

DOI: 10.11931/guihaia.gxzw201711036

引文格式: 成彩霞, 苏雪, 高婷, 等. 美国西部不同地区 *Grayia spinosa* 核 DNA ITS 序列分析 [J]. 广西植物, 2018, 38(5): 617-625
CHENG CX, SU X, GAO T, et al. nrDNA ITS sequence analysis of *Grayia spinosa* (Chenopodiaceae) from different regions in the Western United States [J]. *Guihaia*, 2018, 38(5): 617-625

美国西部不同地区 *Grayia spinosa* 核 DNA ITS 序列分析

成彩霞, 苏雪*, 高婷, 周璇

(西北师范大学 生命科学学院, 兰州 730070)

摘要: 藜科植物 *Grayia spinosa* 是美国西部地区的特有种, 多生长在干旱盐碱地, 具有重要的生态价值。该研究测定了采自美国西部犹他州 *G. spinosa* 的 nrDNA ITS 序列, 与 GenBank 中已提交的 *G. spinosa* 的所有 ITS 序列以及 *G. spinosa* 的四个近缘种作为外类群进行比较, 分析了美国西部不同地区 *G. spinosa* ITS 序列的一级结构与其 RNA 二级结构的变异规律。结果表明: 所有 *G. spinosa* 样品的 nrDNA ITS 序列长度在 611~623 bp 之间, GC 含量在 60.35%~61.0% 之间, 序列间共存在 22 个变异位点, 5 个为简约信息位点。各样品间的遗传距离在 0.001 8~0.008 9 之间, 不同样品间的遗传距离与地理距离的相关性不显著。邻接法构建的系统发育树显示所有 *G. spinosa* 聚为一大支, 与外类群形成明显分支。此外, 利用 RNAfold 在线软件预测了 *G. spinosa* ITS 序列的 RNA 二级结构, 将 8 个 *G. spinosa* 样品的 RNA 二级结构根据构型差异大体上分为四类, 分别记为 type A, B, C 和 D 四类, 主要变异出现在 ITS1 和 ITS2 区。所不同的是在 *G. spinosa* ITS 的一级结构分析中 GSNE1 与 GSWA8 体现出更近的亲缘关系, 但二者的 RNA 二级结构差异明显, 同时 GSNE2、GSUT3、GSUT4、GSCA5、GSCA6、GSCO7 在 ITS 序列一级结构分析中也体现出较近的亲缘关系, 但是他们的 RNA 二级结构差异明显。这可能与 ITS 序列的 RNA 二级结构在进化中体现出更大的保守性有关。

关键词: *Grayia spinosa*, ITS 序列变异, 二级结构, 藜科

中图分类号: Q949.745.1 文献标识码: A 文章编号: 1000-3142(2018)05-0617-09

nrDNA ITS sequence analysis of *Grayia spinosa* (Chenopodiaceae) from different regions in the Western United States

CHENG Caixia, SU Xue*, GAO Ting, ZHOU Xuan

(College of Life Sciences, Northwest Normal University, Lanzhou 730070, China)

Abstract: *Grayia spinosa*, a species endemic to the Western United States, is mainly distributed in arid and saline areas and has important ecological value. ITS sequence of *G. spinosa* collected from Utah State of the United States was sequenced, and aligned with other sequences from GenBank of this species by Blast. The relationship among *G. spinosa* from the West United States were analyzed based on ITS sequences and four related species of *Grayia* were selected as

收稿日期: 2018-03-23

基金项目: 国家自然科学基金(31260054) [Supported by the National Natural Science Foundation of China (31260054)].

作者简介: 成彩霞(1990-), 女, 甘肃秦安人, 硕士研究生, 研究方向为系统与进化植物学, (E-mail) 1395101933@qq.com。

*通信作者: 苏雪, 副教授, 硕士, 主要从事植物繁殖生态学研究, (E-mail) suxue@nwnu.edu.cn。

outgroup. ITS sequence variation of *G. spinosa* and its RNA secondary structure was analyzed in different regions of the United States. The results showed that the length of ITS sequences ranged from 611 bp to 623 bp and the GC contents was 60.35%–61.0%. There were 22 variants sites and five parsimony informative sites in the ITS sequences of *G. spinosa* from different regions in the United States. Genetic distance ranged from 0.001 8 to 0.008 9 between samples, the correlation was not significant between genetic distances and geographical distances of different samples. The phylogenetic tree constructed by Neighbor-joining method indicated that all samples of *G. spinosa* were clustered together, forming a distinct branch with the outer group. In addition, the RNA secondary structure of *G. spinosa* ITS sequence was predicted by RNA fold online software. The RNA secondary structure of *G. spinosa* samples were roughly divided into four types according to the configuration differences, which were named type A, B, C and D. Unlike the primary structure analysis of *G. spinosa* ITS sequences, RNA secondary structure did not reflect the more related phylogenetic relationships between GSNE1 and GSWA8. GSUT3, GSUT4, GSCA5, GSCA6 and GSCO7 indicated closer phylogenetic relationships in the primary structure analysis of ITS sequences, but their RNA secondary structures were significantly different, which was related to the higher conservation of RNA secondary structure.

Key words: *Grayia spinosa*, ITS sequence variation, secondary structure, Chenopodiaceae

藜科植物是被子植物的大科之一,是干旱、荒漠生态环境中的主要植物类群,全球共有 130 属,约 1 500 种。我国的藜科植物资源相当丰富,尤其是在以干旱、盐碱土为主的西北部大面积生态脆弱区,藜科植物成为其植被的最重要构成部分之一,具有十分重要的生态维持与建设价值。参考中国植物志英文版发现,猪毛菜属、盐爪爪属、碱蓬属等为我国西北地区的主要藜科植物,这些类群大多为草本,植株低矮、叶片较小,花很小无观赏性,生长于农田周围或城市绿化区域,而且多以杂草对待。*Grayia spinosa* 为藜科中少有的灌木,植物体较大,寿命长,花被红色、大而鲜艳,不同于大多数藜科植物,具有一定的观赏价值。此外,*G. spinosa* 耐干旱、盐碱及贫瘠的能力极强,是植被恢复重建的极好材料,但仅分布于北美,美国西部是它最重要的分布区。

近年来,随着分子标记技术的迅速发展,DNA 序列分析被广泛应用于植物系统发育分析和遗传多样性研究。与此同时,由这些分子标记所揭示的 DNA 等生物大分子的变异式样及特征成为了研究了解一个物种的最为基本和重要的环节,并为进一步的科学研究提供基础。核 DNA ITS 序列是介于 18S 至 26S 之间的非编码内转录间隔区,包含一个 5.8S 编码区,由于 ITS 序列在进化过程中受到的选择压力小且进化速率较快等特点,被广泛用于属及

种水平等较低分类阶元亲缘关系和系统进化研究 (Jaehav & Shama, 2014; Liang & Wu, 2017; Guo et al, 2016)。近年来,ITS 序列也开始应用于种内系统关系的研究和药用植物真伪品鉴定等方面 (蒋玲艳和郭志刚, 2009; Chen et al, 2010; Khan et al, 2014)。同时,ITS 序列虽然在部分区域出现较高的突变率,却可能会通过碱基补偿替换等方式抵消这些位点的突变对二级结构的干扰,使得在进化过程中其 RNA 二级结构体现出更大的保守性。因此,本研究通过对美国西部不同地区 *G. spinosa* 的 ITS 序列一级结构及其 RNA 二级结构的分析,以期了解以下问题:(1)不同 *G. spinosa* 样品 ITS 序列的变异规律。(2)*G. spinosa* ITS 一级结构与其 RNA 二级结构的变化是否一致?

1 材料

鉴于该研究的主要目的是探讨 *G. spinosa* 分布范围内不同区域 ITS 序列的变异样式,取样地点的选择以涵盖其主要分布区为准,并且在每个地区尽可能取 2~3 个不同地点的材料,凡 NCBI 已有记录的样点不再另行采样而直接下载其序列的相关信息 (包括分布在科罗拉多、加利福尼亚、内华达、华盛顿和犹他州等 *G. spinosa* 主要分布区的所有 *G. spinosa* ITS 序列共 7 个样品)。样品 GSUT4 由朱

格麟教授采自犹他州一野外居群, 叶片干燥后测序。另外, 选择与 *G. spinosa* 亲缘关系较近的 *G. brandegeei*、*Holmbergia tweedii*、*Atriplex joaquinana* 和 *Chenopodium fremontii* 为外类群, 它们的来源、产地和序列登录号见表 1。

表 1 研究材料来源
Table 1 Resources of materials

样品编号 Sample code	物种 Species	采样地点 Sampling location	GenBank 登录号 GenBank accession	作者 Author
GSNE1	<i>Grayia spinosa</i>	Nevada	AM849224	Heklau H
GSNE2	<i>G. spinosa</i>	Nevada	HM005844	Zacharias EH
GSUT3	<i>G. spinosa</i>	Utah	HM005843	Zacharias EH
GSUT4	<i>G. spinosa</i>	Utah	MG584455	Su, X
GSCA5	<i>G. spinosa</i>	California	HE577356	Fuentes S
GSCA6	<i>G. spinosa</i>	California	HM587572	Kadereit G
GSC07	<i>G. spinosa</i>	Colorado	HM587571	Kadereit G
GSWA8	<i>G. spinosa</i>	Washington	AM849225	Heklau H
GBCO9	<i>G. brandegeei</i>	Colorado	HM587570	Kadereit G
HTU10	<i>Holmbergia tweedii</i>	USA	HM005841	Zacharias EH
HTU11	<i>Atriplex joaquinana</i>	USA	HM005852	Zacharias EH
CFCA12	<i>Chenopodium fremontii</i>	California	HE577408	Fuentes S

注: 以下分析均用样品编号代替各样品(编号中的四个字母分别代表属名的首字母, 种名的首字母, 采样地点地名的前两个字母, 其中 *Holmbergia tweedii* 与 *Atriplex joaquinana* 由于作者未提交具体地名, 用“U”表示 USA)。下同。

Note: Following analysis will use sample code instead of sample (There are four letters which represent the initials of generic name, the species name and the first two letters of the region name, as the author did not indicate the location of *Holmbergia tweedii* and *Atriplex joaquinana*, “U” means USA). The same below.

2 方法

2.1 基因组 DNA 的提取

对采自犹他州的 *G. spinosa* 样品用变色硅胶干燥储藏, 带回实验室后采用改良的 CTAB 法 (Doyle, 1987) 提取基因组总 DNA。

2.2 ITS 序列的 PCR 扩增

采用 ITS 通用引物, 即 ITS1 (5' AGA AGT CGT

AAC AAG GTT TCC GTA GC 3') 和 ITS4 (5' TCC TCC GCT TAT TGA TAT GC 3')。PCR 反应总体系为 20 μ L, 其中 PCR Mix (宝生物公司) 10 μ L, 上下游引物 (20 μ mol \cdot L⁻¹) 及 DNA 模板各 1 μ L, 加 ddH₂O 至 20 μ L。扩增程序: (1) 70 $^{\circ}$ C 预变性 4 min; (2) 94 $^{\circ}$ C 变性 1 min, 退火温度 52 $^{\circ}$ C, 退火 1 min; (3) 72 $^{\circ}$ C 延伸 1 min, 循环 35 次; (4) 保持 72 $^{\circ}$ C 4 min; (5) 4 $^{\circ}$ C 保存 PCR 产物。

2.3 PCR 产物检测与测序

PCR 产物在经过琼脂糖凝胶电泳、EB (溴化乙锭) 染色、在紫外灯检测后, 送苏州金唯智公司进行测序 (为双向测序, 结果进行互相验证)。

2.4 数据处理

利用 NCBI 中的 Blast 进行序列比对, 并选择与 *G. spinosa* 亲缘关系近的几个种作为外类群。用 DNASTAR 软件包中的 EditSeq 计算 ITS 序列长度与 GC 含量, 用 DNAMAN 计算遗传距离, R 软件包 vegan 做 Mantel 检验。用 Mage6.0 计算变异位点、简约信息位点并用邻接法构建系统发育树 (利用 1 000 次重复的自展分析检验各分支的置信度)。利用 RNAfold (RNA 二级结构在线预测软件, <http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) 预测 *G. spinosa* 的 ITS 序列的 RNA 二级结构, 该软件采用最小自由能原理, 参数设为默认值。

3 结果与分析

3.1 PCR 产物电泳结果

利用 1.5% 的琼脂糖电泳检测 PCR 产物, 结果见图 1。ITS 序列条带清晰, 与 DL2000 bp 的 DNA Marker 比对后长度在 500 ~ 750 bp, 符合 ITS 序列的长度范围。

3.2 不同样品 ITS 序列比较

对 8 个 *G. spinosa* 样品进行比较发现, 除 GSCA6 的 ITS 序列不完整 (575 bp) 外, 其余来自不同地区的 7 个 *G. spinosa* 样品 ITS 区 (包括 ITS1、5.8S 和 ITS2 序列) 长度范围为 611 ~ 623 bp, 长度变化 12 bp。其中各样品 ITS1 区长度除 GSNE1 为 223 bp 外其他均为 222 bp, 长度变化不大。多数样品 5.8S 区的长度为 164 bp, 仅 GSWA8 出现了 3

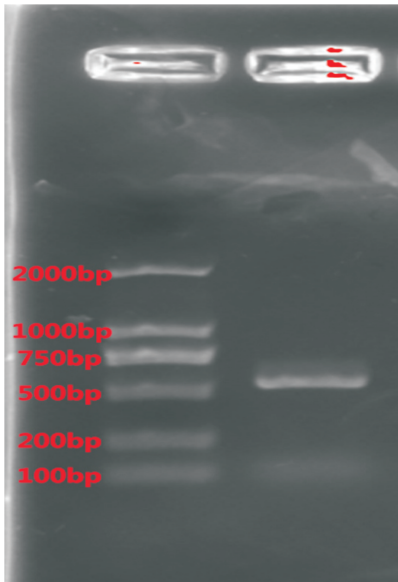


图1 *G. spinosa* ITS序列PCR结果

Fig. 1 PCR results of *G. spinosa* ITS sequence

个碱基的缺失。美国西部不同地区 *G. spinosa* ITS2 区的变化范围在 225~237 bp 之间,但 GSCA6 的 ITS2 区长度仅为 189 bp。ITS 序列的 GC 含量为 60.35%~61.0%,ITS1 与 ITS2 的 GC 量相近且较 5.8S 高(表 2)。ITS 区全序列排列后长度为 623 bp,共 22 个变异位点,5 个简约信息位点,其中 ITS1 区含 5 个变异位点,位点 79、103、113 为简约信息位点,位点 88 为碱基 G 颠换为碱基 T,且包含有一个简并碱基,位点 111 为碱基 C 转换为碱基 T。5.8S 区含 3 个变异位点,无简约信息位点,其中位点 223、248 为碱基 C 转换为 T,位点 266 为碱基 G 转换为碱基 C。ITS2 区有 14 个变异位点,其中除位点 425、616 为简约信息位点外,其余均为变异位点,其中位点 563、590、601、614、615 为碱基 G-C 之间的转换,位点 613 为碱基 T 颠换为碱基 G,448 为碱基 C 颠换为碱基 T,位点 600 出现了碱基 G-C 间的转换以及碱基 G-T 间的颠换。位点 542、622 为碱基 G-A 间的颠换,位点 617 为碱基 A 颠换为碱基 C,位点 621 为碱基 C 颠换为碱基 A。表 3 结果表明,各样品间的长度变化主要出现在 ITS2 区。GC 含量 ITS2 区较高,ITS1 区次之,5.8S 区较低。此外,简约信息位点 ITS1

区较 ITS2 区多一个,而就变异位点而言,ITS2 区变异位点较 ITS1 区多,5.8S 区除个别样品可能存在有变异位点外,其他位点均未出现变异。因此,ITS 序列的主要变异也出现在 ITS2 区。*G. spinosa* 的 ITS 序列碱基的主要变化是碱基的缺失,碱基的转换主要出现在 G-C 间,碱基的颠换主要出现在 G-T 间,而与分布区间并没有表现出相关联系。整个 ITS 区变异位点与简约信息位点分别占总位点的 3.53%和 0.80%。

3.3 样品间遗传距离分析

G. spinosa 各样品间的遗传距离在 0.0018~0.0089 之间(表 4),其中 GSNE1 与 GSWA8、GSNE2 与 GSUT4、GSUT3 与 GSUT4、GSUT4 与 GSCA5、GSUT4 与 GSCA6 间的遗传距离最小均为 0.0018。GSC07 与 GSWA8 间的遗传距离最大,为 0.0089。外类群与 *G. spinosa* 各样品间的遗传距离均较 *G. spinosa* 各样品间的遗传距离大,在 0.0107~0.0815 之间。GSNE1 与 GBC09 间的遗传距离相对较小,为 0.0107,GSWA8 与 CFCA12 间的遗传距离最大为 0.0815。对不同产地 *G. spinosa* 地理距离和遗传距离的 Mantel 检验表明,来自不同地区 *G. spinosa* 各样品间的遗传距离与地理距离的相关性不显著($R^2=0.325, P=0.359$)。

3.4 系统发育树分析

邻接法构建的系统发育树见图 2,各分支的支持率均达到 50 以上。8 个 *G. spinosa* 样品聚为一支,支持率为 93,并与同属的 GBC09 为姊妹群关系。其中,GSUT3,GSUT4,GSCA5,GSCA6,GSC07 与 GSNE2 聚为一支,GSNE1 与 GSWA8 聚为一支,但均分布于内华达地区的 GSNE1 与 GSNE2 却分属于这两支。由此可见,分布在同一地区的 *G. spinosa* 并未表现出更近的亲缘关系。

由于 GSCA6 未提交完整 ITS2 序列,且 GSC07 与 GSWA8(GSC07 和 GSWA8 的 5.8S 区出现了碱基的缺失与转换)可能存在假基因,因此我们仅以 ITS1 区序列重新构建了邻接树(图 3)。结果与基于 ITS 全序列构建的邻接树的分支结构相同,不同仅表现在支持率较低、分支长度较短。

3.5 ITS 序列 RNA 二级结构预测

利用在线软件 RNAfold 对美国西部不同分布

表 2 各样品 ITS 序列长度和 GC 含量

Table 2 Length and GC contents of ITS sequences from each sample

样品编号 Sample code	ITS		ITS1		5.8S		ITS2	
	长度 Length (bp)	GC (%)	长度 Length (bp)	GC (%)	长度 Length (bp)	GC (%)	长度 Length (bp)	GC (%)
GSNE1	623	60.51	223	62.23	164	54.27	236	63.14
GSNE2	622	60.94	222	62.16	164	54.27	236	63.98
GSUT3	622	60.94	222	62.16	164	54.27	236	63.98
GSUT4	622	61.35	222	63.96	164	54.27	235	63.83
GSCA5	622	61.00	222	63.51	164	54.27	225	63.11
GSCA6	622	60.87	222	63.51	164	54.27	189	63.49
GSC07	622	60.89	222	63.51	164	53.05	225	63.56
GSWA8	622	60.58	222	62.61	161	53.42	236	63.56
GBC09	622	61.05	222	63.96	164	54.88	225	62.67
HTU10	622	60.39	222	62.61	164	54.88	235	62.12
HTU11	622	61.96	222	63.93	164	54.88	237	64.98
CFCA12	622	56.70	222	55.86	164	54.88	226	58.85

表 3 不同地区 8 个 *G. spinosa* 样品的 ITS 序列变异位点Table 3 Variant sites of eight *G. spinosa* ITS sequences from different regions

样品编号 Sample code	ITS1					5.8S								ITS2								
	79	88	103	111	113	223	248	266	425	448	542	563	590	600	601	613	614	615	616	617	621	622
GSNE1	A	R	A	C	T	C	C	G	A	C	G	G	G	G	C	T	G	C	A	A	C	A
GSNE2	R	G	G	T	C	C	C	G	C	C	G	G	G	G	C	T	G	C	G	A	C	A
GSUT3	R	G	G	C	C	C	C	G	M	C	A	G	G	G	C	T	G	C	G	A	C	A
GSUT4	G	G	G	C	C	C	C	G	C	C	G	G	C	T	G	G	C	G	A	C	A	G
GSCA5	A	G	G	C	C	C	C	G	M	C	G	C	G	G	C	T	-	-	-	-	-	-
GSCA6	A	G	G	C	C	C	C	G	C	T	G	G	-	-	-	-	-	-	-	-	-	-
GSC07	A	G	G	C	C	T	T	G	C	C	G	G	G	C	C	-	-	-	-	-	-	-
GSWA8	G	T	A	C	T	C	C	A	A	C	G	G	G	C	C	T	G	G	G	A	C	A

注:“-”表示碱基缺失。

Note:“-” indicates the missing base.

区 *G. spinosa* ITS 的 RNA 二级结构进行了预测(图 4)。ITS1 区与 ITS2 区均含有封闭的环状结构,8 个 *G. spinosa* ITS 序列的 RNA 二级结构依它们的相似状况大体上可以分为四类,分别记为 type A, B, C 和 D。GSNE1、GSNE2 与 GSUT3 三者的 RNA 二级结构的构象相似记为 type A。

RNA 二级结构构象相似的 GSUT4 和 GSCA6 记为 type B,他们在 5.8S 区与 ITS2 区均出现了差异,且 GSCA6 的最小自由能更低,推测与 GSCA6 ITS 序列的缺失有关。GSCA5 与 GSWA8(type C)的构象相似,主要差异出现在 ITS1 区,由于 GSWA8 的最小自由能较 GSCA5 的高,因此推测

表 4 来自不同地区各样品间的遗传距离

Table 4 Genetic distance among samples from different regions

	GSNE1	GSNE2	GSUT3	GSUT4	GSCA5	GSCA6	GSCO7	GSWA8	GBCO9	HTU10	HTU11	CFCA12
GSNE1												
GSNE2	0.005 3											
GSUT3	0.005 3	0.003 6										
GSUT4	0.003 6	0.001 8	0.001 8									
GSCA5	0.005 3	0.003 6	0.003 6	0.001 8								
GSCA6	0.005 3	0.003 6	0.003 6	0.001 8	0.003 6							
GSCO7	0.007 1	0.005 3	0.005 3	0.003 6	0.005 3	0.005 3						
GSWA8	0.001 8	0.007 1	0.007 1	0.005 3	0.007 1	0.007 1	0.008 9					
GBCO9	0.010 7	0.012 6	0.012 6	0.010 8	0.012 6	0.012 6	0.014 4	0.012 6				
HTU 10	0.036 6	0.038 5	0.038 5	0.036 6	0.038 5	0.038 5	0.040 3	0.038 5	0.032 8			
HTU 11	0.038 5	0.036 6	0.040 4	0.038 5	0.040 4	0.040 4	0.042 2	0.040 4	0.032 8	0.057 7		
CFCA12	0.079 4	0.073 5	0.077 5	0.075 5	0.077 6	0.077 4	0.079 4	0.081 5	0.075 7	0.083 8	0.086 4	

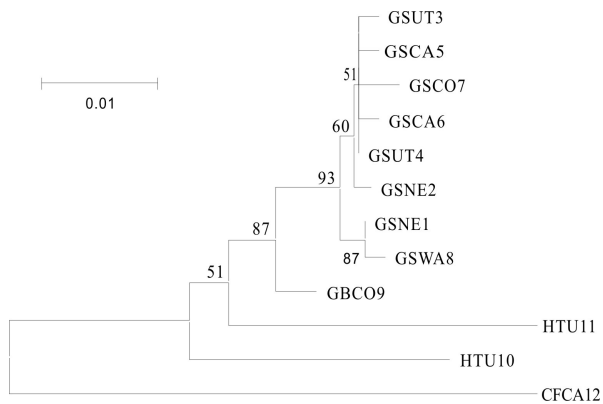


图 2 基于 ITS 序列分析构建的系统树
Fig. 2 Phylogenetic tree based on ITS sequence

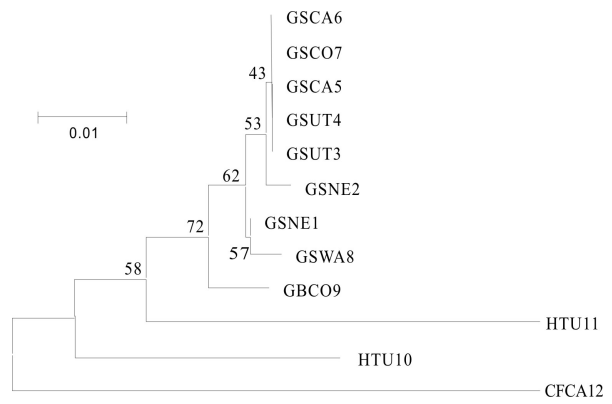
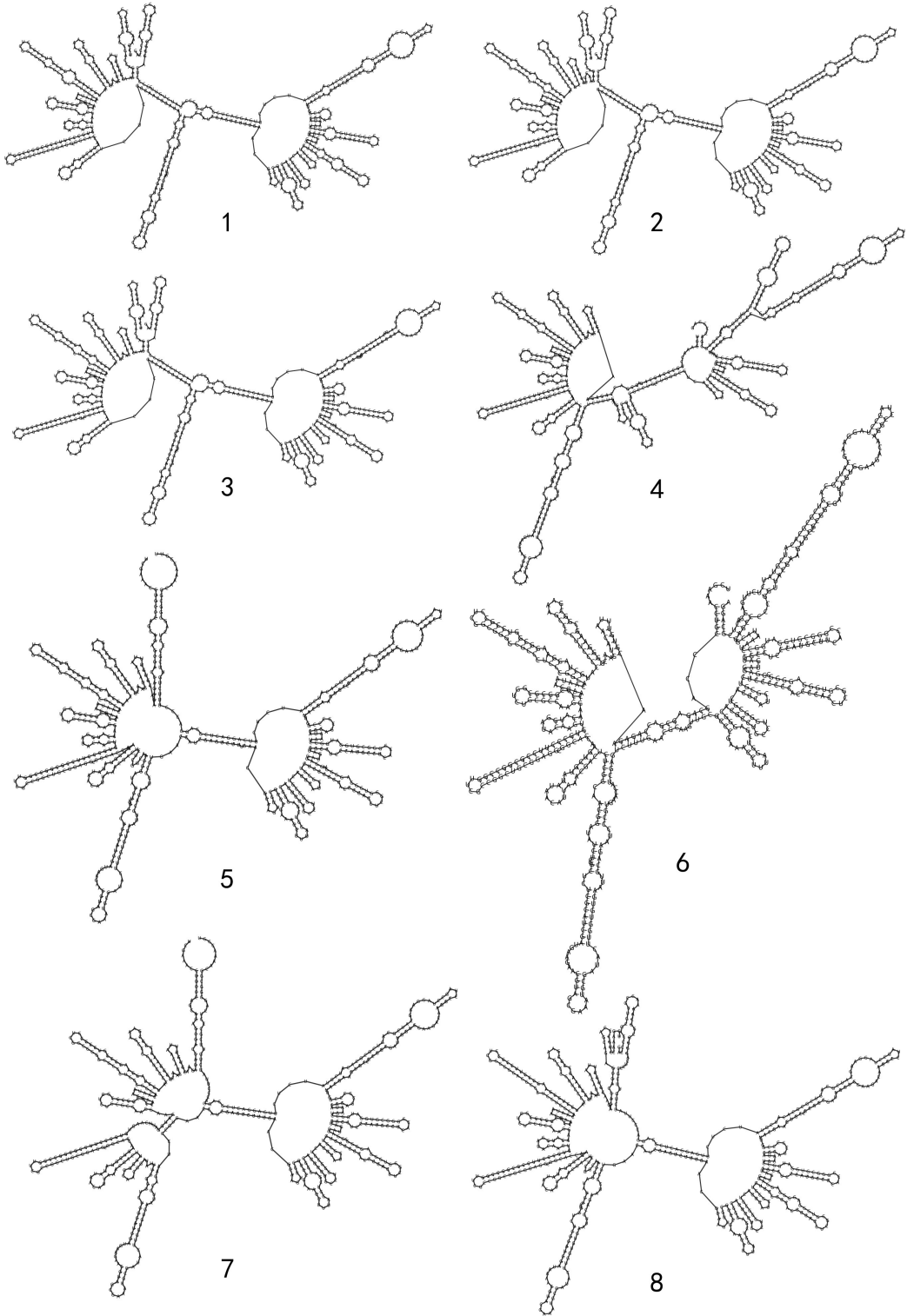


图 3 基于 ITS1 序列分析构建的系统树
Fig. 3 Phylogenetic tree based on ITS1 sequence

GSCA5 较 GSWA8 的 RNA 二级结构稳定。GSCO7 (type D) 在 ITS1 区出现了两个封闭的环, 其 RNA 二级结构均不同于以上三类构象, 在 ITS1 区出现了两个封闭的环。总体而言, *G. spinosa* ITS 序列 RNA 二级结构在 ITS1, ITS2 区以及 5.8S 区均出现了差异。虽然 8 个 *G. spinosa* ITS 样品彼此间 RNA 二级结构的相似性状况与一级结构的并不相同, 但同样与其地理分布区间没有体现出一定的相关性。

4 讨论

ITS 序列是近年来用于探讨植物近缘属间、种间系统关系和种内变异的重要分子标记。被子植物的 ITS 区长度比较稳定, 包括 5.8S rDNA 在内, 总长度为 600 ~ 700 bp。该研究序列分析结果显示, 8 个 *G. spinosa* ITS 序列的长度仅 GSNE1 为 623



注: 1, 2, 3. 类型 A, 编号分别为 GSNE1, GSNE2, GSUT3, 最小自由能分别为 -273.10 、 -272.20 和 -271.20 kcal \cdot mol $^{-1}$; 4, 6. 类型 B, 编号分别为 GSUT4 和 GSCA6, 最小自由能分别为 -277.70 和 -253.80 kcal \cdot mol $^{-1}$; 5, 8. 类型 C, 编号分别为 GSCA5 和 GSWA8, 最小自由能分别为 -271.80 和 -279.40 kcal \cdot mol $^{-1}$; 7. 类型 D, 编号为 GSC07, 最小自由能为 -271.90 kcal \cdot mol $^{-1}$ 。

Note: 1, 2, 3. Type A, codes are GSNE1, GSNE2, GSUT3, the minimum free energy are -273.10 , -272.20 and -271.20 kcal \cdot mol $^{-1}$ respectively; 4, 6. Type B, codes are GSUT4 and GSCA6, the minimum free energy are -277.70 and -253.80 kcal \cdot mol $^{-1}$ respectively; 5, 8. Type C, codes are GSCA5 and GSWA8, the minimum free energy are -271.80 and -279.40 kcal \cdot mol $^{-1}$ respectively; 7. Type D, code is GSC07, the minimum free energy is -271.90 kcal \cdot mol $^{-1}$.

图 4 *G. spinosa* ITS 序列 RNA 二级结构图

Fig. 4 RNA secondary structure of *G. spinosa* ITS sequence

bp,其余均为 622 bp,符合被子植物 ITS 序列长度的变化范围。*G. spinosa* ITS1 区序列长度保守,除 GSNE1 为 223 bp 外,其余均为 222 bp,ITS 序列长度的主要变化出现在 ITS2 区,多数样品 5.8S 区长度为 164 bp,但是 GSWA8 为 161 bp。Baldwin (1992)最早对被子植物 ITS 序列假基因的研究表明:被子植物 5.8 S 序列一般长度都在 163~164 bp 之间,若出现了插入或缺失就有可能形成假基因(人们把含有不具编码功能 5.8S 区 ITS 序列定义为 ITS 假基因)。*G. spinosa* ITS 序列包括 22 个变异位点,说明 ITS 序列在 *G. spinosa* 系统发育关系分析中能够提供足够的证据。5.8S 区最为保守,无简约信息位点,变异位点少。从 ITS 序列的变异位点来看,*G. spinosa* 共有 22 个变异位点,这些变异位点主要存在于 ITS2 区,包含 14 个变异位点,占总变异位点数的 63.63%,ITS1 区仅有 5 个变异位点,5.8S 区最为保守,含 3 个变异位点,无简约信息位点,因此仅以 ITS1 区构建系统发育树并不能提供足够信息量。屈良鹄和陈月琴(1999)通过对不同生物类群的 ITS 序列的比较得出:被子植物大多数科属 ITS 序列的种间差异值为 1.2%~10.2%,曾明等(2003)和车建等(2007)在利用 ITS 序列分别对葛根和西红柿正伪品的鉴定中均认为 ITS 序列种内差异值小于 1%。该研究测得 5 个地区的 8 个 *G. spinosa* 样品在 ITS 区的碱基差异为 3.53%,表现出较大的遗传分化。同样,在对同一地区的 *G. spinosa* 样品系统发育分析中得出分布在内华达的 GSNE1 与 GSNE2、犹他州的 GSUT3 与 GSUT4 以及分布在加利福尼亚的 GSCA5 与 GSCA6,相互间的遗传距离除 GSUT3 与 GSUT4 外,其余并非最小,采用邻接法构建的系统发育树,显示出同一地区的 GSNE1 与 GSNE2 也未聚在一起。这与地理距离与遗传距离 Mantel 检验的结果相吻合。

由于 RNA 二级结构比一级结构更具保守性,因此可以通过二级结构反映一级结构所反映不出的生物信息。该研究对 8 个 *G. spinosa* 样品的 ITS 序列 RNA 二级结构进行了比较,发现主要差异出现在 ITS1 区和 ITS2 区,这与 ITS 序列一级结构分析中其主要的变异位于 ITS1 区和 ITS2 区的结论

相一致。但是,有研究表明多数真核生物中 ITS1 区的 RNA 二级结构具有较高保守性 (Musters et al, 1990),并且 3'端无论是在 ITS 序列一级结构还是其 RNA 二级结构中都比 5'端更保守。而在我们的研究中对 8 个 *G. spinosa* 样品的 ITS 序列 RNA 二级结构的比较中发现仅 GSUT4 与 GSCA6 的 ITS2 区的 RNA 二级结构与其他 6 个样品较大的不同,较 ITS1 区 RNA 二级结构更为保守,这与 ITS2 区在进化中所受的选择压力较高有关。该研究对 ITS 序列一级结构与其 RNA 二级结构的分析还发现,在 ITS 序列一级结构系统发育分析中 GSNE1 与 GSWA8 体现出更近的亲缘关系,但是二者的 RNA 二级结构差异明显,分属于 2 个不同的类型中,同时 GSNE2, GSUT3, GSUT4, GSCA5, GSCA6, GSCO7 在 ITS 序列一级结构分析中也体现了较近的亲缘关系,但是在它们的 RNA 二级结构分析中发现,GSNE2 却在一级结构分析中处在另一分支上的 GSNE1 的 RNA 二级结构基本一致,相似的情况也出现在 GSWA8 与 GSCA5 之间。研究发现突变位点若位于茎区,则有可能使二级结构改变,若突变发生在环上或自由基则对二级结构的影响不大,这可能与环状结构与维持 RNA 的功能有关。比较该研究中 8 个 *G. spinosa* 样品 ITS 一级结构的突变位点仅 ITS1 区的位点 88、103、111,ITS2 区的位点 621、622 位于茎区,其余均位于环上,则对其二级结构的影响不大。总之,*G. spinosa* 样品 ITS 一级结构与其 RNA 二级结构的变化并未表现出完美的一一对应关系, RNA 二级结构体现出了更高的保守性。

致谢 衷心地感谢朱格麟老师为该研究提供实验材料。

参考文献:

- BARAKET G, SADDODU O, CHATTI K, 2009. Sequence analysis of the internal transcribed spacers (ITSs) region of the nuclear ribosomal DNA (nrDNA) in fig cultivars (*Ficus carica* L.) [J]. *Sci Horti*, 120(1):34-40.
- BALDWIN BG, 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the compositae [J]. *Mol Phylogenet Evol*, 1(1): 3.

- CHEN LH, YU Z, JIN HP, 2010. Comparison of ribosomal DNA ITS regions among *Hippophae rhamnoides* from different geographic areas in China [J]. *Plant Mol Biol Rep*, 28(4): 635-645.
- CHE J, TANG L, LIU YJ, et al, 2007. Identification of *Crocus sativus* L. and its mixable Chinese medicinal materials by ITS sequence [J]. *Chin J Chin Mat Med*, 32(8):668-671. [车建, 唐琳, 刘彦君, 等, 2007. ITS 序列鉴定西红花与其易混中药材 [J]. *中国中药杂志*, 32(8):668-671.]
- DOYLE JA, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue [J]. *Phytochem Bull*, 19(1): 11-15.
- GALSTONA W, 1973. *Flora of North America* [J]. *America: Bot Gaz*; 3(7):64-64.
- GIUDICELLIGC, MADER G, SILVAARIAS GA, et al, 2017. Secondary structure of nrDNA Internal transcribed spacers as a useful tool to align highly divergent species in phylogenetic studies [J]. *Genet Mol Biol*, 40(1 Suppl. 1): 191-199.
- GUO GY, CHEN F, SHI XD, et al, 2016. Genetic variation and phylogenetic relationship analysis of *Jatropha curcas* L. inferred from nrDNA ITS sequences [J]. *CR Biol*, 339: 9-10.
- HE YH, 2012. Bioinformatic analysis of the Elaeagnaceae nrDNA ITS sequences [D]. Lanzhou: Northwest Normal University: 21-65. [何懿茜, 2012. 胡颓子科 nrDNA ITS 序列的生物信息学分析 [D]. 兰州:西北师范大学:21-65.]
- JADHAV MS, SHAMA TR, 2014. Molecular authentication of *Hippophae* species and subspecies and their phylogenetic analysis based on nuclear ITS sequences from north western Himalayan region of India [J]. *Res J Biotechnol*, 9(10): 62-68.
- JIANG LY, GUO ZG, WANG C, et al, 2009. ITS Sequence analysis of *Gynostemma pentaphyllum* from different habitats in China [J]. *Chin Trad Herb Drugs*, 7(7):1123-1127. [蒋玲艳, 郭志刚, 王翀, 等, 2009. 中国不同地区绞股蓝 ITS 序列分析 [J]. *中草药*, 7(7):1123-1127.]
- KANG Y, ZHANG M L, CHEN Z D, 2003. A preliminary phylogenetic study of the subgenus *Pogonophace* (*Astragalus*) in China based on ITS sequence data [J]. *Acta Bot Sin*, (45): 140-145.
- KHAN G, ZHANG F, GAO Q, et al, 2014. Molecular phylogeography and intraspecific divergence of *Spiraea alpina* (Rosaceae) distributed in the Qinghai-Tibetan Plateau and adjacent regions inferred from nrDNA [J]. *Biochem Syst Ecol*, 57:278-286.
- LI N, JIANG YF, SU X, et al, 2016. Genetic diversity and genetic structure of northern margin populations of *Hippophae neurocarpa* [J]. *Guihaia*, 36(5):557-563. [李霓, 蒋严妃, 苏雪, 等, 2016. 肋果沙棘北缘居群的遗传多样性与遗传结构 [J]. *广西植物*, 36(5):557-563.]
- LIANG XH, WU YX, 2017. Identification of *Kalidium* species (Chenopodiaceae) by DNA barcoding [J]. *Sci Cold Arid Reg*, 9(1):89-96.
- MUSTERS W, BOON K, SANDE CAVD, et al, 1990. Functional analysis of transcribed spacers of yeast ribosomal DNA [J]. *Embo J*, 9(12):3989.
- QU LH, CHEN YQ, 1999. Principle and method in the classification key of Biological molecular [J]. *J Sun Yat-sen Univ (Nat Sci Ed)*, (1):1-6. [屈良鹤, 陈月琴, 1999. 生物分子分类检索表—原理与方法 [J]. *中山大学学报(自然科学版)*, (1):1-6.]
- TAKAMIVA T, WONGSWAD P, TAJIMA N, et al, 2011. Identification of dendrobium species used for herbal medicines based on ribosomal DNA internal transcribed spacer sequence [J]. *Biol Pharm Bull*, 34(5):779.
- WANG J, YANG X, YANG ZL, et al, 2013. nrDNA ITS sequences analysis and genetic relationship identification of *Magnolia officinalis* from different geographical region [J]. *Guihaia*, 33(1):35-41. [王洁, 杨旭, 杨志玲, 等, 2013. 不同产区厚朴 nrDNA ITS 序列分析及亲缘关系鉴定 [J]. *广西植物*, 33(1):35-41.]
- YANG P, SHU JF, CAI SS, et al, 2017. nrDNA ITS sequence analysis and genetic relationship of *Allium macrostemon* from different geographical regions in Guizhou [J]. *Plant Sci J*, 35(2):171-176. [杨鹏, 舒建锋, 蔡莎莎, 等, 2017. 贵州不同产地薤白 nrDNA ITS 序列分析及亲缘关系研究 [J]. *植物科学学报*, 35(2):171-176.]
- YANO O, HOSHINO T, 2007. Phylogenetic relationships and chromosomal evolution of Japanese *Fimbristylis* (Cyperaceae) using nrDNA ITS and ETS sequence data [J]. *Appl Acta Phytotax Geobot*, 57(3):205-217.
- ZACHARIAS EH, BALDWIN BG, 2010. A molecular phylogeny of North American *Atripliceae* (Chenopodiaceae), with implications for floral and photosynthetic pathway evolution [J]. *Syst Bot*, 35(4):32-34.
- ZENG M, MA YJ, ZHEN SQ, et al, 2003. Studies on ribosomal DNA sequence analyses of *Radix puerariae* and its sibling species [J]. *Chin Pharm J*, 38(3):173-175. [曾明, 马雅军, 郑水庆, 等, 2003. 中药葛根及其近缘种的 rDNA-ITS 序列分析 [J]. *中国药学杂志*, 38(3):173-175.]
- ZHANG LL, WANG J, WAN XB, et al, 2017. Analysis of internal transcribed spacers (its) sequences and phylogenetics of main bast fiber crops [J]. *Acta Agron Sin*, 43(6):862-874. [张力岚, 王俊, 万雪贝, 等, 2017. 主要麻类作物的 ITS 序列分析与系统进化 [J]. *作物学报*, 43(6):862-874.]
- ZHU GL, MOSYAKIN SL, CLEMANTS SE, 2003. *Flora of China Chenopodiaceae* [M]. Beijing: Science Press, 5: 351-414.
- ZHU GL, SANDERSON SC, 2017. Genera and a new evolutionary system of world Chenopodiaceae [M]. Beijing: Science Press: 56-96.
- ZHOU Y, DU XL, ZHENG X, et al, 2017. ITS2 barcode for identifying the officinal rhubarb source plants from its adulterants [J]. *Biochem Syst Ecol*, 70:177-185.
- ZHUO L, 2010. A study of phytogeographic distribution and molecular systematics of the subfamily atriplicioideae in China [D]. Urumqi: Xinjiang Agricultural University:19-36. [卓立, 2010. 中国滨藜亚科的地理分布与分子系统学研究 [D]. 乌鲁木齐:新疆农业大学:19-36.]