of the GenBank download

Chinese name	Speices	Location	GenBank accession No.
Changshu Gouqi		Wolfberry Germplasm Resources Garden	JQ320139
Korea Gouqi		Korea	JQ320140
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320141
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320142
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320143
Lycium hybrid cultivar		Wolfberry Germplasm Resources Garden	JQ320144
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320145
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320146
Lycium hybrid cultivar		Wolfberry Germplasm Resources Garden	JQ320147
Qinghaiheiguo Gouqi	<i>L. ruthenicum</i>	Wolfberry Germplasm Resources Garden	JQ320148
Lycium hybrid cultivar		Wolfberry Germplasm Resources Garden	JQ320149
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320150
Zibing Gouqi		Wolfberry Germplasm Resources Garden	JQ320151
Lycium hybrid cultivar		Wolfberry Germplasm Resources Garden	JQ320152
Lycium hybrid cultivar		Wolfberry Germplasm Resources Garden	JQ320153
Jice Gouqi	<i>L. truncatum</i>	Wolfberry Germplasm Resources Garden	JQ320154
Zhutong Gouqi	<i>L. cylindricum</i>	Wolfberry Germplasm Resources Garden	JQ320155
Xinjiang Gouqi	<i>L. dasystemum</i>	Wolfberry Germplasm Resources Garden	JQ320156
Yunnan Gouqi	<i>L. yunnanense</i>	Wolfberry Germplasm Resources Garden	JQ320157
Mansheng Gouqi		Wolfberry Germplasm Resources Garden	JQ320158
Hongzhi Gouqi	<i>L. cylindricum</i>	Wolfberry Germplasm Resources Garden	JQ320159
Chinese Gouqi	<i>L. chinense</i>	Wolfberry Germplasm Resources Garden	JQ320160
Lycium hybrid cultivar		Wolfberry Germplasm Resources Garden	JQ320161
Heiguo Gouqi	<i>L. ruthenicum</i>	Wolfberry Germplasm Resources Garden	JQ320162
Huangguo Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320163
Beifang Gouqi		Wolfberry Germplasm Resources Garden	JQ320164
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320165
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320166
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320167
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320168
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320169
Ningxia Gouqi	<i>L. barbarum</i>	Wolfberry Germplasm Resources Garden	JQ320170
Zhongning Gouqi	<i>L. ruthenicum</i>	Wolfberry Germplasm Resources Garden	JQ320171
Meiguo Gouqi		America	JQ320172
0702		Jingtai, Gansu	KJ189761
Ningqi No.4	<i>L. barbarum</i>	Jingtai, Gansu	KJ189762
Qinghaiheiguo Gouqi	<i>L. ruthenicum</i>	Nuomuhong, Qinghai	KJ189763
Ningqi No.1	<i>L. barbarum</i>	Jingtai, Gansu	KJ189764
Bianguo Gouqi		Jingtai, Gansu	KJ189765
Ningqi No.3	<i>L. barbarum</i>	Jingtai, Gansu	KJ189766
0901		Jingtai, Gansu	KJ189767
Ningqi No.7	<i>L. barbarum</i>	Jingtai, Gansu	KJ189768
Ningqi No.5	<i>L. barbarum</i>	Jingtai, Gansu	KJ189769
Mengqi No.1		Jingtai, Gansu	KJ189770

构建 NJ 无根进化树, 以 bootstrap 方法 2 000 次重复取样分析分支支持度; ② 检测是否存在种子植物 5.8S 特有的 3 个保守基序 (motif1: CGATGAAGAACGTAGC; motif2: GAATTGCAGAATCC; motif3: TTTGAACGCA)^[27]; ③ 通过在线软件 Mfold^[28] 预测 5.8S 和 ITS2 区的最小自由能 (37 °C 时的 ΔG)。在初步确定假基因的基础上, 再用 DnaSP 计算功能序列和假基因的核苷酸差异平均值 (average number of nucleotide differences, K)、核苷酸多样性 (nucleotide diversity, π) 并进行中性检验 (Tajima's D 或 Fu and Li's D), 进一步对前面的分析结果加以验证。

结果与分析

1 ITS 序列长度、GC 含量、最小自由能分析及假基因的鉴定

本研究共得到 100 条不同的 ITS 序列, 其中包括 10 条直接测序 ITS 序列和 90 条克隆 ITS 序列 (57 条基因组 DNA ITS 序列和 33 条 cDNA ITS 序列)。直接测序 ITS 序列长度变异范围为 601~648 bp, 基因组 DNA ITS 序列长度变异范围为 535~694 bp, cDNA ITS 序列长度变异范围为 536~648 bp。与功能基因相比, 假基因具有明显的长度变异、较低的 GC 含量及较

高的二级结构最小自由能等特征,通过对144条ITS序列(本研究100条和GenBank 44条)的碱基组成、长度变异、5.8S和ITS2二级结构最小自由能进行统计分析(表3),发现83条序列符合假基因特征,因此鉴定为假基因。

枸杞5号克隆测序ITS序列中(N5、QH-N5-1和QH-N5-5)既有功能序列,也存在假基因序列。而所有样品的直接测序结果和其他材料(N1、N7和ZH)的克隆序列得到的全部为假基因序列。GenBank下载的44条ITS序列分析发现功能序列和假基因序列各占一半,均为22条。假基因序列的ITS1、5.8S和ITS2三个区域GC含量显著低于功能序列,表明其甲基化C位点碱基发生频繁突变,而5.8S和ITS2的最小自由能却显著高于功能序列,表明假基因序列的二级结构稳定性较功能序列低。

种子植物5.8S特有的3个保守基序(motif1、motif2和motif3)的碱基突变是假基因判定的主要标准。本研究的144条ITS序列中,motif3保守型最高,95%的序列中均含有该基序。其中,28条序列含有5.8S全部的3个保守基序且没有发生碱基变异,被鉴定为功能序列。33条cDNA功能序列的3个保守基序中均以motif1第13位T→C和motif2第13位C→T组合变异。83条序列的5.8S不含有保守基序或3个保守

基序中至少有一个存在碱基变异,被鉴定为假基因。假基因的3个保守基序碱基变异呈现一定的规律性,其中,以motif1第12位G→A、16位C→A和motif2第3位A→G组合变异为主,占总体变异的67.5%。

2 核苷酸多样性及中性检测

核苷酸多样性分析显示,假基因序列5.8S区的核苷酸多样性(π)和平均核苷酸差异数(K)相对于功能序列高(表4),其 π 值和 K 值均是功能序列的1.34倍,说明5.8S假基因序列具有较多的变异。对于所有假基因序列,3个区域的核苷酸多样性和平均核苷酸差异数数值高低排列顺序不同,核苷酸多样性排列顺序为ITS1 > 5.8S > ITS2,平均核苷酸差异数排列顺序为ITS2 > ITS1 > 5.8S。

GenBank的44条序列和本研究的100条序列因测序引物的不同分别进行中性检验。通过Tajima's D或Fu and Li's D检验结果发现(表5),ITS功能序列D值的绝对值较大,且5.8S区呈现显著性差异($P < 0.001$ 或 $P < 0.02$),表明ITS功能序列偏离中性进化;而假基因序列D值的绝对值较小,且未呈现显著性差异($P > 0.10$),表明假基因序列倾向于中性进化^[29,30]。

3 系统发育分析

基于所有ITS序列的遗传距离构建的NJ树中,除枸杞5号的部分cDNA(cDN5-2、cDN5-4、cDN5-9和

Table 3 Characteristic information of ITS regions of *Lycium*. F: Functional ITS sequences; P: ITS pseudogene sequences; SD: Standard deviation; Δ : Direct sequencing succeeded

Code	Type	No.	Length/bp (SD)			GC content/% (SD)			5.8S Δ G (SD)	ITS2 Δ G (SD)
			ITS1	5.8S	ITS2	ITS1	5.8S	ITS2	/kcal·mol ⁻¹	/kcal·mol ⁻¹
N1	P	10	220.3 (6.8)	158.2 (1.6)	254.4 (23.0)	53.1 (0.0)	50.5 (0.0)	57.5 (0.0)	-48.0 (2.0)	-100.5 (13.1)
N5 ^Δ	F	2	260.0 (0.0)	159.0 (0.0)	275.0 (0.0)	56.9 (0.0)	56.0 (0.0)	58.4 (0.0)	-51.9 (0.0)	-101.7 (0.8)
	P	8	209.8 (23.2)	158.3 (2.0)	252.6 (21.2)	52.6 (0.0)	50.4 (0.0)	55.5 (0.0)	-48.3 (1.6)	-95.1 (20.1)
N7 ^Δ	P	10	217.2 (15.4)	160.4 (1.8)	253.6 (14.6)	53.9 (0.0)	50.2 (0.0)	56.9 (0.0)	-50.4 (1.7)	-98.65 (14.7)
QH-N5 ^Δ	F	4	258.0 (0.0)	161.0 (0.0)	269.0 (0.0)	57.0 (0.0)	56.1 (0.0)	58.3 (0.0)	-52.5 (0.6)	-95.9 (1.7)
	P	25	227.39 (10.3)	161.0 (0.0)	165.2 (12.9)	57.5 (0.0)	52.0 (0.0)	57.9 (0.0)	-50.5 (0.7)	-100.3 (6.0)
ZH	P	8	230.8 (13.1)	161.0 (0.0)	266.3 (8.1)	57.9 (0.0)	51.9 (0.0)	57.8 (0.0)	-49.5 (2.9)	-100.2 (4.2)
JQ320139-JQ320172	F	18	267.8 (6.6)	154.0 (0.0)	257.8 (0.5)	67.0 (1.4)	55.3 (1.5)	69.0 (1.0)	-41.6 (3.7)	-123.3 (3.5)
	P	16	249.9 (21.7)	154.0 (0.0)	251.3 (7.8)	59.9 (4.1)	52.4 (1.4)	62.6 (3.4)	-39.1 (4.4)	-103.2 (9.7)
KJ189761_KJ189770	F	4	262.3 (11.7)	154.0 (0.0)	231.5 (1.1)	69.1 (0.9)	51.3 (0.8)	68.0 (2.3)	-44.5 (2.4)	-113.9 (1.8)
	P	6	215.0 (0.0)	154.0 (0.0)	213 (0.0)	58.9 (0.2)	51.3 (0.0)	54.4 (0.2)	-44.0 (0.1)	-85.0 (0.6)
ALL	F	28	265.1 (7.9)	155.4 (2.6)	256.9 (11.7)	65.2 (0.0)	54.9 (0.0)	66.5 (0.0)	-44.3 (5.3)	-116.5 (10.8)
	P	83	227.4 (19.5)	158.4 (17.5)	255.3 (19.4)	56.7 (0.0)	51.5 (0.0)	58.1 (0.0)	-47.2 (4.9)	-99.2 (11.6)

Table 4 Nucleotide diversity of the ITS regions in *Lycium*. F: Functional ITS sequences; P: ITS pseudogene sequences; Nucleotide diversity (π); Average number of nucleotide differences (K)

ITS region	Sequence type	Number of sequences	Polymorphic site No.	Total mutation	π	K
ITS1	F	28	157	307	0.466 77	74.682 54
	P	83	169	472	0.548 26	92.656 48
5.8S	F	28	154	357	0.406 77	62.642 86
	P	83	154	391	0.545 15	83.953 57
ITS2	F	28	229	436	0.479 63	110.314 81
	P	83	204	539	0.483 75	98.684 98
ITS	F	28	624	1 565	0.596 07	373.738 10
	P	83	535	1 563	0.586 25	313.641 20

Table 5 Neutrality test of the ITS regions in *Lycium*. F: Functional ITS sequences; P: ITS pseudogene sequences. Significant difference at *** $P < 0.001$ in Tajima's D analysis and *** $P < 0.02$ in Fu and Li's D analysis

Source	ITS region	Sequence type	Number of sequences	Neutrality test		
				Tajima's D	Fu and Li's D	
KJ189761-KJ189770	ITS1	F	4	-1.506 60***	-0.421 63	
		P	6	-1.554 42	-1.587 40	
	5.8S	F	4	-1.494 75***	-0.504 97	
		P	6	-0.293 73	-0.169 71	
	ITS2	F	4	-1.570 81***	-0.420 82	
		P	6	1.352 61	1.721 56	
	ITS	F	4	-1.716 38***	-0.464 80	
		P	6	0.011 58	0.190 80	
	JQ320139-JQ320172	ITS1	F	18	-0.758 65	-1.260 87
			P	16	-0.977 50	-0.396 39
5.8S		F	18	-2.532 72***	-3.647 94***	
		P	16	-0.846 39	0.026 44	
ITS2		F	18	-1.144 55	-0.749 57	
		P	16	-0.167 98	0.518 16	
ITS		F	18	-0.600 75	0.348 76	
		P	16	-0.823 45	0.096 04	
This study		ITS1	F	6	-1.131 97	-1.155 29
			P	61	-0.856 03	0.723 26
	5.8S	F	6	1.385 27	1.731 03***	
		P	61	0.554 42	-1.267 97	
	ITS2	F	6	-1.367 32	-1.399 92	
		P	61	-0.827 50	-0.335 27	
	ITS	F	6	-1.453 26	-1.489 86	
		P	61	-0.739 80	1.055 59	

cDN5-10) 外, 所有 cDNA 功能序列单独聚为一支; 除 JQ320148、JQ320154、JQ320163、JQ320172 外, 枸杞属植物 ITS 功能序列和假基因序列全部分离出来 (图 1), 但不同品种的功能序列并没有完全聚为一支, 而是以较高的支持度聚为 3 支 (I、II、III), 其中第 I 支多为不同品种的宁夏枸杞及其杂交种; 第 II 支为黑果枸杞; 第 III 支为宁杞 5 号 (N5、QH-N5-1 和 QH-N5-5) 的 6 条克隆序列。

本研究鉴定的 83 条假基因序列, 同样以较高的支持度聚为 3 支 (A、B、C), 其中 A 支主要为宁夏枸杞和栽培品种 (N1、N5、N7、ZH、QH-N5) 的直接测序和 DNA 克隆序列, 均只有 motif3 保守基序; B 支为枸杞属其他种和变种 (包括红枝枸杞、新疆枸杞、韩国枸杞、蔓生枸杞、云南枸杞和常熟枸杞), 均只有 motif2 保守基序; C 支主要为 QH-N5-1、QH-N5-5 和 ZH 克隆序列, 均含有 motif2 和 motif3 两个保守基序。

N1、N7 和 ZH 的 cDNA ITS 克隆序列全部分布在同一支, 而 N5 的 cDNA ITS 克隆序列及 N1、N5、N7 的 DNA ITS 克隆序列均分布在 cDNA 支和 A 支两支上。这两支序列长度相差约 100 bp, 且 cDNA 支 ITS1、5.8S、ITS2 三个区域的 GC 含量均低于 A 支, 说明枸杞属植物 DNA 和 cDNA 的 ITS 序列存在明显的差异。此外, 本研究的系统发育聚类结果与其他两个研究团队的报道一致^[31,32], 推测假基因可能对枸杞属植物种属

系统构建影响不大。从假基因的角度分析聚类形成, 能更好地解释前人研究中聚类分支形成的原因。

讨论

nrDNA ITS 序列在高等植物中通常为致同进化, 不同的 ITS 序列基本趋于一致或完全一致, 这为其作为 PCR 扩增产物直接测序奠定了基础。同时, ITS 能够提供较为合适的变异和信息位点, 加之双亲遗传的特点, 已经广泛地应用于属内种间或种内群体的系统进化研究, 成为 DNA 条形码鉴定的核心序列。在应用 ITS 序列进行枸杞属植物的鉴定研究中, 有报道因 ITS 序列 PCR 直接测序合格率低, 无法获得理想结果, 不纳入统计^[33]; 也有研究结果^[25]显示直接测序合格率为 64%; 本研究中直接测序合格率为 71%。为提高合格率, 在 PCR 扩增中加入 DMSO 进行改善^[16], 筛选了 7 个不同的浓度梯度 (0%、1%、2%、4%、6%、8%、10%), 但直接测序合格率并未得到改善。由此可见, ITS 在枸杞属存在较低的扩增和鉴定成功率。

本研究对 ITS 序列在枸杞属的应用开展了进一步研究, 通过多次重复试验, 排除 PCR 过程污染、测序错误等可能性。发现枸杞属植物不同种间及个体间的基因组内、cDNA 组内的 ITS 序列均存在广泛的多态性, 不仅 GC 含量有一定的变化, 不同序列长度之间也存在变化。枸杞属 ITS 序列总长度为 535~694 bp, GC

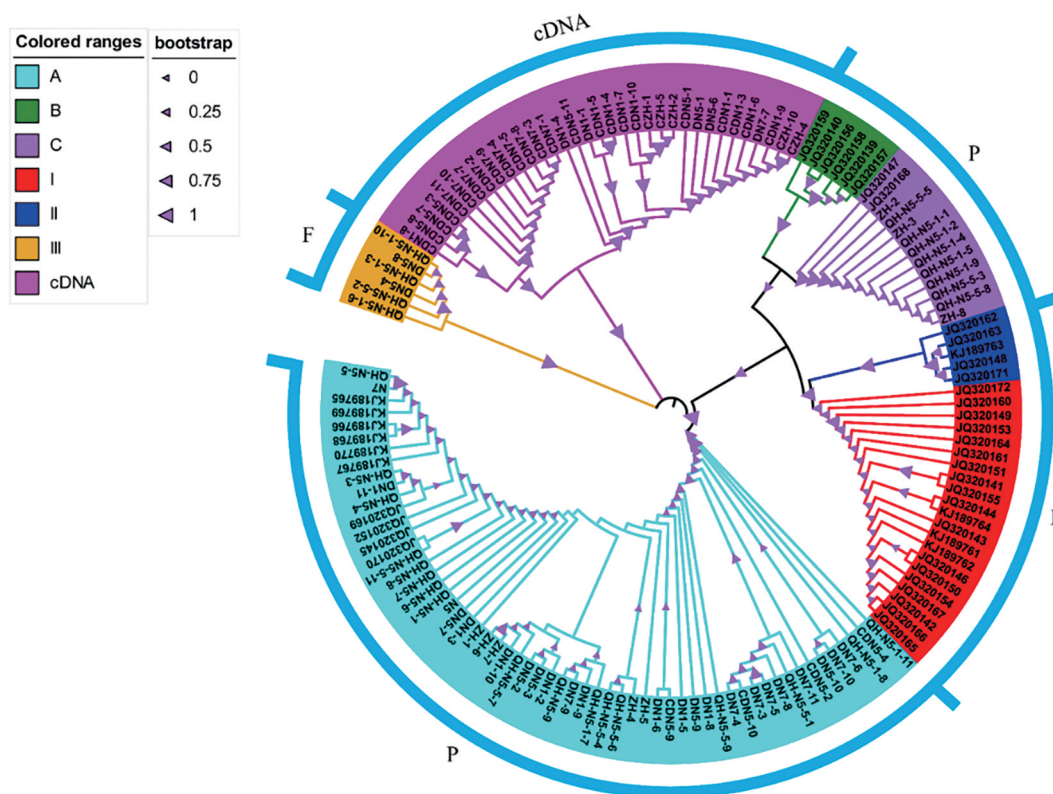


Figure 1 Neighbor-joining tree topology constructed on Kimura two-parameter distance matrix using ITS entire region of all paralogs in *Lycium*. A, B and C represent the three different types of pseudogenes, respectively. I, II and III represent the three different types of functional gene, respectively. cDNA represents cDNA sequence

含量为 46.17%~59.03%。不同序列长度间存在超过 100 bp 的差异且主要在 ITS1 区。基因组内不同克隆序列 GC 含量变化最大, 为 47.29%~59.03%, 而 cDNA 组 GC 含量变化范围为 46.17%~55.40%。此外, 枸杞属植物存在多种 ITS 变异类型, 如 DN5 检测到 3 种类型, 包括 1 种功能基因和 2 种假基因, 而其他样品均检测到一到两种变异类型。

高度重复的 nrDNA ITS 有成千上万的拷贝, 其基因位点在进化过程中易发生功能缺失突变, 如插入、缺失或移码, 从而形成假基因^[34]。本研究发现枸杞属植物种间和个体内均存在大量的 ITS 假基因。与功能基因相比, 假基因的 GC 含量明显降低, 表明其甲基化 C 位点碱基发生频繁突变, 低 GC 含量可能会导致基因功能的丧失, 也可能是假基因的起源^[35]; 而 5.8S 和 ITS2 区二级结构最小自由能明显偏高, 表明其二级结构稳定性较功能序列显著降低。进一步分析发现, 假基因序列在高度保守的 5.8S 区的核苷酸多态性和平均核苷酸差异数明显偏大, 说明其序列发生了较多的变异, 包括碱基替换和长度变异。同时, 由于假基因不受选择压力的约束, 在进化过程中核苷酸置换脱离了功能限制, 倾向于中性进化, 因此, 比功能基因的进化

速率要快。

5.8S 序列通常被认为高度保守, 有报道在系统发育分析时往往将其排除不加以考虑^[31,32]。然而在 ITS 假基因的鉴定研究中, 5.8S 是 ITS 拷贝中极具价值的功能指标。通过检测 5.8S 区是否含有种子植物特有的 3 个保守基序或是否发生碱基变异及 GC 含量、二级结构最小自由能等可对假基因进行判定。本研究通过 5.8S 序列的变异进行了相关评估, 判定方法简便有效, 同时发现了前人在枸杞属系统发育研究中使用的 ITS 序列部分为假基因。

基于系统发育树分析, 本研究中枸杞属植物共存在 3 种类型的假基因且彼此间存在较大的差异, 类型 A 以宁夏枸杞和栽培品种的直接测序和 DNA 克隆序列为主; 类型 B 以枸杞属其他种和变种的克隆序列为主; 而类型 C 以 QH-N5-1、QH-N5-5 和 ZH 克隆序列为主。此外, ZH、QH-N5-1 和 QH-N5-5 三个个体中存在两种类型的 ITS 假基因, 在系统发育树中分布于 A 和 C 两支, 其中, A 支仅含有 motif1 保守基序, C 支含有 motif2 和 motif3 两个保守基序。这些结果表明, 枸杞属植物不同种间及个体基因组内均存在显著的致同进化不完全现象。

基因家族成员以协同的方式进化,这种进化模式称为“致同进化”,即当某个成员发生突变时,它能够通过反复的基因交换或不等交换传播到其他成员中去^[36]。以往的研究显示,物种致同进化不完全受多种因素的影响,既有物种进化和生活史方面的因素,如多倍体化^[37]、异源多倍体化^[38]、孤雌生殖^[39]、世代周期长^[40]和杂交^[12]等,也有基因家族在染色体上、基因组中组织结构方面的因素。枸杞在中国的栽培历史悠久,存在很多的农家品种和品系。杂交是其主要的育种手段之一^[41]。如雄性不育种质“宁杞5号”的样品种内和个体内均存在明显的多态性,克隆测序后序列均为功能序列和假基因序列两种类型。甚至其 cDNA 功能序列也分为两类,在系统发育树上分布于 cDNA 支和 A 支,且这两类 cDNA 长度存在 100 bp 的差异。此外,枸杞属植物虽存在多种变异类型,但同一个体的不同克隆序列大多被置于一到两个不同的分支,并且位于同一分支上的序列间具有很高的相似度,仅存在少数碱基的差异。因此,推测基因组内串联重复序列之间的近期分化、物种间的杂交可能是导致枸杞属植物个体内 ITS 相似序列的变异及致同进化不完全的主要原因。

与四照花属^[42]、苏铁属^[10]的结果相似,枸杞属绝大部分假基因序列与功能序列在基因树上各自独立形成分支,表明假基因与功能基因起源时间存在明显的差异,又因假基因较功能基因进化速率快,倾向于中性进化,因此推测枸杞属 ITS 假基因起源时间可能相对较早,并且具备了较多的遗传变异,但假基因没有形成一个独立的单系,说明他们可能不是一次起源,或者一次起源后变异分化的速度差异较大。物种的进化起源对物种的分类鉴定、生物多样性保护及遗传育种具有重要意义,ITS 多态性对枸杞属植物起源进化和系统发育具有重要影响,但是克隆测序结果往往偏向于获取高丰度的序列,而低丰度的序列难以获取^[24],因此,后续研究可采用高通量测序技术,更深层次地揭示 ITS 多态性对物种进化形成的影响。

本文首次对枸杞属植物 ITS 序列致同进化不完全和假基因现象进行了较为系统地报道,发现枸杞属种间、种内及个体内均存在 ITS 多态性和假基因现象。对枸杞属致同进化不完全进行探讨,发现枸杞属 ITS 致同进化不完全或许是导致枸杞属植物 DNA 条形码研究中测序成功率和鉴定率较低的重要原因之一。初步推测杂交可能是导致枸杞属 ITS 多态性的主要原因,后续研究将进一步扩大样品量,采用高通量测序进行相关研究,为假基因在枸杞属的应用提供更多 DNA 序列的分子数据。本研究为 ITS 序列在枸杞属内亲缘关系、系统发育重建和 DNA 条形码鉴定的研究提供了

一定的参考和科学依据。

致谢:感谢中国医学科学院药用植物研究所姚入宇博士为本研究提供部分材料。

作者贡献:黄璐琦构思并设计了实验方案;钱丹完成样品采集,数据分析和文章修改;张娇完成样品采集,实验并撰写论文初稿;聂颖兰和郭威参与数据分析;所有作者均阅读并参与了文章的修改。

利益冲突:本文无相关利益冲突。

References

- [1] Chinese Pharmacopoeia Committee. Pharmacopoeia of the People's Republic of China (中华人民共和国药典) [S]. Beijing: China Medical Science Press, 2020: 260.
- [2] Wang SS, Zhang Z, Chen NH. Research progress on main active components and pharmacological effects of wolfberry [J]. Acta Neuropharmacol (神经药理学报), 2018, 8: 53.
- [3] Qian D, Ji R, Guo W, et al. Advances in research on relationships among Lycium species and origin of cultivated Lycium in China [J]. China J Clin Mater Med (中国中药杂志), 2017, 42: 3282-3285.
- [4] Eickbush TH, Eickbush DG. Finely orchestrated movements: evolution of the ribosomal RNA genes [J]. Genetics, 2007, 175: 477-485.
- [5] Tian X, Li DZ. Application of DNA sequences in plant phylogenetic study [J]. Acta Bot Yunnan (云南植物研究), 2002, 24: 170-184.
- [6] Álvarez I. Ribosomal ITS sequences and plant phylogenetic inference [J]. Mol Phylogenet Evol, 2003, 29: 417-434.
- [7] Li DZ, Gao LM, Li HT, et al. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants [J]. Proc Natl Acad Sci U S A, 2011, 108: 19641-19646.
- [8] Bailey C. Characterization of angiosperm nrDNA polymorphism paralogy, and pseudogenes [J]. Mol Phylogenet Evol, 2003, 29: 435-455.
- [9] Kovarik A, Pires JC, Leitch AR, et al. Rapid concerted evolution of nuclear ribosomal DNA in two tragopogon allopolyploids of recent and recurrent origin sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AY458586, AY458588, AY458589, and AY458587 [J]. Genetics, 2005, 169: 931-944.
- [10] Xiao LQ, Möller M, Zhu H. High nrDNA ITS polymorphism in the ancient extant seed plant Cycas: incomplete concerted evolution and the origin of pseudogenes [J]. Mol Phylogenet Evol, 2010, 55: 168-177.
- [11] Zheng XY, Cai DY, Yao LH, et al. Non-concerted ITS evolution, early origin and phylogenetic utility of ITS pseudogenes in Pyrus [J]. Mol Phylogenet Evol, 2008, 48: 892-903.

- [12] Muir G, Fleming CC, Schlötterer C. Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. [J]. Mol Biol Evol, 2001, 18: 112-119.
- [13] Weu XX, Wang XQ. Recolonization and radiation in Larix (Pinaceae): evidence from nuclear ribosomal DNA paralogues [J]. Mol Ecol, 2004, 13: 3115-3123.
- [14] Xu Y, Xu J, Gao JY, et al. Polymorphism of ITS in Camellia: an example of widespread escape from consistent evolution [J]. Chin Bull Bot (植物学报), 2011, 46: 162-169.
- [15] You H, Zhou AT, Yue LL, et al. Amplification and sequence characterization of ITS in Camellia [J]. Bull Bot Res (植物研究), 2014, 34: 403-408.
- [16] Zhou AT, Yue LL, Li M, et al. Intra-genomic polymorphism in the nrDNA ITS sequence of *Camellia reticulata* [J]. Plant Sci J (植物科学学报), 2013, 31: 1-10.
- [17] Huang JF, Li L, van der Werff H, et al. Origins and evolution of cinnamon and camphor: a phylogenetic and historical biogeographical analysis of the *Cinnamomum* group (Lauraceae) [J]. Mol Phylogenet Evol, 2016, 96: 33-44.
- [18] Xu B, Zeng XM, Gao XF, et al. ITS non-concerted evolution and rampant hybridization in the legume genus *Lespedeza* (Fabaceae) [J]. Sci Rep, 2017, 7: 40057.
- [19] Harpke D, Peterson A. Non-concerted ITS evolution in *Mammillaria* (Cactaceae) [J]. Mol Phylogenet Evol, 2006, 41: 579-593.
- [20] Harpke D, Peterson A. Extensive 5.8S nrDNA polymorphism in *Mammillaria* (Cactaceae) with special reference to the identification of pseudogenic internal transcribed spacer regions [J]. J Plant Res, 2008, 121: 261-270.
- [21] Razafimandimbison SG, Kellogg EA, Bremer B. Recent origin and phylogenetic utility of divergent ITS putative pseudogenes: a case study from Naucleaeae (Rubiaceae) [J]. Syst Biol, 2004, 53: 177-192.
- [22] de Sousa Queiroz C, de Carvalho Batista FR, de Oliveira LO. Evolution of the 5.8S nrDNA gene and internal transcribed spacers in *Carapichea ipeacacuanha* (Rubiaceae) within a phylogeographic context [J]. Mol Phylogenet Evol, 2011, 59: 293-302.
- [23] Li M, Zhao H, Zhao F, et al. Alternative analyses of compensatory base changes in an ITS2 phylogeny of *Corydalis* (Papaveraceae) [J]. Ann Bot, 2019, 124: 233-243.
- [24] Song J, Shi L, Li D, et al. Extensive pyrosequencing reveals frequent intra-genomic variations of internal transcribed spacer regions of nuclear ribosomal DNA [J]. PLoS One, 2012, 7: e43971.
- [25] Chen LL, Zhao MX, Jiang S, et al. Identification, phylogenetic relationship analysis of *Lycium* based on rbcL-a and ITS sequences and the discovery of ITS pseudogene [J]. Biotechnol Bull (生物技术通报), 2017, 33: 123-130.
- [26] White T, Bruns T, Lee S, et al. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics [M] // Innis M, Gel-fand J, Sainsky J, et al, editors. PCR Protocols, a Guide to Methods and Applications. San Diego: Academic Press, Inc., 1990: 315-321
- [27] Harpke D, Peterson A. 5.8S motifs for the identification of pseudogenic ITS regions [J]. Botany, 2008, 86: 300-305.
- [28] Zuker M. Mfold web server for nucleic acid folding and hybridization prediction [J]. Nucleic Acids Res, 2003, 31: 3406-3415.
- [29] Fu YX. New statistical tests of neutrality for DNA samples from a population [J]. Genetics, 1996, 143: 557-570.
- [30] Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection [J]. Genetics, 1997, 147: 915-925.
- [31] Shi ZG, Wan R, Li YL, et al. Identification of *Lycium* L. germplasm resources based on nrDNA ITS sequence [J]. Hubei Agric Sci (湖北农业科学), 2016, 55: 5966-5968.
- [32] Li XG, Wang YK, Li J, et al. Study on 10 species of *Lycium* L. genetic relationship based on nrDNA-ITS sequence [J]. Chin Agric Sci Bull (中国农学通报), 2014, 30: 128-135.
- [33] Xin TY, Yao H, Gao HH, et al. Super food *Lycium barbarum* (Solanaceae) traceability via an internal transcribed spacer 2 barcode [J]. Food Res Int, 2013, 54: 1699-1704.
- [34] You H, Zhou AT, Yue LL, et al. Research on nrDNA ITS pseudogenes in plants [J]. Biotechnol Bull (生物技术通报), 2013, 11: 14-18.
- [35] Bailey JA, Liu G, Eichler EE. An alu transposition model for the origin and expansion of human segmental duplications [J]. Am J Hum Genet, 2003, 73: 823-834.
- [36] Nei M, Rooney AP. Concerted and birth-and-death evolution of multigene families [J]. Annu Rev Genet, 2005, 39: 121-152.
- [37] Suh Y, Thien LB, Zimmer EA. Nucleotide sequences of the internal transcribed spacers and 5.8S rRNA gene in *Canella winterana* (Magnoliales; Canellaceae) [J]. Nucleic Acids Res, 1992, 20: 6101-6102.
- [38] Karvonen P, Savolainen O. Variation and inheritance of ribosomal DNA in *Pinus sylvestris* L. (Scots pine) [J]. Heredity, 1993, 71: 614-622.
- [39] Campbell CS, Wright WA, Cox M, et al. Nuclear ribosomal DNA internal transcribed spacer 1 (ITS1) in *Picea* (Pinaceae): sequence divergence and structure [J]. Mol Phylogenet Evol, 2005, 35: 165-185.
- [40] Sang T, Crawford DJ, Stuessy TF. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution [J]. Proc Natl Acad Sci U S A, 1995, 92: 6813-6817.
- [41] An W, Zhan HX, He J, et al. Progress and prospects of wolfberry breeding [J]. Nor Hort (北方园艺), 2009, 33: 125-128.
- [42] Hu ZF, Xiang QY, Huang HW. Incomplete concerted evolution of nrDNA ITS in *Cornus subg. Syncarpea* (Cornaceae) [J]. J Trop Subtrop Bot (热带亚热带植物学报), 2013, 21: 29-37.