

云南山茶 (*Camellia reticulata*) nrDNA ITS 序列多态性分析

周阿涛¹, 岳亮亮², 李旻², 刘迪秋^{1*}, 丁元明^{2*}

(1. 昆明理工大学生命科学与技术学院, 昆明 650224; 2. 云南出入境检验检疫局, 昆明 650228)

摘 要: 对云南山茶 (*Camellia reticulata*) 8 个品种的核糖体 DNA 内转录间隔区 (nrDNA ITS) 进行克隆测序, 将获得的 ITS 序列进行 GC 含量、5.8S 区二级结构的稳定性、替代模式、核苷酸多样性及系统发育关系的相关分析。实验结果显示, 云南山茶 8 个品种的 ITS 序列存在丰富的基因组内多态性, 同时包含中性进化的假基因, 表明其 ITS 序列逃离了一致性进化。云南山茶 ITS 序列多态性的原因可能来自广泛的种间杂交, 以及 rDNA 位点在基因组中有不确定的物理位置。ITS 假基因为品种的物种形成研究提供了更全面的遗传证据, 同时也提示了利用 ITS 假基因进行系统发育分析可能会对其真实的系统关系造成影响。

关键词: 云南山茶; nrDNA ITS; 多态性; 假基因

中图分类号: Q941⁺.1; Q949.758.4

文献标识码: A

文章编号: 2095-0837(2013)01-0001-10

Intra-genomic Polymorphism in the nrDNA ITS Sequence of *Camellia reticulata*

ZHOU A-Tao¹, YUE Liang-Liang², LI Min², LIU Di-Qiu^{1*}, DING Yuan-Ming^{2*}

(1. Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming 650224, China;

2. Yunnan Entry-Exit Inspection and Quarantine, Kunming 650228, China)

Abstract: In this study, nrDNA ITS copies from *Camellia reticulata* were cloned and sequenced. Complementary analyses were performed including GC content, estimated secondary structure stability of the 5.8S region, pattern of base substitutions, nucleotide diversity, and phylogenetic reconstruction. The ITS copies showed a high degree of polymorphism including pseudogenes in *C. reticulata*, suggesting an incomplete concerted evolution of the ITS region that may have resulted from hybridization events and dispersed distribution of rDNA loci among chromosomes. Diversity of the pseudogenes can help in understanding the origin of the cultivar *C. reticulata*. However, great caution should be shown when using these pseudogenes to reconstruct phylogenetic trees.

Key words: *Camellia reticulata*; nrDNA ITS; Polymorphism; Pseudogenes

高等植物的核糖体 DNA (rDNA) 是一些高度重复序列组成的多基因家族, 在基因组内往往有成千上万的拷贝, 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^[1]。每个重复单元由 18S rRNA、5.8S rRNA 和 26S rRNA 共同构成, 位于重

复单元之间的转录间隔区 (internal transcribed spacer, ITS1 和 ITS2) 是研究较低分类群系统关系的重要标记^[2-4]。nrDNA (包括 ITS) 通常是以一致进化模式 (concerted evolution) 进化的, 在这种进化模式下, 不同的 ITS 拷贝序列会基本趋于一致或

收稿日期: 2012-09-12, 修回日期: 2012-11-09。

基金项目: 国家质检总局科技计划项目: 云南出口花卉物种信息平台建立及分子生物学检测技术研究 (2010IK273)。

作者简介: 周阿涛 (1988 -), 男, 硕士研究生, 研究方向为分子生物学 (E-mail: zhouatao_520@163.com)。

* 通讯作者: 丁元明 (1964 -), 男, 研究员, 主要从事植物及进出口检疫相关研究; 刘迪秋 (1979 -), 女, 副教授, 研究方向为分子生物学 (Author for correspondence. E-mail: dingdym2008@yahoo.com.cn; diqiuliu@gmail.com)。

完全一致^[1,5,6]。因此,在同一基因组内,ITS 基因多样性会倾向于降低或消失。驱动一致性进化等分子机制包括不等交换(unequal crossover)、基因转换(gene conversion)和基因扩增(gene amplification)等^[7-9]。然而,尽管一致进化在 rDNA 的进化中起着很大的作用,但事实上 ITS 多态性并非只存在于不同的基因组之间。越来越多的研究发现,某些物种内甚至个体内的 ITS 序列一致进化不完全,拷贝之间存在明显的差异^[10-12]。此外,在很多植物中还发现有 ITS 假基因拷贝,如苏铁属(*Cycas*)^[10]、梨属(*Pyrus*)^[11]、栎属(*Quercus*)^[12]、落叶松属(*Larix*)^[13]等植物中均存在 ITS 假基因,这些假基因不会在短时间内从基因组中消失,它们或许以不同于功能基因的进化速率发生着独立的变化^[5]。ITS 多态性的发现,在一定程度上增加了研究工作的难度^[5]。但是,如果对这些拷贝进行充分取样分析,研究 ITS 多态性的成因和发展规律,可能会为系统进化和物种形成研究提供更全面的遗传证据^[13-16]。

山茶属(*Camellia* L.)是山茶科(Theaceae)中最大的一个属。近年来,ITS 作为常用分子标记已被广泛应用于山茶属植物的系统分类^[17-22],同时也发现了部分种类的 ITS 存在多态性。这些研究认为山茶属植物 ITS 多态性可能是因其基因组拥有较多的 rDNA 位点所致,同时也提出种间杂交可能导致 ITS 多态性的观点^[23-25],但没有给出直接的 ITS 分子证据。云南山茶(*C. reticulata*)又称滇山茶,主要分布于中国云南、四川等西南地区,是山茶属植物中人们较为熟悉的一个物种,其观赏价值是在历史长河中通过杂交、变异、自然选择以及人工选择而来。云南山茶中的 nrDNA ITS 是否也逃离了一致进化,存在多态性呢?我们以云南山茶为研究对象,采用通用引物对 ITS 片段进行克隆测序,以期达到如下目标:(1)检测 ITS 序列的多态性;(2)推断云南山茶的起源和演化。

1 材料与方法

1.1 实验材料

本实验采集山茶属植物云南山茶(*C. reticulata* Lindl.)的新鲜叶片,使用变色硅胶迅速干燥,常温

条件下保存,供总 DNA 提取。本实验所用植物材料为昆明金殿公园和昆明植物园引种栽培的个体,详见表 1。

表 1 实验材料
Table 1 Experimental materials

类群 Taxon	品种名 Cultivar	简称 Abbreviation	采集地 Locality
云南山茶 <i>Camellia reticulata</i> Lindl.	‘菊瓣’ ‘Juban’	JB	昆明金殿公园
	‘小桂叶’ ‘Xiaoguiye’	XGY	昆明植物园
	‘早桃红’ ‘Zaotaohong’	ZTH	昆明植物园
	‘靖安茶’ ‘Jing’an Cha’	JAC	昆明金殿公园
	‘牡丹茶’ ‘Mudan Cha’	MDC	昆明植物园
	‘云峰茶’ ‘Yunfeng Cha’	YFC	昆明金殿公园
	‘麻叶银红’ ‘Mayeyinhong’	MYYH	昆明植物园
	‘红花油茶’ ‘Honghuayou Cha’	HHYC	昆明金殿公园

1.2 植物总 DNA 提取、扩增与纯化

使用 DNA 提取试剂盒(Takara MiniBEST universal Genomic DNA Extraction Kit)提取样品总 DNA。扩增引物选用通用引物 ITS4: 5'-TCCTC-CGCTTATTGATATGC-3' 和 ITS5: 5'-GGAAGTAAAGTCGTAACAAGG-3'。PCR 扩增反应总体积为 50 μL,反应液包含:1 ×buffer、MgCl₂ 2 mmol/L、dNTP 0.2 mmol/L、引物 0.4 μmol/L、Taq DNA 聚合酶 1 U。扩增程序为:95℃预变性 5 min; 95℃变性 30 s, 56℃退火 30 s, 72℃延伸 30 s, 30 个循环;最后 72℃延伸 7 min。PCR 反应在 ABI Veriti 梯度 PCR 仪上进行。此外,通过添加 4% (V/V) 的 DMSO 优化 PCR 扩增条件,扩增 ITS 片段。并对两种条件下的 PCR 目的产物进行回收纯化。

1.3 PCR 扩增产物的克隆测序

用纯化的 PCR 产物、PMD19-T 载体和连接酶构建重组质粒(16℃)。将重组质粒转入 DH5α 感受态细胞中。过夜培养后,鉴定阳性克隆并测序,同时也进行 PCR 产物直接测序。

1.4 序列处理与分析

所获得序列经校对后,参照 GenBank 上 ITS 序列,确定其边界。然后通过 MEGALIGN 软件中 ClustalW 对序列进行排列,并进行人工校正。本研究对所获得的 ITS 序列进行假基因初步鉴定:(1)利用 BioEdit V5.0.6^[26] 计算各个 ITS 序列长度以及鸟嘌呤(Guanine)和胞嘧啶(Cytosine)的含量(GC content)。(2)通过在线软件 Mfold (<http://>

mfold. rna. albany. edu/? q = mfold/RNA-Folding-Form)测定 5. 8S 区的最小自由能(ΔG , 37℃下的 RNA 模式)。(3) 通过在线软件 HYPERMUT^[27] (<http://www. hiv. lanl. gov/content/sequence/HYPERMUT/hypermur. html>)进行序列的替代模型检测。经过初步判定为 ITS 假基因的序列,用符号‘P’加以区别。在初步确定假基因的基础上,利用 DnaSP^[28]进行核苷酸差异平均值(Average number of nucleotide differences, *K*)、核苷酸多样性(nucleotide diversity, π)和中性检验(Tajima’s *D*)测定。同时,利用 Recombination Detection Program pack-age3b27 软件包中的 GENECONV^[29]进行替代模型下的重组子检测。

为更好地研究云南山茶的系统关系,下载 Gen-Bank 中山茶组(sect. *Camellia*)^[30]部分 ITS 序列以及 *Pyrenaria menglaensis* 为外类群(GenBank 登录号见附录)。应用 Modeltest 3. 7^[31]获得最适核苷酸替代模型为 SYM + I + G 和相关参数估值。采用贝叶斯推断(Bayesian inference, BI)构建相应

的基因树。同时应用 SplitsTree 4 软件,基于 uncorrected p-distances 的 NeighborNet (NN)法构建 NN split 系统发育网络。

2 实验结果

2.1 ITS 序列长度、GC 含量、最小自由能及替代模式分析

本实验从 8 个云南山茶品种中克隆得到 75 条不同的 ITS 序列,其中普通 PCR 条件下得到 44 条 ITS 序列,其序列长度变异范围为 602~624 bp;而优化 PCR 条件后得到的 31 条 ITS 序列长度变异范围为 648~665 bp(详见表 2)。

通过对所有 ITS 序列进行 GC 含量测定发现,普通 PCR 条件下所得序列 ITS1、ITS2 和 5. 8S 区的 GC 含量平均值分别为(54. 9±1. 8)%、(59. 1±2. 8)%、(46. 7±1. 5)%;而在优化 PCR 条件下所得序列的 ITS1、ITS2 和 5. 8S 区的 GC 含量平均值分别为(70. 6±0. 6)%、(75. 4±1. 0)%、(54. 0±0. 4)%。显然,云南山茶 ITS 序列中 ITS2 区 GC 含

表 2 云南山茶 8 个品种的 ITS 序列平均长度、GC 含量以及 5. 8S 区二级结构自由能
Table 2 Average lengths and GC content of ITS regions and free energy of 5. 8S secondary structures of eight *Camellia reticulata* cultivars

品种 Cultivar	类型 Type	数目 No.	长度 Length (bp) (SD)			GC 含量 GC content (%) (SD)			最小自由能 5. 8S ΔG (SD) (kcal/mol)
			ITS1	ITS2	5. 8S	ITS1	ITS2	5. 8S	
‘菊瓣’JB	F	4	251.2(3.2)	247.0(2.8)	164.0(0.0)	70.4(0.6)	75.2(1.0)	53.8(0.6)	-50.8(2.0)
	P	4	229.2(0.5)	213.5(5.0)	164.0(0.0)	55.4(0.7)	58.9(1.9)	46.8(1.4)	-44.6(1.2)
‘小桂叶’XGY	F	3	251.0(0.0)	245.6(1.2)	164.0(0.0)	70.8(0.2)	74.8(0.4)	54.2(0.0)	-52.1(1.6)
	P	5	226.6(4.6)	201.2(29.0)	161.2(6.2)	54.4(1.8)	57.8(3.7)	47.1(1.5)	-40.0(3.4)
‘早桃红’ZTH	F	6	251.6(4.6)	242.3(3.7)	164.0(0.0)	70.1(0.8)	75.1(0.8)	53.8(0.4)	-51.6(0.8)
	P	8	227.2(2.4)	214.0(4.6)	164.0(0.0)	54.8(2.1)	59.2(2.1)	46.7(1.4)	-43.8(3.4)
‘靖安茶’JAC	F	5	247.6(3.8)	237.2(0.8)	164.0(0.0)	71.2(0.2)	76.7(0.2)	54.2(0.6)	-52.4(1.0)
	P	4	228.0(3.3)	213.8(4.8)	164.2(0.5)	54.7(2.4)	59.0(1.8)	47.4(0.7)	-45.4(4.0)
‘牡丹茶’MDC	F	3	253.3(3.8)	238.0(9.6)	164.0(0.0)	70.8(0.2)	74.8(2.0)	54.0(0.4)	-51.8(0.0)
	P	6	225.3(5.6)	193.2(32.4)	164.0(0.0)	54.2(1.2)	59.6(3.0)	46.5(1.2)	-39.6(14.0)
‘云峰茶’YFC	F	3	250.6(2.0)	246.3(2.0)	164.0(0.0)	70.6(0.6)	75.6(0.2)	53.6(0.0)	-51.8(0.0)
	P	8	228.2(1.4)	211.2(0.5)	164.0(0.0)	54.6(0.8)	58.9(2.0)	46.0(1.1)	-43.6(3.4)
‘麻叶银红’MYYH	F	3	252.3(5.7)	239.6(4.6)	164.0(0.0)	70.6(0.6)	76.0(0.8)	54.0(0.4)	-51.3(0.8)
	P	4	230.3(7.4)	212.8(10.2)	164.0(0.0)	57.2(3.4)	60.8(6.8)	48.0(3.2)	-45.4(5.3)
‘红花油茶’HHYC	F	4	252.2(2.5)	246.0(0.0)	164.0(0.0)	70.8(0.8)	74.9(0.5)	54.1(0.3)	-51.8(0.0)
	P	5	228.0(1.8)	215.2(5.3)	164.2(0.4)	54.6(1.2)	58.4(1.6)	46.2(1.3)	-43.2(2.2)
所有序列 All	F	31	251.0(3.6)	242.6(5.0)	164.0(0.0)	70.6(0.6)	75.4(1.0)	54.0(0.4)	-51.7(1.0)
	P	44	227.7(3.6)	209.2(16.6)	163.7(2.1)	54.9(1.8)	59.1(2.8)	46.7(1.5)	-43.1(3.8)

注: F, ITS 功能序列; P, ITS 假基因序列; SD, 标准方差。
Notes: F, Functional ITS sequences; P, ITS Pseudogenes; SD, Standard deviation.

量高于 ITS1 区,优化 PCR 条件下得到的序列 GC 含量明显高于普通 PCR 条件下所得序列。此外,经优化 PCR 条件后得到 ITS 序列存在 GC 含量非常高的区域,如 ITS1 区的 118~182 位点处、ITS2 的 75~105 位点处,而普通 PCR 条件下得到的 ITS 序列却没有类似的区域。

相对于 ITS1 和 ITS2, 5.8S 长度更为保守,因此在比较 ITS 序列最低自由能时,仅对 5.8S 区进行了测定。结果发现,在普通 PCR 条件下获得的 ITS 序列的最低自由能为 (-42.8 ± 3.4) kcal/mol; 而优化 PCR 条件获得的 ITS 序列的最低自由能为 (-51.7 ± 1.0) kcal/mol。显然,在普通 PCR 条件下获得的 ITS 序列的最低自由能要高于 PCR 条件优化后获得的 ITS 序列。在普通 PCR 条件下获得的 MYYH_P4 总长 624 bp, 具备较高的 GC 含量 (ITS1、ITS2 和 5.8S 区的 GC 含量分别为: 62.08%、70.45% 和 52.44%), ΔG 也较低 (-53.0) kcal/mol, 但 5.8S 区存在多处碱基变异,同时 ITS1 和 ITS2 区存在多处与假基因类似的碱基缺失和碱基变异。以上结果表明,通过普通 PCR 条件获得的 ITS 序列均为假基因拷贝,而通过优化 PCR 条件获得的 ITS 序列均为功能拷贝。

以对照序列为基准对所获得序列进行了核苷酸替代模式分析(见图 1)。ITS 功能拷贝之间的核苷酸突变主要发生在 ITS1 和 ITS2 区,其中 G→A、C→T 的替代占有所有这些核苷酸替代的 22.9%; 而可疑的假基因明显积累了更多的突变,并且随机分散在 ITS1、5.8S 及 ITS2 三个区域,其中 G→A、C→T 的替代占 49.9%。

2.2 核苷酸多样性及中性检测

如表 3 所示,所有可疑假基因序列的核苷酸多样性(π)和平均核苷酸差异数(K)明显高于功能序列,假基因序列的 π 值和 K 值分别为功能序列的 5.6 倍和 4.2 倍。对于所有假基因序列,各个区的核苷酸多样性和平均核苷酸差异数值高低排列均为 ITS1 > ITS2 > 5.8S。

本研究对所得序列进行 Tajima's D 检验,发现功能序列及假基因序列的 D 值均为负值(表 3), ITS 功能序列 D 值的绝对值较大,且 5.8S 区呈现显著性差异($p < 0.05$),表明 ITS 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,表明假基因序列倾向于中性进化^[32,33]。

2.3 重组鉴定及系统分析

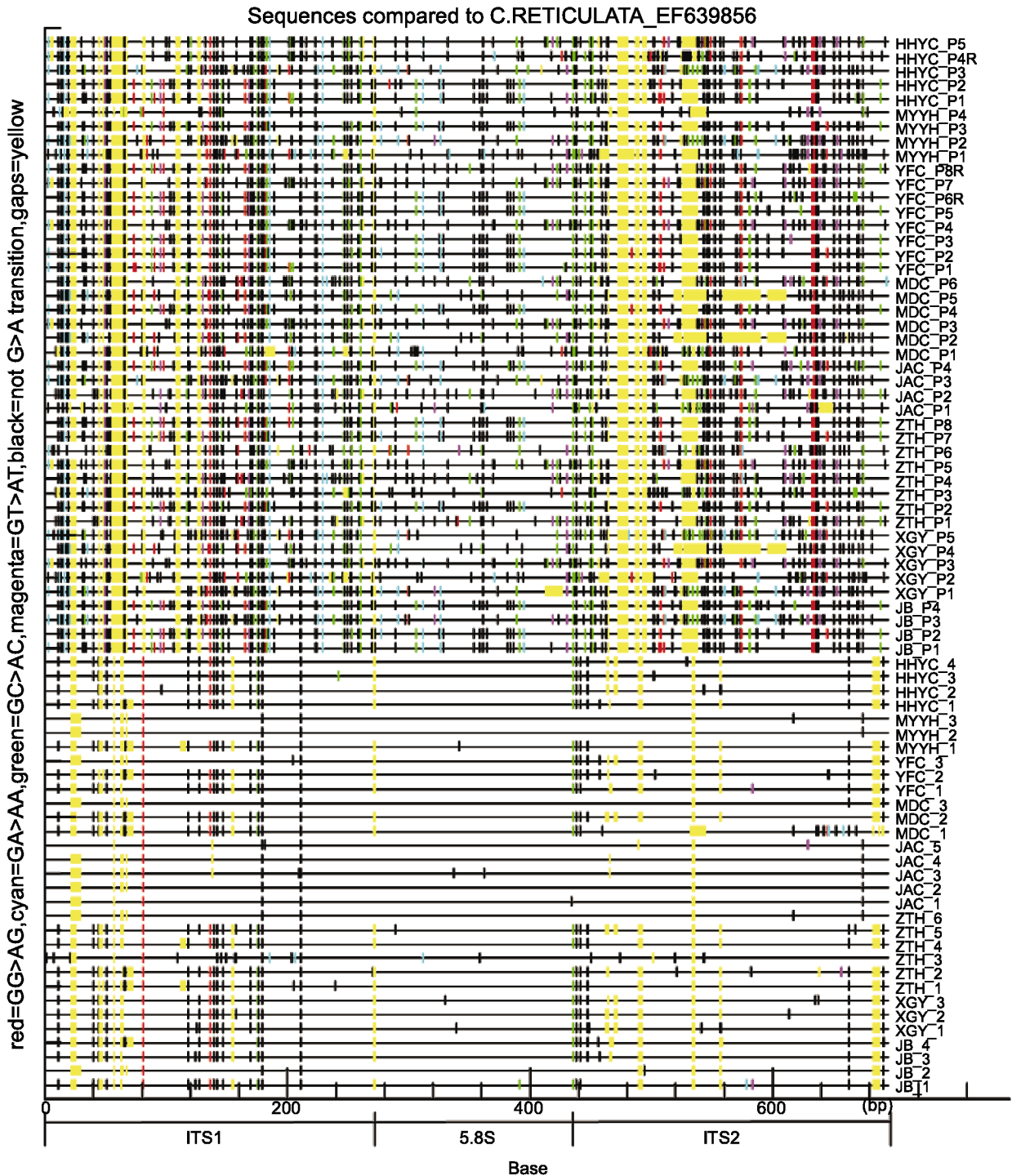
应用 GENECONV 对所得序列进行重组检测发现,可疑假基因 YFC_P6、YFC_P8、HHYC_P4 被判断为重组子,其重组亲本分别为 YFC_P5 和 ZTH_P3、ZTH_P1 和 JB_P1、MDC_P2 和 HHYC_P5,可疑重组子在图 1 中用 R 标记。由于重组子可能会影响进化树的结构,甚至掩盖其真正的亲缘关系,在构建系统树时应除去重组子^[34]。

在基于所有 ITS 序列构建的系统树中,除 MYYH_P4 拷贝外,云南山茶 ITS 功能序列与 ITS 假基因分离开来(图 2)。在 ITS 功能序列形成的大组(I)中,不同品种的功能序列并没有完全聚为一支,而是以较高的支持率分别聚在两支上(A 和 B),记为 A、B 两种类型。其中,A 支包含云南山茶和西南山茶(*Camellia pitardii*),显示了较亲密的关系;B 支包含云南山茶、毛蕊山茶(*Camellia mairei*)、石果毛蕊山茶(*Camellia mairei* var. *lapidea*)、怒江山茶(*Camellia saluenensis*)、多齿山茶(*Camellia polyodonta*)和全缘叶山茶(*Camellia subintegra*),并含有 MYYH_P4 拷贝。其中,云南山茶的两个品种‘红花油茶’和‘小桂叶’的 ITS 拷贝全部分布在 A 支中,‘靖安茶’的 ITS 拷贝全部分布在 B 支中,而来自其余 5 个品种的 ITS 功能序列则在 A、B 两支中都有分布。NNsplit 网状进化分析结果与 Bayes 系统发育重建结果相互印证(图 3)。在云南山茶假基因拷贝形成的单系中(II),大多数拷贝的分支都显示出较高的后验概率。但是同一个体内的 ITS 假基因序列并非单系,序列分散在分支 II 中不同的小分支中。

3 讨论

3.1 云南山茶 ITS 多态性

近年来,nrDNA ITS 序列多态性现象在越来越多的植物中被发现,如苏铁属(*Cycas*)^[10]、梨属(*Pyrus*)^[11]、栎属(*Quercus*)^[12]等。多倍化、异源多倍化、杂交以及基因家族在染色体上的结构等因素都可能导致一致进化不完全^[12,35]。本研究也发现云南山茶中 ITS 序列在基因组内存在明显的多态



(1)JB: ‘菊瓣’; XGY: ‘小桂叶’; ZTH: ‘早桃红’; JAC: ‘靖安茶’; MDC: ‘牡丹茶’; YFC: ‘云峰茶’; MYYH: ‘麻叶银红’; HHYC: ‘红花油茶’。

(2)序列后字母‘P’代表该序列为假基因拷贝; 序列后数字代表不同的拷贝。

(1)JB: ‘Juban’; XGY: ‘Xiaoguiye’; ZTH: ‘Zaotaohong’; JAC: ‘Jing’an Cha’; MDC: ‘Mudan Cha’; YFC: ‘Yunfeng Cha’; MYYH: ‘Mayeyinhong’; HHYC: ‘Honghuayou Cha’。

(2)‘P’ indicates that the sequence of the gene is pseudogene; the number represents a different copy.

图 1 以云南山茶中的功能拷贝 EF639856 为 ITS 功能性参照序列,
云南山茶 8 个品种中所有 ITS 序列替代位点的分布示意图

Fig. 1 Schematic illustration of the distribution of substitution sites across the entire ITS region obtained from eight *C. reticulata* cultivars using the functional sequence of *C. reticulata* (EF639856) as reference

表 3 云南山茶 ITS 核苷酸多样性
Table 3 Nucleotide diversity of the ITS regions in *Camellia reticulata*

ITS 区域 ITS region	序列数目 Number of sequences	多态位点数 Polymorphic sites	总突变数 Total mutation	核苷酸差异平均值 Average number of nucleotide differences (K)	核苷酸多态性 Nucleotide diversity(π)	中性检验 Neutrality test (Tajima's D)
ITS1 region						
All functional	31	40	43	10.024	0.04212	-0.25281
All pseudogenes	44	131	174	34.646	0.17587	-0.48984
ITS2 region						
All functional	31	36	38	5.420	0.02463	-1.57101
All pseudogenes	44	90	116	21.637	0.17037	-0.68434
5.8S region						
All functional	31	15	15	1.673	0.01020	-1.85499*
All pseudogenes	44	76	87	17.394	0.11596	-0.46882
Entire region						
All functional	31	91	94	16.527	0.02674	-1.12935
All pseudogenes	44	294	369	70.501	0.15000	-0.62374

注：中性检验值，*： $p < 0.05$ ；无*： $p > 0.10$ 。
Note: Tajima's D, *： $p < 0.05$ ；no *： $p > 0.10$ 。

性。一方面,ITS 功能序列存在广泛变异,长度变异范围为 648~665 bp,并且在受试样品中普遍存在两种类型的 ITS 功能拷贝(A、B 两种类型),如‘早桃红’中得到 4 条不同的 A 类型序列,2 条不同的 B 类型序列等。另一方面,云南山茶个体内存在大量的 ITS 假基因,这些假基因彼此间的差异较大(表 3)。显然,云南山茶 nrDNA ITS 逃离了一致进化。云南山茶个体内功能性 ITS 序列为何会呈现多态性呢?根据云南山茶的遗传背景分析,主要有以下几种可能:

(1)杂交起源是形成云南山茶 ITS 拷贝多态性的原因。有研究认为云南山茶可能是西南山茶和怒江山茶自然杂交的异源多倍体复合体。这些研究表明云南山茶多倍体复合体与西南山茶、怒江山茶、石果毛蕊山茶、南山茶(*C. semiserrata*)和多齿山茶关系较近。其中,云南山茶四倍体由云南山茶二倍体和西南山茶杂交形成,云南山茶六倍体由云南山茶二倍体、西南山茶和怒江山茶杂交形成,并含有华东山茶(*C. japonica*)的基因组成分^[36,37]。在本研究中,根据系统发育重建结果(图 2)以及 NN split 网络分析(图 3),云南山茶 ITS 功能序列并没有形成一个单系,ITS 拷贝分散在两个分支中,其中 A 支是由云南山茶和西南山茶构成的单系分支,显示了密切的亲缘关系;B 支则包含云南山茶、毛蕊山茶、石

果毛蕊山茶、怒江山茶、多齿山茶和全缘叶山茶。此外,从分支结构可以看出,同一个体内的不同 ITS 拷贝并不是完全聚在一支上,而是分散聚在 A、B 两支上。如‘早桃红’、‘麻叶银红’、‘云峰茶’、‘牡丹茶’和‘菊瓣’同一个体内不同的拷贝分散聚在 A、B 两支上,推断西南山茶可能是其亲本之一,另一亲本则来自毛蕊山茶和怒江山茶中的一个或多个,这与刘利勤(2009)^[37]的研究结果较为一致。就此可推断云南山茶多倍体种是通过杂交起源的,并且其它物种的基因渗透也促进了云南山茶的遗传多样性。而‘红花油茶’、‘靖安茶’和‘小桂叶’这三个云南山茶品种只检测到一种类型的 ITS 拷贝。

(2)本研究中 ITS 多态性的起源还可能来源于同一细胞内不同 ITS 拷贝的基因结构和座位的不同。有研究表明,云南山茶的 rDNA 位点较多,且大都不在同一染色体上^[38],而位点之间的基因重组的频率要比位点内的重组低得多^[1,39],因而在 rDNA 一致进化过程中,部分 ITS 难以与其它位点基因进行不等交换和重组等过程,这在一定程度上阻碍了一致进化的进程,很可能造成位点之间的 ITS 拷贝的不同。甚至有些位点变成了假基因,逃离了一致进化,从而产生了个体内 ITS 多态性。这种推断还需要更多的研究将 ITS 拷贝定位在染色体上,才能取得更加直接的证据,方可加以证明。

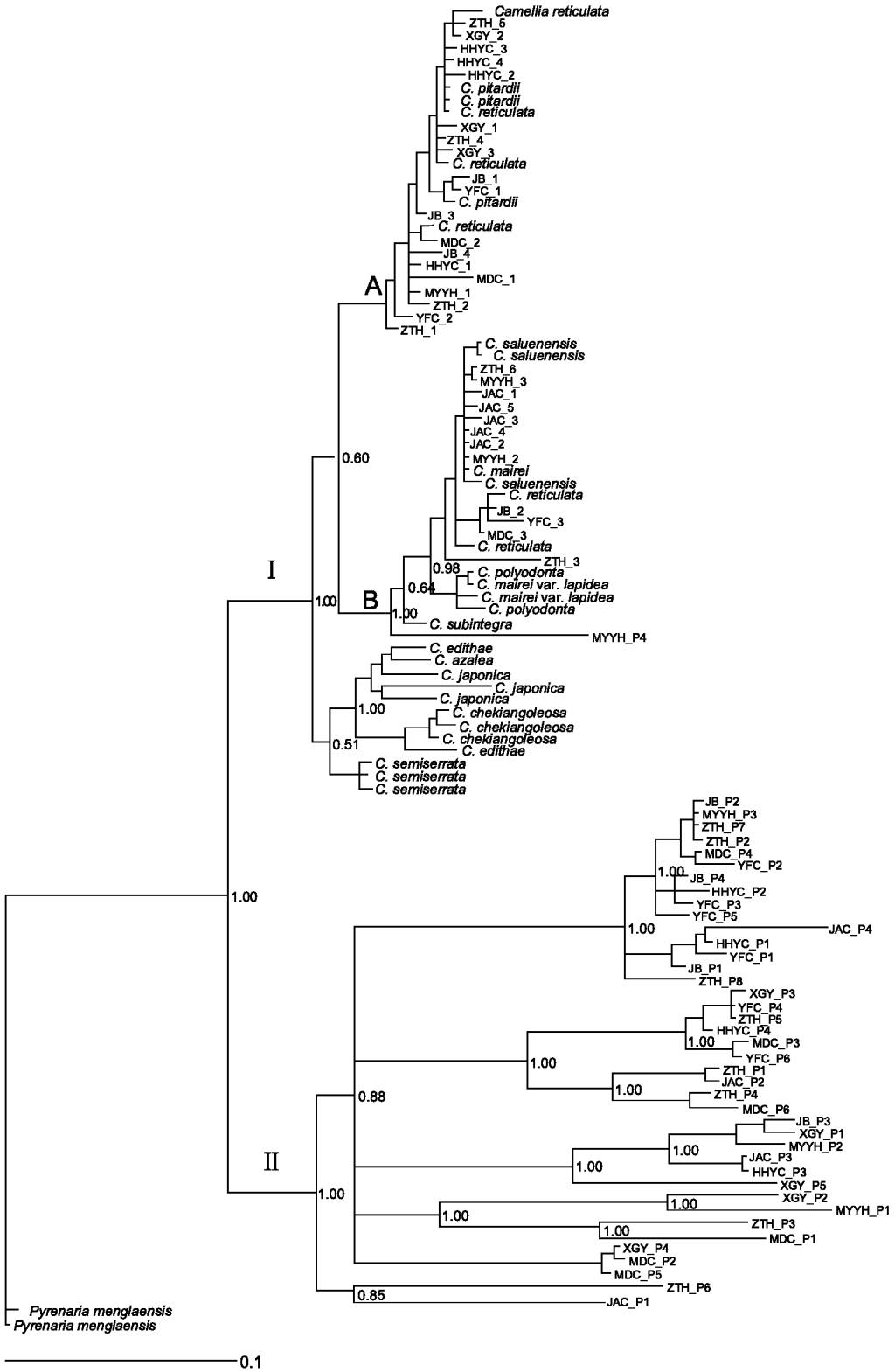


图 2 基于 ITS 序列构建的 Bayes 系统进化树

(节点上的数字为贝叶斯后验概率;图中相关内容信息参阅图 1 注解)

Fig.2 Bayes tree based on ITS sequences data (numbers at the node are Bayesian Posterior Probabilities (BPP); see Figure 1 for graph information)

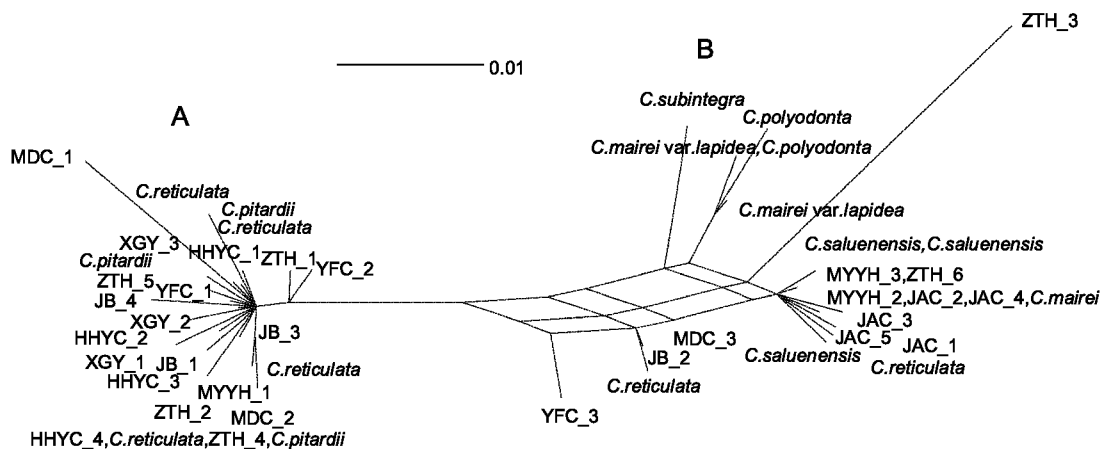


图3 基于 uncorrected p-distances 的 NeighborNet 法构建的 NN split 系统发育网络
(图中相关内容信息参阅图1注解)

Fig.3 NeighborNet split network based on uncorrected p-distances
(see Figure 1 for graph information)

3.2 云南山茶 ITS 假基因

nrDNA ITS 基因家族有成千上万的拷贝,在进化过程中一些拷贝可能会功能退化而变成假基因。本研究获得的 ITS 假基因大都具有以下特性:(1)假基因拷贝在长度上存在约 40 bp 的缺失。(2)5.8 区长度较为保守,除 XGY_P1 存在 14 bp 的缺失以及 JAC_P3 和 HHYC_P3 存在 1 bp 的插入,其余序列 5.8S 为完整的 164 bp,但都具有较高的核苷酸变异(图 1),高度保守的 5.8S 区存在如此高的变异正是假基因的特征之一。(3)GC 含量明显偏低,表明其 GC 位点碱基突变频繁;5.8S 二级结构最小自由能明显偏高,表明其二级结构稳定性显著降低(表 2)。由于 GC 含量和二级结构稳定性都明显下降,因而在普通 PCR 条件下因首先熔解而被优先扩增。(4)存在较高的核苷酸替代速率(图 1),其中 G→A、C→T 的替代占 49.9%,由此导致 GC 含量显著降低,可能与基因功能的缺失和假基因的起源有关^[5]。(5)进化脱离了功能限制,在中性进化模式下以较高的替代速率进化,不像功能序列要维持特定的二级结构需要进行补偿性突变(compensatory mutation)^[40]。然而,也存在较特殊的假基因序列,如 MYH_P4 拷贝的 GC 含量较其它假基因高,5.8S 区的自由能也较低。但是,将该序列与其它功能序列比对,MYH_P4 拷贝在长度上短约 22 bp,5.8S 区存在 6 bp 的变异(功能拷贝仅存在 0~2 bp 的变异),同时该拷贝存在多处与

假基因类似的碱基变异和碱基缺失(图 1),因此也被鉴定为假基因。在系统树上(图 2),MYH_P4 假基因拷贝与云南山茶功能序列的 A 支聚在一起,表明 MYH_P4 序列与该类型的 ITS 序列具有较近的亲缘关系,这似乎说明 MYH_P4 序列从功能基因正向假基因转化,同时也提示了应用 ITS 假基因进行系统发育分析可能会对其真实的系统关系造成影响。

云南山茶 ITS 假基因同样具有非常高的多态性,同一品种的 ITS 假基因拷贝分散聚在不同的小分支上(图 2)。由于所获得的 ITS 假基因以中性模式进化,与功能基因的进化速率不同,因此在系统重建结果中并不能反映 ITS 假基因真实的系统位置,但可以明显地看出,除 MYH_P4 假基因拷贝外,其它 ITS 假基因拷贝均与功能序列分离。在这些假基因中,不同品种之间存在彼此差异很小甚至完全一致的 ITS 假基因拷贝。如 ZTH_P1 和 JAC_P2 拷贝之间仅存在 4 bp 的差异;MDC_P3 和 YFC_P7 拷贝仅存在 3 bp 的差异;JB_P4 和 YFC_P3 拷贝仅存在 5 bp 的差异;ZTH_P5 和 YFC_P4 序列相同,XGY_P3 和 ZTH_P5、YFC_P4 拷贝之间仅存在 2 bp 的差异;HHYC_P3 和 JAC_P3 序列完全相同。考虑到假基因往往被认为以较高的速率进行独立的中性进化^[39],上述不同品种之间存在的相同或差异很小的中性进化的假基因拷贝现象表明,这些相近的假基因可能在近期通过品种之间的杂交等方式互

相融合,在供试品种中存在时间较短,暗示供试品种形成时间也可能较短。由此看来,广泛研究这些假基因在云南山茶中的保留状况、起源和分歧时间,可能会提供一些有意义的进化史信息。

参考文献:

- [1] Eickbush T H, Eickbush D G. Finely orchestrated movements: evolution of the ribosomal RNA genes [J]. *Genetics*, 2007, 175(2): 477–485.
- [2] 田欣,李德铎. DNA 序列在植物系统学研究中的应用[J]. *云南植物研究*, 2002, 24(2): 170–184.
- [3] Álvarez I, Wendel J F. Ribosomal ITS sequences and plant phylogenetic inference [J]. *Mol Phylogenet Evol*, 2003, 29(3): 417–434.
- [4] Li D Z, Gao L M, Li H T, 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 USA*, 2011, 108(49): 19641–19646.
- [5] Bailey C D, Carr T G, Harris S A, et al. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes [J]. *Mol Phylogenet Evol*, 2003, 29(3): 435–455.
- [6] Kovarik A, Pires J C, Leitch A R, et al. Rapid concerted evolution of nuclear ribosomal DNA in two tragopogon allopolyploids of recent and recurrent origin [J]. *Genetics*, 2005, 169(2): 931–944.
- [7] Hillis D M, Moritz C, Porter C A, et al. Evidence for biased gene conversion in concerted evolution of ribosomal DNA [J]. *Science*, 1991, 251(4991): 308–310.
- [8] Gonzalez I L, Sylvester J E. Human rDNA: Evolutionary patterns within the genes and tandem arrays derived from multiple chromosomes [J]. *Genomics*, 2001, 73(3): 255–263.
- [9] Kovari A, Matyasek R, Lim K Y, et al. Concerted evolution of 18–5.8–26S rDNA repeats in *Nicotiana* allotetraploids [J]. *Biol J Linn Soc*, 2004, 82(4): 615–625.
- [10] Xiao L Q, 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(1): 168–177.
- [11] Zheng X, Cai D, Yao L, et al. Non-concerted ITS evolution, early origin and phylogenetic utility of ITS pseudogenes in *Pyrus* [J]. *Mol Phylogenet Evol*, 2008, 48(3): 892–903.
- [12] Muir G, Fleming C C, Schlötterer C. Tree divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. [J]. *Mol Biol Evol*, 2001, 18(2): 112–119.
- [13] Wei X X, Wang X Q. Recolonization and radiation in *Larix* (Pinaceae): evidence from nuclear ribosomal DNA paralogues [J]. *Mol Ecol*, 2004, 13(10): 3115–3123.
- [14] 肖龙骞,朱华. 植物 nrDNA ITS 致同进化不完全现象及其进化意义 [J]. *西北植物学报*, 2009, 29(8): 1708–1713.
- [15] 马长乐,周浙昆. ITS 假基因对栎属系统学研究的影响及其对分子系统学研究的启示 [J]. *云南植物研究*, 2006, 28(2): 127–132.
- [16] Wei X X, Wang X Q, Hong D Y. Marked intragenomic heterogeneity and geographical differentiation of nrDNA ITS in *Larix potaninii* (Pinaceae) [J]. *J Mol Evol*, 2003, 57(6): 623–635.
- [17] Li R, Yang J B, Yang S X, et al. Phylogeny and taxonomy of the *Pyrenaria* complex (Theaceae) based on nuclear ribosomal ITS sequences [J]. *Nord J Bot*, 2011, 29(6): 780–787.
- [18] Vijayan K, Zhang W J, Tsou C H. Molecular taxonomy of *Camellia* (Theaceae) inferred from nrITS sequences [J]. *Am J Bot*, 2009, 96(7): 1348–1360.
- [19] 杨俊波,李洪涛,杨世雄,等. 四个 DNA 片段在山茶属分子系统学研究中的应用 [J]. *云南植物研究*, 2006, 28(2): 108–114.
- [20] 田敏,李纪元,倪穗,等. 基于 ITS 序列的红山茶组植物系统发育关系的研究 [J]. *园艺学报*, 2008, 35(11): 1685–1688.
- [21] 唐绍清,施苏华,钟扬,等. 基于 ITS 序列探讨山茶属金花茶组的系统发育关系 [J]. *广西植物*, 2004, 24(6): 488–492.
- [22] 唐绍清,钟扬,施苏华,等. *Camellia nitidissima* 与 *C. petelotii* 之间的关系研究——来自 nrDNA ITS 的证据 [J]. *武汉植物学研究*, 2001, 19(6): 449–452.
- [23] 徐颖,徐晶,高继银,等. 山茶属植物 ITS 的多态性——一个广泛逃离一致性进化的实例 [J]. *植物学报*, 2011, 46(2): 162–169.
- [24] Vijayan K, Tsou C H. Technical report on the molecular phylogeny of *Camellia* with nrITS: the need for high quality DNA and PCR amplification with

Pfu-DNA polymerase[J]. *Bot Stud*, 2008, 49(2): 177–188.

[25] 徐晶. 茶 16S–26S rDNA 多态性与进化[D]. 上海: 复旦大学生命科学学院, 2008.

[26] Hall T A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT[J]. *Nucleic Acids Symp Ser*, 1999, 41: 95–98.

[27] Rose P R, Korber B T. Detecting hypermutations in viral sequences with an emphasis on G→A hypermutation[J]. *Bioinformatics*, 2000, 16(4): 400–401.

[28] Rozas J, Rozas R. DnaSP version 3, an integrated program for molecular population genetics and molecular evolution analysis [J]. *Bioinformatics*, 1999, 15(2): 174–175.

[29] Padidam M, Sawyer S, Fauquet C M. Possible emergence of new geminiviruses by frequent recombination[J]. *Virology*, 1999, 265(2): 218–225.

[30] Wu Z Y, Hong D Y. Flora of China: Vol. 12 [M]. Beijing: Science Press, 2007: 367–412.

[31] Posada D, Crandall K A. Modeltest: testing the model of DNA substitution [J]. *Bioinformatics*, 1998, 14(9): 817–818.

[32] Fu Y X. New statistical tests of neutrality for DNA samples from a population [J]. *Genetics*, 1996, 143(1): 557–570.

[33] Fu Y X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection [J]. *Genetics*, 1997, 147(2): 915–925.

[34] McDade L A. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis [J]. *Evolution*, 1992, 46(5): 1329–1346.

[35] 肖龙骞, 朱华. 苏铁 nrDNA ITS 区的序列多态性: 不完全致同进化的证据[J]. 生物多样性, 2009, 17(5): 476–481.

[36] 夏丽芳, 顾志健, 王仲朗, 等. 探讨云南山茶起源的一线曙光——野生二倍体类型在金沙江流域的发现[J]. 云南植物研究, 1994, 16(3): 255–262.

[37] 刘利勤. 滇山茶多倍体复合体起源和演化的分子细胞遗传学研究[D]. 昆明: 中国科学院昆明植物研究所, 2009.

[38] Gu Z J, Xiao H. Physical mapping of the 18S–26S rDNA by fluorescent *in situ* hybridization (FISH) in *Camellia reticulata* polyploidy complex (Theaceae) [J]. *Plant Sci*, 2003, 164(2): 279–285.

[39] Zhang D M, Sang T. Physical mapping of ribosomal RNA genes in peonies (*Paeonia*, Paeoniaceae) by fluorescent *in situ* hybridization: implications for phylogeny and concerted evolution [J]. *Am J Bot*, 1999, 86(5): 735–740.

[40] Petrov D A, Hartl D L. Pseudogene evolution and natural selection for a compact genome [J]. *J Heredity*, 2000, 91(3): 221–227.

附录 (Appendix)

Taxon	GenBank accession NO.	Taxon	GenBank accession NO.
<i>Camellia reticulata</i>	HM061380, HM061398, HM061392, HM061397, HM061401, EF639856	‘早桃红’ ZTH	
<i>Camellia japonica</i>	EF639856, EU579723, HM061428	Functional	JX457339, JX500464 ~ JX500468
<i>Camellia semiserrata</i>	EF649688, EF649691, EU579770	Pseudogene	JX457338, JX457340, JX500496 ~ JX500501
<i>Camellia subintegra</i>	EU579776	‘靖安茶’ JAC	
<i>Camellia saluenensis</i>	EU579768, HM061345, EU579767	Functional	JX500469 ~ JX500473
<i>Camellia polyodonta</i>	EF646285, EU579760	Pseudogene	JX500502 ~ JX500505
<i>Camellia pitardii</i>	FJ432115, EF646288, EU579759	‘牡丹茶’ MDC	
<i>Camellia mairei</i>	EU579686	Functional	JX500474 ~ JX500476
<i>Camellia edithae</i>	EU579696, EU579695	Pseudogene	JX500506 ~ JX500511
<i>Camellia azalea</i>	EU579681	‘云峰茶’ YFC	
<i>Pyrenaria menglaensis</i>	EU579798, AF456266	Functional	JX500477 ~ JX500479
‘菊瓣’ JB		Pseudogene	JX500512 ~ JX500519
Functional	JX500457 ~ JX500460	‘麻叶银红’ MYYH	
Pseudogene	JX500487 ~ JX500490	Functional	JX500480 ~ JX500482
‘小桂叶’ XGY		Pseudogene	JX500520 ~ JX500522, JX666613
Functional	JX500461 ~ JX500463	‘红花油茶’ HHYC	
Pseudogene	JX500491 ~ JX500495	Functional	JX500483 ~ JX500486
		Pseudogene	JX500523 ~ JX500527