J Mol Evol (2001) 53:124–134
DOI: 10.1007/s002390010200

JOURNAL OF MOLECULAR EVOLUTION

Extensive Ribosomal DNA Genic Variation in the Columnar Cactus *Lophocereus*

Stefanie Hartmann,[1] John D. Nason,[1,2] Debashish Bhattacharya[1]

[1] University of Iowa, Department of Biological Sciences, 239 Biology Building, Iowa City, IA, 52242-1324, USA
[2] Iowa State University, Department of Botany, Ames, IA, 50011, USA

Received: 28 July 2000 / Accepted: 7 March 2001

Abstract. Sequence analysis of the hypervariable internal transcribed spacer (ITS) regions of ribosomal DNA (rDNA) is commonly used to gain insights into plant and animal population structure and phylogeny. We characterized ITS1, ITS2, and the 5.8S coding region of 18 senita (*Lophocereus*) individuals from 12 different populations in Baja as well as from closely related cactus species. Analyses of multiple clones demonstrated extensive paralogy in the senita rDNA gene family. We identified at least two putatively non-recombining rDNA operons in senita as well as multiple paralogous sequences within each operon. Usage of PCR, reverse transcriptase (RT)-PCR, Southern blot, primary sequence analyses of the 18S rDNA gene, and secondary structure analyses of the 5.8S rRNA showed that one of the operons encodes rDNA pseudogenes in a low copy-number (*Truncated*), whereas the second operon encodes an expressed rRNA (*Functional*). Surprisingly, we found extensive paralogy not only in the ITS regions but also in the 5.8S coding regions in senita both within and between operons. Phylogenetic analyses suggest that the second rDNA operon originated prior to the divergence of *Lophocereus.* A significant ($p < 0.05$) divergence-rate acceleration was found in the *Lophocereus* 5.8S rDNA coding region in the *Functional* operon in comparison to *Pereskiopsis porteri* (Cactaceae) and *Portulaca molokiniensis* (Portulacaceae) with *Silene dioica* and *Spinacia oleracea* as the outgroups.

Correspondence to: Debashish Bhattacharya; *email*: dbhattac@blue.weeg.uiowa.edu

Key words: ITS — *Lophocereus* — Pachycereeae — *Pachycereus* — Paralogy — Phylogeny — Pseudogene — Ribosomal RNA — Senita

Introduction

Ribosomal DNA (rDNA) coding regions are present in all cells and have been widely used for evolutionary studies (Baldwin et al. 1995; Soltis and Soltis 1998; Hershkovitz et al. 1999). The canonical rDNA operon consists of the 18S coding region, the internal transcribed spacer 1 (ITS1) region, the 5.8S coding region, the ITS2 region, the 28S coding region, and the non-transcribed spacer (NTS). The operons are tandemly repeated and, because of their high copy number (up to thousands of copies per cell), are easily amplified using PCR methods and conserved primers. The different regions of the rDNA operon repeat unit have different rates of sequence evolution that have been used to study variation at all systematic levels (Hamby and Zimmer 1992; Hershkovitz and Lewis 1996; Mayer and Soltis 1999). Whereas, the ITS and NTS regions have been widely applied at the interspecific-intergenetic level, few studies document population-level patterns of variation that are presumed to give rise to higher-level patterns of divergence.

Concerted evolution, driven by unequal crossing over and gene conversion, is an important factor in the evolution of gene families such as that encoding rDNA (Dover 1982). In extreme cases, all repeat members may be homogenized by this process and paralogy is not a concern in phylogenetic analyses, i.e., any rDNA copy can

be used as a representative of the family. A gene family that is not subject to concerted evolution also presents no problems of paralogy as long as the different orthologs are correctly identified in taxa. Theory predicts, however, that there may be cases of intermediate concerted evolution and partial homogenization, for example, through introgression or migration, which introduces new variants faster than they can be homogenized within the genome (Sanderson and Doyle 1992). This is a particularly vexing problem because, gone unnoticed, rDNA paralogy can lead to substantial errors in the reconstruction of species/population phylogenies (Buckler et al. 1997; Rich et al. 1997). It is, therefore, of utmost importance that the presence and extent of rDNA paralogy be determined, particularly in studies conducted at lower taxonomic levels. Growing evidence points out, in fact, that intra-individual rDNA sequence variation, be it microheterogeneity (i.e., paralog polymorphism) or the existence of distinct operons, is a common problem with this molecular marker (e.g., Suh et al. 1993; Buckler and Holtsford 1996b; Miller et al. 1996; Campbell et al. 1997; Hershkovitz and Zimmer 1997; Hershkovitz et al. 1999). Distinguishing orthologs from paralogs is of greatest concern in intraspecific studies in which ancestral polymorphisms, in combination with incomplete concerted evolution, may render rDNA ITS data unusable for uncovering processes such as gene flow. In these cases, incomplete homogenization of rDNA sequences (or any gene family) is better used to understand introgression or recent migration events.

In this study, we conduct a detailed examination of nuclear rDNA structure in the Sonoran Desert columnar cactus, senita (genus *Lophocereus*, Cactaceae). We show the existence of two distinct rDNA operons that share limited sequence identity and have likely remained distinct entities since the origin of *Lophocereus.* One operon (*Functional*) encodes an expressed 18S rRNA and a 5.8S coding region that folds into a conserved secondary structure (Hershkovitz and Lewis 1996; Yeh and Lee 1991). Paralagous copies of this operon were found with sequence divergence, not only in the ITS regions of the rDNA, but, remarkably, also within the 5.8S rRNA. The second operon (*Truncated*) encodes a truncated, nonfunctional 18S sequence, has a 157 bp deletion in ITS2, and a highly divergent 5.8S rRNA. *Truncated* also has paralagous copies with sequence variation both in the 5.8S gene and the ITS regions. Two rDNA operons also exist in *Pachycereus marginatus,* the sister taxa of *Lophocereus.* From these data, we conclude that the second operon predates the origin of *Lophocereus* and *Pachycereus marginatus.*

Materials and Methods

Study Species and Sample Populations. Senita is a large, multistemmed and long-lived Sonoran Desert species in the tribe Pachycereae, that ranges from northwestern Sinaloa, western Sonora, Baja California, and adjacent islands in Mexico into southern Arizona in Organ Pipe Cactus National Monument (Lindsay 1963). The genus *Lophocereus* presently consists of only two described species, *L. schottii* and *L. gatesii* (Lindsay 1963). The predominant *L. schottii schottii* is restricted to, but widely distributed within the Sonoran Desert of northwestern Mexico and the southwestern United States. The two subspecies *L. s. tenuis* and *L. s. australis* occur in the coastal regions of southern Sonora/northern Sinaloa and the southernmost region of Baja, respectively (Lindsay 1963). Eighteen senita populations were sampled from Baja California as shown in Fig. 1. Seeds from a senita individual in population 18 have been deposited in the University of Iowa Herbarium. Seeds from several of the other populations used in this study are available upon request.

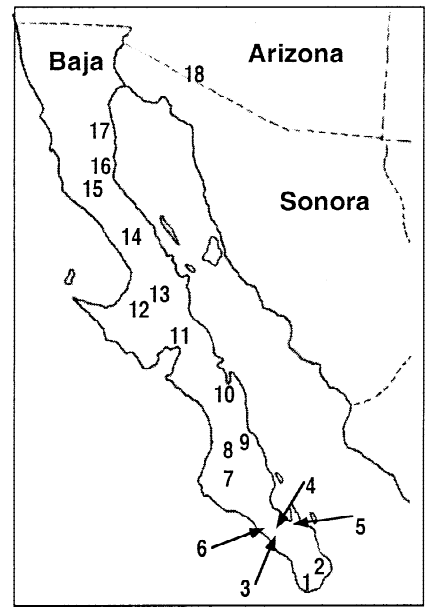

Fig. 1. Location of the 18 populations of *Lophocereus* that were sampled in Baja, California.

DNA Extraction and PCR Methods. Senita total genomic DNA was extracted with the CTAB method (Doyle and Doyle 1987). Modifications were implemented to optimize DNA extraction from the polysaccharide-rich cactus tissues. In our method, approximately 50 mg of frozen tissue (flowerbuds or stem tissue) was ground to a fine powder in liquid nitrogen for some samples. For other samples, anthers were scraped out of unopened flowerbuds. The frozen powder was added to 500 μl CTAB buffer (2.5% CTAB, 3.5 M NaCl, 50 mM ETDA, 250 mM Tris Cl pH 8.0) and a small amount of glass beads (0.2 mm diameter), and vortexed vigorously. After incubation at 60 °C for 30–60 min, 500 μl Phenol:Chloroform:Isoamylalcohol (25:24:1) was added and the mixture was vortexed. The sample was centrifuged at 14000 rpm for 5 min at 4 °C. The aqueous phase was transferred to a fresh tube, and the extraction was repeated with an equal volume of Phenol:Chloroform:Isoamylalcohol. Ten μg/ml RNAse (A) was added to the aqueous phase; incubation was carried out at 37 °C for 20–60 min. Where necessary, a second Phenol:Chloroform:Isoamylalcohol extraction was carried out and/or CTAB was removed by an overnight dialysis step against TE buffer at 4 °C (Pierce microdialyzer; dialysis tubing: Gibco BRL, Molecular weight exclusion: 12000–14000 Da). To remove polysaccharides, further purification of the DNA with glassmilk (GeneClean, Bio101) was carried out for several of the samples. The DNA was then resuspended in 20–50 μl TE buffer and a small amount was analyzed on a 0.8% agarose gel to assess its quality and yield.

Different sets of PCR primers were used to amplify the fragment containing the ITS1 region, the 5.8S rDNA coding region, and the ITS2 region. Forward primers were either 1800F at the 3′ terminus of the 18S sequence (5′-ACCTGCGGAAGGATCATTG-3′) or LsITS1-F (5′-CAAGCAAAAAGACCCGTGAAC-3′) which was complementary to the 5′-terminus of ITS1. The reverse PCR primers were either ITS4-R (5′-TCCTCCGCTTATTGATATGC-3′) at the 5′ terminus of the 28S coding region and LsITS2-R (5′-CTTTYAACACGCGTTAGGGTCC-3′) near the 3′ terminus of ITS2. The template DNA was first denatured at 94 °C for 10 min and then subjected to 35 cycles of the following PCR program: 94 °C for 1 min, 53 °C (for 1800F, ITS4) or 60 °C (for LsITS1-F,-ITS2-R) for 1 min, 72 °C for 1 min. The final cycle included an extension of 10 min at 72 °C. These PCR reactions always resulted in the amplification of two bands: one higher-yield, larger band of ca. 700 bp (upper) and another, smaller band of ca. 550 bp (lower). For this reason, the entire reaction volume was loaded onto a 1% agarose gel and the 700 bp PCR product was excised and purified with either Wizard PCR Preps (Promega DNA purification system) or with glassmilk according to the manufacturer's instructions. The purified DNA product was quantified on a 1% agarose gel and sequenced using a dye termination sequencing protocol. The reactions were done on a 373 A Fluorescent Automated Sequencer (Perkin Elmer-Applied Biosystems).

Sequence analysis of the upper band showed that it contained the major, multi-copy rDNA operon which, however, encoded at least two distinct paralogs. We decided then to further analyze the lower band, which was also determined to be rDNA from senita but with a 157 bp deletion in the ITS2 fragment. Motivation for this thorough analysis of the lower band rDNA operon was twofold: first, we wanted to determine whether this operon also showed intra-individual heterogeneity and thus explore the possibility of using it as a phylogeographic marker; second, we wanted to determine if this second operon encoded a functional rDNA sequence.

Heterogeneity Within the Lower Band Sequence. To study intra-individual heterogeneity of the ITS fragments, the lower band from one individual (population 2) was PCR-amplified and subcloned into the pCR2.1 vector (Invitrogen). For the ligation, 12.5 ng vector and the purified PCR product were incubated in a 10 μl volume at 4 °C overnight with T4 DNA ligase (Promega). The entire ligation reaction was used for transformation of competent JM109 cells according to standard protocols. Single colonies were picked, grown in liquid medium over night at 37 °C, and plasmid DNA was isolated using a standard alkaline lysis mini-prep protocol (Sambrook et al. 1989). Plasmids were digested with *Eco*RI to confirm the presence of an insert.

Individual clones were used as a template for PCR amplification of the lower band using the 1800F and ITS4-R primers and the conditions described above. PCR products were double-digested with *Alu*I and *Hinf*I, ethanol-precipitated, collected by centrifugation for 20 min at 4 °C, and resuspended in 10 μl water. Loading buffer was added and the samples were heated to 95 °C for 10 min, quick-cooled, and loaded on an 8% polyacrylamide gel.

Amplification and Sequencing of ITS from Related Cactus Species. Other members of the tribe Pachycereeae were also studied to understand more broadly rDNA evolution in these cacti. Sampled taxa were *Pachycereus marginatus, P. weberi* (both purchased from Mesa Garden, New Mexico), *P. pringlei,* and *P. pecten-aboriginum* (both collected from wild populations in the Sonoran Desert). Seeds of an individual from *P. pecten-aboriginum* were placed in the University of Iowa Herbarium. Total genomic DNA was prepared from these taxa as described above. PCR reactions were done with the template DNAs and the ITS4 and 1800F primers to amplify fragments containing the ITS1, ITS2, and the 5.8S rDNA coding region. PCR products were cloned into the pGEM-T Vector (Promega) according to the manufacturer's protocol and sequenced using standard sequencing primers.

Phylogenetic Methods. We used the 5.8S rDNA sequences to reconstruct the phylogeny of the senita populations and related cactus species (Table 1). This alignment, of length 164 nt, was analyzed with the maximum likelihood:(fastDNAml I.1, Olsen et al. 1994) and maximum parsimony (PAUP* V4.0b3a, Swofford 2000) methods. The fast DNAml tree was constructed using the global search option, transition/transversion ration = 2, and a jumbled sequence input. To assess the robustness of each clade, bootstrap analysis (Felsenstein 1985) was done with 100 resamplings. Maximum parsimony analysis was done using the heuristic search of ten replicates with random stepwise addition and tree bisection-reconnection (TBR) and the MulTrees option. One-thousand bootstrap replicates were analyzed with this method. The 5.8S rDNA phylogeny was outgroup-rooted with this sequence from *Silene dioica* and *Spinacia oleracea* which are outside of the clade containing the Cactaceae genera *Pachycereus* and *Lophocereus* in the order Caryophyllales. A second data set of 5.8S and ITS1 regions that included 38 sequences from individuals in 17 Baja populations (see Table 2) was analyzed with the distance method. Corrected distances were calculated with the HKY85 model (Hasegawa et al. 1985) and a phylogeny was built with neighbor-joining. The transition/transversion ratio was set to 2, gaps were excluded in pairwise comparisons, and an uniform mutation rate was used in the calculation of the distance matrix. These data were subjected to a bootstrap distance analysis (1000 replications). All phylogenetic analyses were done with PAUP* (V4.0b3a, Swofford 2000).

Secondary Structure Analysis. A putative secondary structure of the 5.8S rRNA from a sequence isolated from the upper band of an individual from population 17 was reconstructed using existing structures as a guide (e.g., Yeh and Lee 1991; Hershkovitz and Lewis 1996).

Table 1. Caryophyllales 5.8S sequences used for the phylogenetic analysis of the *Functional* and *Truncated* operons. L is the lower PCR fragment in *Lophocereus* spp. and *Pachycereus marginatus*

Species: population		Family	Accession number
Lophocereus:	17-3	Cactaceae	AF328639
	17-2	Cactaceae	AF328640
	17-2L	Cactaceae	AF328629
	17-1	Cactaceae	AF328641
	17-1L	Cactaceae	AF328628
	16-1	Cactaceae	AF328642
	15-2	Cactaceae	AF328644
	15-1	Cactaceae	AF328645
	12-2L	Cactaceae	AF328630
	11-1	Cactaceae	AF328650
	10-2	Cactaceae	AF328652
	9-1	Cactaceae	AF328654
	8-2	Cactaceae	AF328655
	8-1L	Cactaceae	AF328631
	6-2	Cactaceae	AF328658
	6-1L	Cactaceae	AF328632
	5-1L	Cactaceae	AF328633
	2-1L (Type 1)	Cactaceae	AF328636
	2-1L (Type 2)	Cactaceae	AF328635
	2-1L (Type 3)	Cactaceae	AF328634
Pachycereus marginatus L		Cactaceae	AF328667
Pachycereus marginatus		Cactaceae	AF328665
Pachycereus pecten-aboriginum		Cactaceae	AF328669
Pachycereus. weberi		Cactaceae	AF328666
Pachycereus. pringlei		Cactaceae	AF328668
Pereskiopsis porteri		Cactaceae	GSDB.S.109828
Portulaca molokiniensis		Portulacaceae	GSDB.S.76069
Portulaca quadrifida		Portulacaceae	GSDB.S.76070
Portulaca oleracea		Portulacaceae	GSDB.S.109833
Silene diocia		Caryophyllaceae	X86830
Spinacia oleracea		Chenopodiaceae	AF062088

Table 2. *Lophocereus* spp. *Functional* and *Truncated* (L) 5.8S and ITS1 sequences used for the phylogenetic analysis. *Pachycereus* spp. was the outgroup in this tree

Species: population	Accession number	Species: population	Accession number
Lophocereus:		*Lophocereus:*	
17-4	AF328638	9-1	AF328654
17-3	AF328639	8-2	AF328655
17-2	AF328640	8-1L	AF328631
17-2L	AF328629	7-2	AF328656
17-1	AF328641	7-1	AF328657
17-1L	AF328628	6-2	AF328658
16-1	AF328642	6-1L	AF328632
15-3	AF328643	5-1	AF328659
15-2	AF328644	5-1L	AF328633
15-1	AF328645	4-2	AF328660
14-1	AF328646	4-1	AF328661
13-1	AF328647	3-1	AF328662
12-2	AF328648	2-1	AF328663
12-2L	AF328630	2-1L (Type 1)	AF328636
12-1	AF328649	2-1L (Type 2)	AF328635
11-1	AF328650	2-1L (Type 3)	AF328634
10-3	AF328651	1-1	AF328664
10-2	AF328652	1-1L	AF328637
10-1	AF328653		
P. weberi	AF328666	*P. marginatus*	AF328665

Southern Blot Analyses. We did Southern blot analyses to confirm the presence of/and to estimate the relative copy numbers of both upper and lower rDNA fragments in the senita genome. Genomic DNA was digested to completion with *Bam*HI, *Eco*RI, *Hind*III, electrophoresed through an 1% agarose gel, and blotted onto a nylon membrane via capillary transfer. Purified PCR product (50–100 ng) was labeled in the presence of 50 μCi of ^{32}P-CTP using the Random Primer labeling Kit (Boehringer Mannheim) according to the manufacturer's instructions. Prehybridization (90 min) and hybridization (overnight) were done at 62 °C in a solution containing 0.25M Na_2HPO_4 1M EDTA, 1% BSA, 7% SDS. Membranes were washed in 2×SSC and 0.1%SDS at 45 °C twice for 15 min and once for 45 min. Film-exposure was carried out at −70 °C. A PCR product encoding ITSI-5.8s rDNA-ITS2, or a partial region of the ITS2 that recognized specifically the upper rDNA band were used as probes in the Southern blot analyses.

Genomic PCR Analysis of Lophocereus *rDNA.* We used a PCR approach to determine if the lower band had a complete 18S sequence indicative of a functional coding region. Conserved primers within the 18S gene (1800F, 1400F, 1055F, 920F, 690F, 300F, Medlin et al. 1988) were used in combination with a reverse primer specific to the lower band (ITS2P-R: 5′-GCAGTGTGAATGATG-3′) for rDNA genomic amplifications. The DNA was isolated from an individual from population 4. Annealing was done at 58 °C and primer extension was for 2.5 min. All other PCR conditions were as described above. PCR products were separated on a 1% agarose gel. The longest fragment obtained (using 1400F) was gel-excised, purified, and sequenced as described above. We also determined the 18S rDNA sequence for the upper PCR band of an individual from population 5 using the conserved 1400F and the upper band-specific ITS1 (5′-GCCACCCCTGCGCACCGGTG-GATTGAG-3′) primers. This PCR product was cloned into pGEM-T Easy (Promega) and sequenced as described above.

Relative Rate Test. The relative rate test (Takezaki et al. 1995) was implemented to determine whether the 5.8S rDNA genes of *Lophocereus* had an accelerated rate of sequence divergence. The program PHYLTEST (V2.0, Kumar 1996) was used to calculate rates of coding region divergence in the *Lophocereus* upper and lower band rDNA sequences in comparison to other members of the Cactaceae and the Portulacaceae with *Silene dioica* as the outgroup. The Kimura two-parameter model (Kimura 1980) with rate variation modeled as a gamma distribution having the shape parameter $\alpha = 0.5$ was used to calculate the pairwise distances.

RT-PCR. We used the RT-PCR method to determine the expression of the 18S coding region in the upper and lower band rDNA operons. Total RNA was isolated from a senita seedling (population 10) using the Qiagen RNA extraction kit according to the manufacturer's instructions. To exclude DNA contamination, RNA was extensively electrophoresed on a 1% Low Melting Point agarose gel and the 18S rRNA band was excised. cDNA synthesis was done using SuperScript II reverse transcriptase (Gibco BRL) and conserved primers within the 18S sequence (B, 1400R, 1055R, Elwood et al. 1985; Medlin et al. 1988). PCR was done using primers within the 18S coding region (300F forward primer paired with B, 1055R, or 1400R reverse primers). PCR products were subcloned into pGEM-T Easy (Promega) and the entire ligation reaction was used for transformation of competent JM109 cells according to standard protocols. Single colonies were picked, grown in liquid medium over night at 37 °C, and plasmid DNA was isolated using a standard alkaline lysis mini-prep protocol (Sambrook et al. 1989). Plasmid DNA was digested with *Eco*RI to confirm the presence of an insert prior to sequencing with the SP6 primer.

Results

Results of PCR Analyses

PCR reactions, using the species-specific and conserved primers within the 18S and 28S genes to amplify ITS1 and ITS2 (and the 5.8S gene), resulted in two bands in all studied individuals from senita. PCR analyses of *Pachycereus marginatus* with the ITS4 and 1800F primers resulted in two bands of same sizes as in senita (ca. 700 bp, 550 bp). A single band of approximate size 700 bp was found in PCRs with *P. weberi, P. pectenaboriginum, P. pringlei, P. pringlei.* In senita, the 700 bp (upper) band was of higher yield whereas the second lower (550 bp) band was only faintly visible (Fig. 2). The upper band was sequenced from a number of individuals and, using the BLAST search (Altschul et al. 1997), shown to encode rDNA sequence. Our initial analyses showed that the upper band sequences varied between senita individuals, so we tested this region as a potential phylogeographic marker.

As we accumulated more sequence data, however, ITS and 5.8S seqences from the upper band of senita did not cluster into geographically distinct populations like those generated using allozyme data. Furthermore, three individuals from different populations (6, 9, 11) consistently formed a distinct clade (see Fig. 3). These individuals were not only more divergent in their ITS1 and ITS2 than the other senita rDNA sequences, but contained six nucleotide differences within the 5.8S gene that differentiated them from the major clade. We hypothesize that these two clades are paralogous rDNA operons in senita that are found in the upper PCR frag-

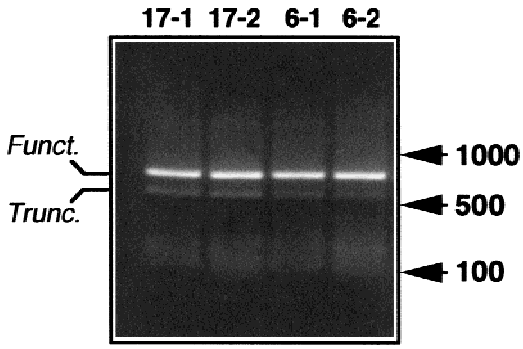

Fig. 2. A Photograph of an agarose gel showing results of the amplification of the ITS1, 5.8S, and ITS2 region in senita using conserved primers within the 18S and 28S genes (1800F and ITS4, respectively). The two bands were found in all studied senita individuals. Shown are the results for two individuals each for populations 17 and 6. *Funct.* is the *Functional* operon, whereas *Trunc.* is the pseudogene *Truncated* operon.

ment. This hypothesis is supported by the remarkable sequence divergence in the 5.8S rDNA sequence of members of these clades. We also believe that there may be additional, as yet undetected, heterogeneity within each of these two paralogs based on the consistent observation of double peaks (ambiguities) at some positions in the sequencing electropherograms.

It was, therefore, clear that without a great deal of effort to clone and characterize all paralogs within the upper PCR band of senita, these rDNA regions could not be used as population markers. We then decided to examine the lower 550 bp band to evaluate its potential as a phylogeographic marker in senita. The lower band sequence was determined from several individuals. This fragment also encoded ITS sequence from senita, although it was found to have a deletion of 157 bp starting after the fourth bp of ITS2 (relative to the upper band). Upper and lower band comparisons were done with rDNA sequences from the same individual (17-2). Within ITS1, there was a 20.7% (34/164) difference between these sequences. Within the 5.8S sequence, there was a remarkable 14.7% (24/163) difference. Within the ITS2 there was a 17.5% (22/126) difference. The lower band ITS1 contained a 36 bp region that was unalignable with this region in the upper band, resulting in a greater number of changes in ITS1 in comparison to the 5.8S sequence. The GC content of the upper ITS1 region was 62.3% (100/160), of its 5.8S gene 50.3% (82/163), and of its ITS2 region, 57.6% (159/276). The GC content of the lower ITS1 region was 65.2% (105/161), of its 5.8S gene 47.2% (77/163), and of its ITS2 region, 52.0% (66/127). These values are within the range of what is found in other angiosperms; i.e., 50–75% for ITS1 and ITS2 (Baldwin et al. 1995).

Lower Band Heterogeneity

RFLP analysis using *Alu*I/*Hinf*I double-digests and cloned PCR products (encoding ITS1-5.8S-ITS2) from

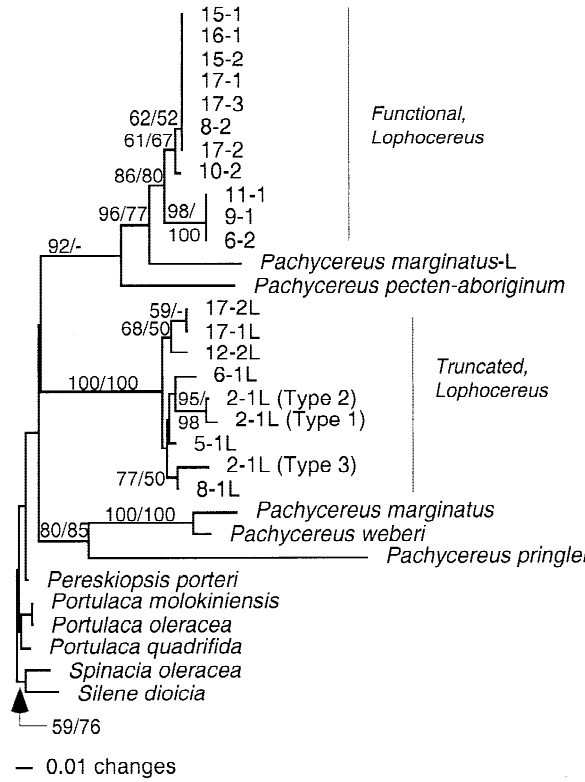

Fig. 3. FastDNAml tree of the 5.8S gene from senita individuals collected in different areas of Baja and of the 5.8S gene of the *Pachycereus* species and other Caryophyllales with *Spinacia* and *Silene* as outgroups. A total of 164 nt (46 parsimony informative sites) were used in the analysis. Results of maximum likelihood bootstrap analysis (100 replications) are shown on the left of the slashmarks above the branches, whereas the bootstrap values shown on the right are from an unweighted maximum parsimony analysis (1000 replications). The locations of the sampled senita in Baja are those shown in Fig. 1. *Functional* is the expressed senita rDNA operon, whereas *Truncated* is the pseudogene operon. Types 1–3 are the different pseudogene sequences shown in Figs. 4, 5.

the lower band of one individual (2-1) showed that there are two dominant sequence types (Type 1, Type 2; see Fig. 4). A third pattern (Type 3) was found in 1/14 clones and encoded a sequence closely related to Type 2 with the only apparent RFLP being the upward shift of the smallest band on this gel. One clone encoding each type of lower band rDNA was sequenced. Types 1 and 2 differed at a total of 12 nt positions (7 in ITS1, 4 in 5.8S, and 1 in ITS2), whereas Types 2 and 3 differed at 2 sites (1 in ITS1 and 1 in 5.8S, Fig. 5). We postulate that the number of changes between Types 1 and 1 and 2 are too great for these to be allelic variants. These two sequences likely represent rDNA paralogs, whereas Types 2 and 3 are likely allelic variants. The apparent band shift is explained by the loss of a *Hinf*I site in the Type 3 ITS that is 17 bp from a shared site in Type 2.

Phylogeny of Upper and Lower Band Sequences

Phylogenetic analyses of the 5.8S rDNA sequences shows a complex history of duplications of *Lophocereus*

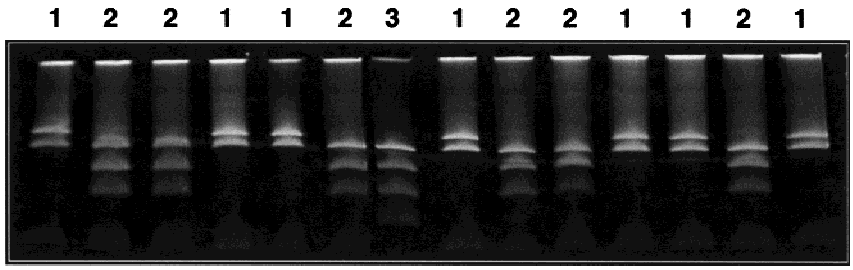

Fig. 4. PAGE gel showing results of the RFLP analysis of cloned PCR products. The ITS1, ITS2, and the 5.8S coding region of an individual from population 2 was cloned and the inserts of 14 clones were isolated with PCR and double-digested with *Alu*I and *Hin*I. The three types of restriction patterns correspond to the three sequence types for the lower band rDNA fragment, labeled 1–3. Types 1 and 2 represent paralog polymorphism within the lower band operon, whereas Type 3 is likely an allelic variant of Type 2 (see text).

```
         5          15         25         35         45
ATGTGGGGAG GACTGCCTCG ACCCCCTTTC CAAGCCGGGA GCAACCCTAG 2-1 (Type 1)
. . . . . C. . . . A. . . . . . . . . . . . . . T. . . . . . . . . . . 2-1 (Type 2)
. . . . . C. . . . A. . . . . . . . . . . . . . T. . . . . . . . . . . 2-1 (Type 3)

        55         65         75         85         95
GCACAGCAAC AAAACCCCAG CGCGAACTGC GCCAAGGAAC ACGAACTCAT 2-1 (Type 1)
. . . . . . . . . . G. . . . . . . . C. . . . . . . . . . . . . . . . 2-1 (Type 2)
. . . . . . . . . . G. . . . . . . . C. . . . . . . . . . . . . . . . 2-1 (Type 3)

       105        115        125        135        145
AGTGTGCCCG CCTGCACCTG GCCCGTTGGC GCGCGGGGGC GGCACCTGTC 2-1 (Type 1)
. . . A. . . . . . G. . . . . . . . . . . . . . . . . . C. . . . . 2-1 (Type 2)
. . . A. . . . . . G. . . . . . . . . . . . . . . . . . T. . . . . 2-1 (Type 3)

       155        165        175        185        195
CCGACTTAAA ACgtaatgat tctcggcaac ggatatctca gctcttgcat 2-1 (Type 1)
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2-1 (Type 2)
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2-1 (Type 3)

       205        215        225        235        245
cgatgaaaaa cgtagtgaaa tgcgatactt agtgtgaatt gcagaatccc 2-1 (Type 1)
. . . . . . . . . . . t. . . . . . . . . g. . . . . . . . . . . . a. . . . . 2-1 (Type 2)
. . . . . . . . . . . t. . . . . . . . . g. . . . . . . . . . . . g. . . . . 2-1 (Type 3)

       255        265        275        285        295
gtgaaccatc gagtctttga acgcaagttg cgcctgaagc cttcaaacta 2-1 (Type 1)
. . . . . . . . . . . . . . . . . . . . . . . . . . . . t. . . c. 2-1 (Type 2)
. . . . . . . . . . . . . . . . . . . . . . . . . . . . t. . . c. 2-1 (Type 3)

       305        315        325        335        345
agggcacgtc tacatgggcg tcacgCATCA CGCACACTGC CGGGAAGGGC 2-1 (Type 1)
. . . . . g . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2-1 (Type 2)
. . . . . g . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2-1 (Type 3)

       355        365        375        385        395
TCGCGGACCC TAACACATGG TGAAAGCACA AAACTATTGA GGCCCCAGGT 2-1 (Type 1)
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2-1 (Type 2)
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2-1 (Type 3)

       405        415        425        435        445
CAAGCGGGGC TACCCGCTGA GTTTAAGCAT ATCAATAAGC GGAGGA      2-1 (Type 1)
. . . . . A. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2-1 (Type 2)
. . . . . A. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2-1 (Type 3)
```

Fig. 5. Alignment of pseudogene sequence Types 1–3 shown in Fig. 4. These sequences were isolated from a single individual from population 2. Identical positions are marked with dots and all nucleotide substitutions are shown. The ITS1 and ITS2 regions are in uppercase, whereas the 5.8S rDNA is in lowercase.

and *Pachycereus* rDNA operons that is consistent with there having been at least two operons in the common ancestor of these taxa (Fig. 3). Distance analysis of the combined 5.8S + ITS1 sequences from a broader sampling of *Lophocereus* upper and lower bands of individuals from Baja substantiates the finding of two distinct operons in this taxon (Fig. 6). Within the 5.8S tree, the well-supported sister group relationship of *P. marginatus* upper and lower bands with two different *Pachycereus* clades is consistent with this scenario.

Southern Blot Analyses

We probed genomic DNA from different senita individuals with the entire ITS1-5.8S-ITS2 fragment to detect the

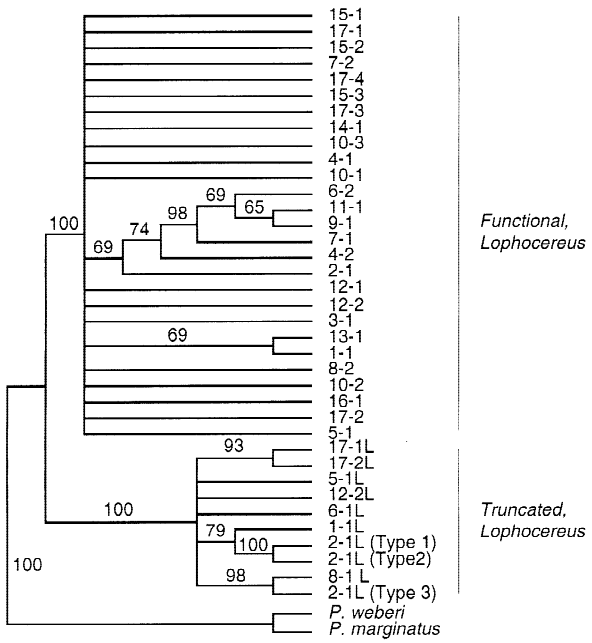

Fig. 6. Bootstrap consensus tree (1000 replications) inferred from a distance analysis (HKY85 model, neighbor-joining) of 5.8S and ITS1 sequences from all sampled senita individuals in Baja (see Table 2). This phylogeny has been rooted on the branch leading to the *Pachycereus* sequences. Only groups found in ≥60% of the bootstrap replicates are shown. The locations of the sampled senita in Baja are those shown in Fig. 1. The *Functional* and *Truncated* operons in senita are identified.

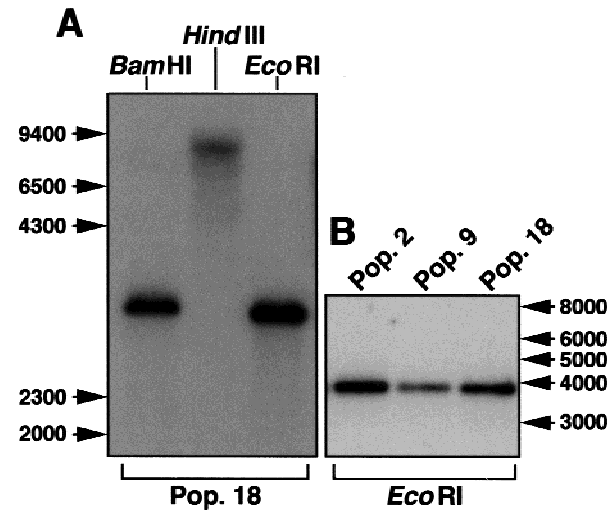

Fig. 7. Southern analyses of rDNA genes in senita. (**A**) Genomic Southern blot of an individual from population 18 that has been digested with *Bam*HI, *Hind*III, and *Eco*RI and probed with a PCR product encoding an entire *Functional* ITS1, 5.8S, ITS2 region. The lambda/*Hind*III molecular weight marker is from NEB. The presence of a distinct band in all digests is consistent with a single locus encoding *Functional* and that *Truncated* exists in too low a copy number to be resolved in this experiment (see text for details). (**B**) Genomic Southern blot of individuals from populations 2, 9, and 18 that have been digested with *Eco*RI and probed with a fragment that spans the ITS2 deletion in *Truncated* (i.e., the probe is *Functional*-specific). The 1 kb ladder is from NEB.

lower band rDNA operon. Because the *Functional* and lower band ITS fragments share a sequence identity of about 80%, a full-length probe derived from this sequence should hybridize to both rDNA regions. A RFLP is, however, expected with the analysis because the upper band contains an *Eco*RI site 248 nt from the 5′ end of the probe that is not found in the lower band DNA. Our prediction was to see two bands in the Southern analysis, corresponding to the two distinct operons. Even after a long (5 days) exposure, however, we were able to detect only a single band (Fig. 7A). This could mean either that one of the operons is in such a low copy number that we are not detecting it with the Southern analysis, or that both operons are closely linked and occur in the single bands. The second scenario is less likely because, in addition to the *Eco*RI polymorphism, there is a 150 bp deletion in the ITS2 region of the lower band operon that should result in a noticeable size shift in agarose gels. To address this issue, we repeated the Southern blot analysis using as probe a 157 bp fragment of the ITS2 in the *Functional* operon that is deleted in the lower band, i.e., the probe was specific for the *Functional* operon. This analysis resulted in the identification of the same band in different populations (Fig. 7B) as when the ITS1-5.8S-ITS2 region was used as the probe. We interpret this as support for the first scenario of a putative low copy number of the lower band operon and hypothesize that its amplification (albeit weakly) results from the greater sensitivity of the PCR method.

The Lophocereus Lower Band Encodes a Pseudogene

To determine if the lower band ITS region in *Lophocereus* is part of a rDNA operon that contains a functional 18S gene, we isolated portions of this coding region with PCR methods. Usage of conserved forward primers (300F, 690F, 920F, 1055F, 1400F, 1800F) within the 18S gene in combination with a reverse primer specific for the lower band (ITS2P-R) failed to amplify the full length 18S gene from this operon. The 5′ end of the 18S gene is apparently deleted because PCR primers upstream of the 1400F primer, in combination with ITS2P-R, did not produce a fragment (Fig. 8). The longest product obtained from these PCRs (1400F+ITS2P-R) was sequenced and confirmed to be the lower band operon. Sequence data for this region of the 18S gene showed 81/100 and 147/161 matches in comparison to the homologous sequence from spinach and the *Lophocereus* upper band, respectively. These data lead us to postulate that the lower band operon (*Truncated*) encodes a pseudogene.

Functional is Expressed in Senita

We used the RT-PCR method to test whether *Functional* is expressed in senita. As described above, gel-purified 18S rRNA was used for cDNA synthesis and RT-PCR using conserved primers within this coding region. RT-

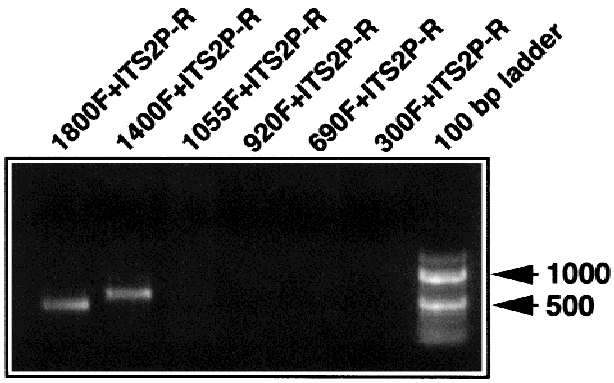

Fig. 8. Genomic PCR analysis of the senita pseudogene operon in an individual from population 4. Only primers 1400F and 1800F, which recognize conserved sequences at the 3′ end of the 18S rDNA gene, paired with the *Truncated*-specific reverse primer ITS2P-R resulted in the amplification of an 18S rDNA fragment. This result suggests that the 5′ end of the 18S coding region has been deleted in senita.

PCR products obtained with primers 300F+1400R or 1055R were cloned and sequenced. All RT-PCR fragments encoded a single sequence which was identical to the homologous genomic region from the senita upper band. This result identifies *Functional* as the expressed coding region in senita.

Relative Rate Tests

Divergence rates comparisons were done of the 5.8S rDNA coding region in the *Lophocereus Functional* (individual 17-2) operon with homologous sequences in other Cactaceae (*Pereskiopsis porteri*) and Portulacaceae (*Portulaca molokiniensis*) with *Silene dioica* as the outgroup. These analyses showed significant rate differences between 17-2 (used to represent all the *Lophocereus* 5.8S genes) and the other ingroup and outgroup sequences. For example, the branch length leading to the 5.8S gene in 17-2 ($L_a = 0.1567$) showed a significant rate elevation in comparison to this gene in *Portulaca* ($L_b = -0.01791$, $L_a - L_b = 0.1746 \pm 0.0698$, $Z = 2.5004$, $p < 0.05$).

Discussion

Senita as a Study System

Our study was originally designed to use ITS1 and ITS2 as markers for phylogeographic analyses of senita. Results from paleoecological studies suggest that the Sonoran Desert flora, including senita, recolonized northern Baja from Pleistocene refugia in southern Baja and/or Sonora within the last 2000–4000 years (Peñalba and Devender 1998). In a previous study, allozyme markers showed directionality in historical pathways of gene migration in senita. In fact, over 60% of the variation in estimated migration rates among senita populations in Baja can be explained by a hypothesis of isolation by linear (north–south) distance (J. N. unpublished data). The rDNA analyses were designed to test this hypothesis and to provide an independent corroboration of the allozyme data.

Evolution of Paralogous rDNA Sequences within Functional and Truncated

Senita has two distinct rDNA operons, one that is functional and one that is a pseudogene. *Functional* clearly demonstrates paralog polymorphism, and we postulate that the same holds for *Truncated* (see Figs. 4, 5). The paralog polymorphism may have resulted from recombination during recent range expansion and nothward migration, or high levels of current pollen- or seed-mediated gene flow. Gene flow may introduce new rDNA variants into populations at a rate faster than their homogenization (Sanderson and Doyle 1992). Analyses of 32 allozyme loci from 12 Baja populations of senita, however, indicate unusually low rates of gene flow relative to other long-lived woody species, including Baja populations of the columnar cacti *Pachycereus pringlei* and *Stenocereus thurberi* (Hamrick et al. in press). Regardless of the source of the rDNA variants, these results are particularly striking because *Functional* was resolved as a single strong band in the Southern analysis. This suggests that the *Functional* repeats may be located at a single chromosomal site and would, therefore, be expected to undergo rapid concerted evolution (Schlötterer and Tautz 1994). Regarding the possible origin of the divergent lower band rDNA sequence from a contaminant, this can likely be excluded because the plant tissue used in our analyses was derived from anthers inside unopened senita flowerbuds or stem tissue trimmed with a sterile razor blade. Furthermore, database searches using the lower band sequence(s) failed to find strong similarity to any published rDNA sequence and identified a similar spectrum of taxa as searches using the upper band rDNA sequences. Results of the similarity searches also allowed us to exclude the possibility that the lower band was of organellar origin. This leads us to hypothesize that the lower band encodes a distinct rDNA operon in senita.

Origin of the Senita Truncated Operon

We postulate that *Truncated* was present in the common ancestor of *Lophocereus* and *Pachycereus* and has maintained an evolutionary history that is distinct from the *Functional* operon. The high sequence divergence between *Functional* and *Truncated* suggests that recombination does not take place between these repeats. This may reflect the relatively large deletions in *Truncated.* Alternatively, *Truncated* may occur in a genomic region

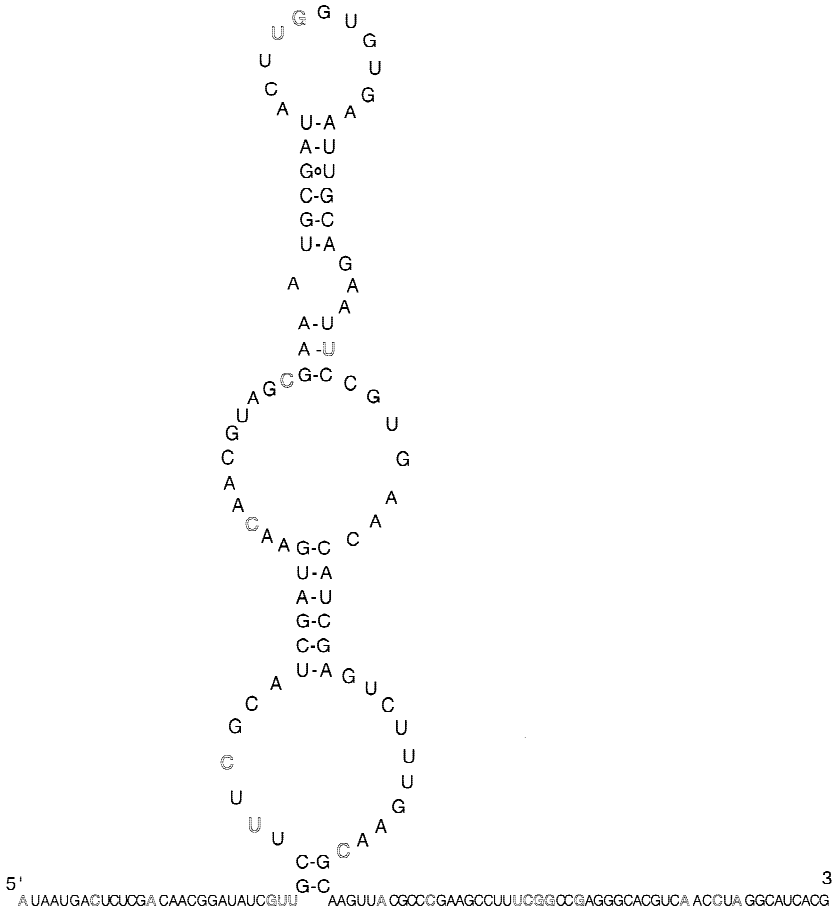

Fig. 9. Putative secondary structure of the 5.8S rRNA from the *Functional* operon. This sequence was isolated from an individual in population 17. The outlined sites differ between the *Functional* and *Truncated* operons in this individual.

that is protected from recombination and/or occurs in very low copy numbers (consistent with the Southern blot analysis). A similar situation exists in *Zea* which contains a functional and a pseudogene rDNA operon, both of which were present prior to the divergence of modern *Zea* (Buckler IV and Holtsford 1996a; Buckler et al. 1997).

To determine the evolutionary origin of *Truncated* in senita, we analyzed related cactus species for the presence of two rDNA operons. Morphological studies (Gibson and Horak 1978) and recent molecular evidence from chloroplast sequence data (S. H. unpublished data) point to the hummingbird pollinated *Pachycereus marginatus* as the sister group of *Lophocereus.* Using conserved PCR primers within the 18S and 28S genes to amplify ITS1 and ITS2 (and the 5.8S gene), we obtained and sequenced two bands from *P. marginatus* that were of the same size as the two PCR products from *Lophocereus* (a stronger band of approximately 700 bp, and a fainter band of ca. 550 bp). We tested *Pachycereus weberi, P. pecten-aboriginum,* and *P. pringlei,* three columnar cacti in the tribe Pachycereeae which are the most closely related to *P. marginatus* and *Lophocereus* (Gibson and Horak 1978; S. H. unpublished data), for the presence of two rDNA operons. Using conserved primers as described above, we were able to amplify only one major band (ca. 700 bp) from these *Pachycereus* species.

We used the 5.8S region to determine the evolutionary history of these operons. Phylogenetic analysis using sequences derived from the upper and lower bands of *P. marginatus* showed that the upper band 5.8S rDNA clustered with the lower band sequences of *Lophocereus,* whereas the lower *P. marginatus* band formed a distinct group that included the single sequences isolated from *P. pringlei* and *P. weberi* (Fig. 3). We did not further characterize these operons within *P. marginatus* but speculate that the lower band may be a non-functional, truncated rDNA operon.

Interestingly, sequence relationships within the two *Lophocereus*/*Pachycereus* clades (i.e., *Lophocereus functional*/*P. marginatus*-L/*P. pecten-aboriginum* and *P. marginatus*/*P. pringlei*/*P. weberi*) are consistent with chloroplast DNA phylogenies of these taxa (S. H. unpublished data). Sequence comparisons of non-coding regions between the following genes *trn*L–*trn*F, *trn*C–*trn*D, *trn*S–*trnf*M show that *Lophocereus* and *P. marginatus* share a sister group relationship to the exclusion of *P. pecten-aboriginum* and that *P. weberi* and *P. marginatus* are more closely related to each other than either is to *P. pringlei.* These data suggest that different rDNA paralogs may have been lost in *P. pecten-aboriginum, P. pringlei,* and *P. weberi,* whereas *Lophocereus* and *Pachycereus marginatus* have maintained both ancestral operons. Alternatively, all *Pachycereus* species may re-

tain the ancestral rDNA operons but we are unable to amplify the second paralog in *P. pecten-aboriginum, P. pringlei,* and *P. weberi* because of extreme sequence divergence at the primer-binding sites or the evolution of nucleotide bias leading to hairpins making these sequences resistant to PCR amplification. That all studied *Lophocereus* and *Pachycereus* rDNA sequences have high relative sequence divergence rates is demonstrated by a comparison of corrected (HKY85 model) distance values between the taxa shown in Fig. 3. The highest distance estimate among the non-*Lophocereus*/*Pachycereus* sequences is 3.85% between *Portulaca quadrifida* and the outgroup *Silene diocia,* whereas between *P. pringlei* and *Lophocereus* 8–2, this value is a magnitude higher at 30.43%. Secondary structure analysis of the 5.8S rRNA from the senita *Functional* operon (Fig. 9) shows that, in spite of a high relative divergence rate, this coding region can be folded into a structure typical of plant 5.8S rRNAs (Yeh and Lee 1991; Hershkovitz and Lewis 1996). The 24 positions that differ between the 5.8S rRNA in *Functional* and *Truncated* are as outlined text in Fig. 9. These changes are dispersed throughout the secondary structure.

In conclusion, our study shows the extreme level of paralogy that is possible in plant rDNA genes. Although other taxa may not have rDNA heterogeneity to the extent found in senita, our results underline the caution that must be used when using rDNA genes in population analyses. A significant finding in senita is the existence of high levels of coding region heterogeneity both within and between operons. These data are a model for understanding the dynamics of concerted evolution in a system that is likely characterized by extensive ancestral polymorphisms, historical gene flow between populations, and/or existence of hybridization during species evolution. However, the lack of evidence for polyploidy at 32 allozyme loci (Hamrick et al. in press) suggests that *Truncated* does not result from a duplication of the genome. Our analysis of rDNA of related cactus species suggests that the second operon was present in the common ancestor of *Lophocereus* and *Pachycereus marginatus.*

Acknowledgments. We acknowledge T. Fleming, J. Hamrick, and T. Van Devender for discussions about senita. D. B. acknowledges support from the Carver Foundation and a CIFRE Award. S. H. thanks Sigma XI for a Grant-in-Aid of Research. We thank M. Harris, D. Fergus, and R. Bhatnagar for help in collecting cactus tissues and two reviewers for their helpful comments.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Ann Miss Bot Garden 82:247–277

Buckler IV ES, Holtsford TP (1996a) *Zea* ribosomal repeat evolution and substitution patterns. Mol Biol Evol 13:623–632

Buckler IV ES, Holtsford TP (1996b) *Zea* systematics: ribosomal ITS evidence. Mol Biol Evol 13:612–622

Buckler ES, Ippolito A, Holtsford TP (1997) The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics 145:821–832

Campbell CS, Wojciechowski MF, Baldwin BG, Alice LA, Donoghue MJ (1997) Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier agamic* complex (Rosaceae). Mol Biol Evol 14:81–90

Dover G (1982) Molecular drive: a choesive mode of species evolution. Nature 299:111–117

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Elwood HJ, Olsen GJ, Sogin ML (1985) The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytrichia nova* and *Stylonychia pustulata.* Mol Biol Evol 2:399–410

Felsenstein J (1985) Confidence limits on phylogenies an approach using the bootstrap. Evolution 39:783–791

Gibson AC, Horak KE (1978) Systematic anatomy and phylogeny of mexican columnar cacti. Ann Miss Bot Garden 65:999–1057

Hamby RK, Zimmer EA (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, New York, pp. 50–91

Hamrick JL, Nason JD, Fleming TH (in press) Genetic diversity of columnar cacti. In: Fleming TH, Valiente Banuet A (eds), Evolution, ecology and conservation of columnar cacti and their mutualists. University of Arizona Press, Arizona

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Hershkovitz MA, Lewis LA (1996) Deep-level diagnostic value of the rDNA-ITS region. Mol Biol Evol 13:1276–1295

Hershkovitz MA, Zimmer EA (1997) On the evolutionary origins of the cacti. Taxon 46:217–242

Hershkovitz MA, Hahn WJ, Zimmer EA (1999) Ribosomal DNA and angiosperm evolution. In: Hollingsworth PM, Bateman RM, Gornall R (eds), Molecular systematics and plant evolution. Taylor and Francis, London

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kumar S (1996) PHYLTES: A program for testing phylogenetic hypotheses. Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University

Lindsay G (1963) The genus *Lophocereus.* Cactus Succulent J 35:176–192

Mayer MS, Soltis PS (1999) Intraspecific phylogeny analysis using ITS sequences: insights from studies of the *Streptanthus glandulosus* complex (cruciferae). Syst Bot 24:47–61

Medlin L, Elwood H, Stickel S, Sogin M (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71:491–9

Miller BR, Crabtree HM, Savage HM (1996) Phylogeny of fourteen *Culex* mosquito species, including the *Culex pipiens* complex, inferred from the internal transcribed spacers of ribosomal DNA. Insect Mol Biol 5:93–107

Olsen GJ, Matsuda H, R. H, R O (1994) fastDNAmL: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput Appl Biosci 10:41–48Peñalba MC, Devender TRV (1998) Cambios de vegetacíon y clima en Baja California, México, durante los últimos 20,000 años. Geol Noroeste 2:21–23

Rich SM, Rosenthal BM, Telford SR, III, Spielman A, Hartl DL, Ayala FJ (1997) Heterogeneity of the internal transcribed spacer (ITS-2) region within individual deer ticks. Insect Mol Biol 6:123–129

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanderson MJ, Doyle JJ (1992) Reconstruction of organismal and gene phylogenies from data on multigene families concerted evolution homoplasy and confidence. Syst Biol 41:4–17

Schlötterer C, Tautz D (1994) Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. Curr Biol 4:777–783

Soltis PS, Soltis DE (1998) Molecular evolution of 18S rDNA in angiosperms: implications for character weighting in phylogenetic analysis. In: Soltis DE, Soltis PS, JJ D (eds) Molecular systematics of plants II: DNA sequencing. Kluwer Academic Publishers, Boston, pp. 188–248

Suh Y, Thien L, Reeve H, Zimmer, EA (1993) Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. Am J Bot 80:1042–1055

Swofford DL (2000) PAUP: Phylogenetic analysis using parsimony. Smithsonian Institution, Washington, DC

Takezaki N, Razhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. Mol Biol Evol 12:823–833

Yeh LCC, Lee JC (1991) Higher-order structure of the 5.8S ribosomal RNA sequence within the yeast 35S precursor ribosomal RNA synthesized *in Vitro*. J Mol Biol 217:649–660