Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 7987–7992, July 1997
Evolution

Phylogeny of the genus *Pistacia* as determined from analysis of the chloroplast genome

(DNA/PCR/pistachio/restriction fragment length polymorphism)

DAN E. PARFITT AND MARIA L. BADENES

Department of Pomology, University of California, Davis, CA 95616

Communicated by S. J. Peloquin, University of Wisconsin, Madison, WI, March 10, 1997 (received for review June 1, 1996)

ABSTRACT **Classification within the genus *Pistacia* has been based on leaf morphology and geographical distribution. Molecular genetic tools (PCR amplification followed by restriction analysis of a 3.2-kb region of variable chloroplast DNA, and restriction fragment length polymorphism analysis of the *Pistacia* cpDNA with tobacco chloroplast DNA probes) provided a new set of variables to study the phylogenetic relationships of 10 *Pistacia* species. Both parsimony and cluster analyses were used to divide the genus into two major groups. *P. vera* was determined to be the least derived species. *P. weinmannifolia*, an Asian species, is most closely related to *P. texana* and *P. mexicana*, New World species. These three species share a common origin, suggesting that a common ancestor of *P. texana* and *P. mexicana* originated in Asia. *P. integerrima* and *P. chinensis* were shown to be distinct whereas the pairs of species were monophyletic within each of two tertiary groups, *P. vera*:*P. khinjuk* and *P. mexicana*:*P. texana*. An evolutionary trend from large to small nuts and leaves with few, large leaflets to many, small leaflets was supported. The genus *Pistacia* was shown to have a low chloroplast DNA mutation rate: 0.05–0.16 times that expected of annual plants.**

Very little information is available concerning the species relationships within the genus *Pistacia*. A general survey of relationships among the Anacardiaceae has been done based on floral morphology (1). The most complete summary and taxonomic descriptions for *Pistacia* are provided by Zohary (2). Zohary described 11 *Pistacia* species, which he divided into four sections: *Lentiscella* Zoh., containing *P. mexicana* HBK, and *P. texana* Swingle; *Eu Lentiscus* Zoh., containing *P. lentiscus* L., *P. saporte* Burnat., and *P. weinmannifolia* Poisson; *Butmela* Zoh., containing *P. atlantica* Desf.; and *Eu Terebinthus*, containing *P. chinensis* Bge., *P. khinjuk* Stocks, *P. palaestina* Bois., *P. terebinthus*, and *P. vera* L. Chromosome counts were performed on three species (2): *P. lentiscus* with 2n = 24, *P. atlantica* with 2n = 28, and *P. vera* with 2n = 30. *P. integerrima* Stewart was considered to be a variety of *P. chinensis* by Zohary (2). An additional species, *Pistacia aethiopica* Kokwaro, has been described by Kokwaro and Gillett on the basis of leaf morphology and tree size (3). The authors consider this species to be synonymous with samples described as *P. lentiscus* L. var. *emarginata* Engl.; however, based on their description, the species could also be a variety of *P. atlantica*.

Leaf characteristics (size, shape, number of leaflets, pubescence, wings on rachis, and presence or absence of terminal leaflet) have been the primary characters used for taxonomic identification of *Pistacia* sp. Wood anatomy also has been evaluated as an identification tool (4) as well as reference to historic distribution patterns. Floral characters have been used less for identification in *Pistacia* but have been used above the genus level (1). This is surprising because there is considerable variability for inflorescence structure among species and flowering dates when grown at the same location (D.E.P., unpublished work). Nut morphology also varies among species. However, for many of the species, these distinctions are more difficult to evaluate.

Pistacia. Species easily form interspecific hybrids, suggesting a very close relationship and raising questions about the accuracy of the reported chromosome counts. F_1 plants are relatively easily generated (all *Pistacia* sp. are dioecious), but F_2 plants are more difficult to produce (D.E.P., unpublished work). The actual level of speciation and relationships within the genus *Pistacia* remains unclear. A better understanding of these relationships is a prerequisite to their use for plant improvement or genetic studies.

Variation at the DNA level in the chloroplast genome has been used since the early 1980s to study species relationships (5). Chloroplast DNA (cpDNA) is ideal for this purpose because it is highly conserved. Variation in most of the nuclear genome is less useful because higher mutation rates can produce significant variation within species at a given locus (6). DNA analysis can be especially useful where the effects of environment can alter observed morphological characters.

Many phylogenetic studies have been done using restricted cpDNA, either separately isolated or as part of a total DNA preparation, probed with either a homologous isolated, restricted, and labeled cpDNA fraction or with cloned cpDNA-labeled probes (7, 8). Well characterized cpDNA libraries have been established, including the tobacco library developed by Olmstead and Palmer (9) and used in this study.

Arnold *et al.* (10), Liston *et al.* (11), and Rieseberg *et al.* (12) have described a procedure using PCR analysis of a hypervariable 3.2-kb fragment of the chloroplast genome for phylogenetic study. Ogihara *et al.* (13) determined that this region is relatively mutable and thus an especially suitable region for PCR amplification and restriction analysis. Major advantages to this approach are speed of analysis and the use of nonradioactive visualization. A potential drawback is the possibility that the nature of the mutations in this region is significantly different than that of those occurring elsewhere in the chloroplast genome. However, having no previous reason to believe that mutations occurring in this region are qualitatively different than other mutations, we chose to use this region for phylogeny construction in conjunction with conventional restriction fragment length polymorphism analysis of the chloroplast genome.

MATERIALS AND METHODS

Plant Materials. Total DNA was isolated from 5 gm of fresh leaves of male and female *P. atlantica*, *P. chinensis*, *P. khinjuk*, *P. lentiscus*, *P. mexicana*, *P. terebinthus*, *P. texana*, *P. vera*, *P. weinmannifolia*, *P. integerrima*, and *Schinus molle* L. using the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/947987-6$2.00/0
PNAS is available online at http://www.pnas.org.

Abbreviation: cpDNA, chloroplast DNA.

hexadecyltrimethyl-ammonium bromide method of Doyle and Doyle (14); 150–600 μg was obtained. *Schinus molle* L., located within the Anacardiaceae, was used as an outgroup taxon for statistical analyses. Species were verified based on collection records (available by request from DEP), comparison of observed leaf and seed morphology with Zohary (2), and observation of flowering dates and morphology over a 7-year period. *P. integerrima* was evaluated as a distinct species although Zohary (2) classified it as a variety of *P. chinensis*.

PCR Amplification and Analysis of Hypervariable cpDNA Region. A 3.2-kb region of relatively variable cpDNA (13) bounded by the conserved sequences 5′-ATGTCACCA-CAAACAGAAACTAAAGCAAGT-3′ [rbcL] and 5′-ACTACAGATCTCATACTACCCC-3′ [ORF 106] (9) was amplified via PCR as described in Arnold *et al.* (10); 0.5–1 μg DNA was used per reaction. Primers for the flanking sequences were synthesized by Operon Technologies (Alameda, CA). Complete amplification was obtained for all 21 samples. Twenty seven restriction endonucleases (*Alu*I, *Ase*I, *Bgl*II, *Bam*HI, *Bst*NI, *Cla*I, *Dde*I, *Dra*I, *Eco*RI, *Eco*RV, *Hha*I, *Hin*dIII, *Hin*fI, *Hpa*I, *Kpn*I, *Mbo*I, *Msp*I, *Pal*I, *Pst*I, *Rsa*I, *Sal*I, *Sa*I, *Scr*FI, *Sty*I, *Taq*I, *Xba*I, and *Xho*I) were tested for their ability to reveal polymorphisms within the amplified fragment among the 21 samples. Initially several common 6-base restriction enzymes were tested and did not reveal polymorphisms. Additional 4- and 5-base restriction enzymes, reported to have high restriction frequencies for cpDNA and mtDNA (5), were used. Approximately 1 μg of amplified, restricted DNA was electrophoresed in 2% agarose gels at 50 V for 4 h (Fig. 1). After the observation that small fragments were not being detected with the 2% agarose gels, all digestions with 4-base restriction enzymes were run on 5% FMC Metaphor agarose, which permitted resolution of fragments as small as 40-bp (Fig. 2). Restriction sites for 17 restriction endonucleases were mapped to the amplified fragment via double digestion (Fig. 3) and were referenced to the reported tobacco sequence in this region to determine the orientation of the map (15).

Restriction Analysis of Chloroplast Genome. The entire chloroplast genome was analyzed for cpDNA restriction site polymorphisms using the general procedures described by Dowling *et al.* (5). Total DNA was isolated (14), and 10 μg

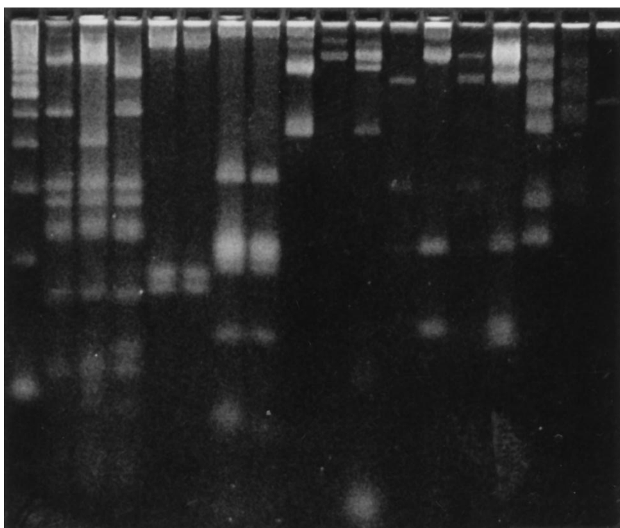

Fig. 1. *Pistacia* polymorphic cpDNA restriction fragments from 3.2-kb amplified region in a 2% agarose gel. Lanes 1–18 with enzyme and restricted cpDNA, respectively: 100-bp molecular weight marker ladder, *Hin*fI:*P. vera*, *Hin*fI:*P. weinmannifolia*, *Hin*fI:*P. atlantica*, *Pal*I:*P. vera*, *Pal*I:*P. weinmannifolia*, *Rsa*I:*P. vera*, *Rsa*I:*P. weinmannifolia*, *Sty*I:*P. weinmannifolia*, *Sty*I:*P. lentiscus*, *Sty*I:*P. vera*, *Bst*NI:*P. lentiscus*, *Bst*NI:*P. vera*, *Scr*FI:*P. lentiscus*, *Scr*FI:*P. vera*, *Dde*I:*P. weinmannifolia*, *Dde*I:*P. vera*, and *Hin*dIII:λ DNA.

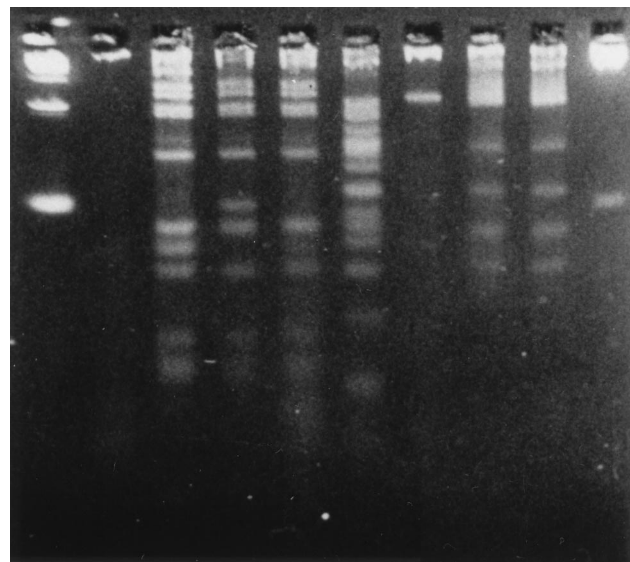

Fig. 2. Improved resolution of <200-bp fragments in 5% FMC Metaphor agarose gel. Lanes 1–10, left to right, with restriction enzyme and restricted cpDNA, respectively: 100-bp molecular weight marker ladder, undigested:*P. vera*, *Hin*fI:*P. weinmannifolia*, *Hin*fI:*P. atlantica*, *Hin*fI:*P. vera*, *Mbo*I:*P. vera*, *Msp*I:*P. vera*, *Alu*I:*P. vera*, *Alu*I:*P. weinmannifolia*, and *Hin*dIII:λ DNA.

DNA was digested with *Bam*HI, *Eco*RI, *Hin*dIII, *Xba*I, and *Hae*III, separated in 1.2% agarose gels at 20 V for 18 h, and blotted to Amersham Hybond N membranes (16) followed by

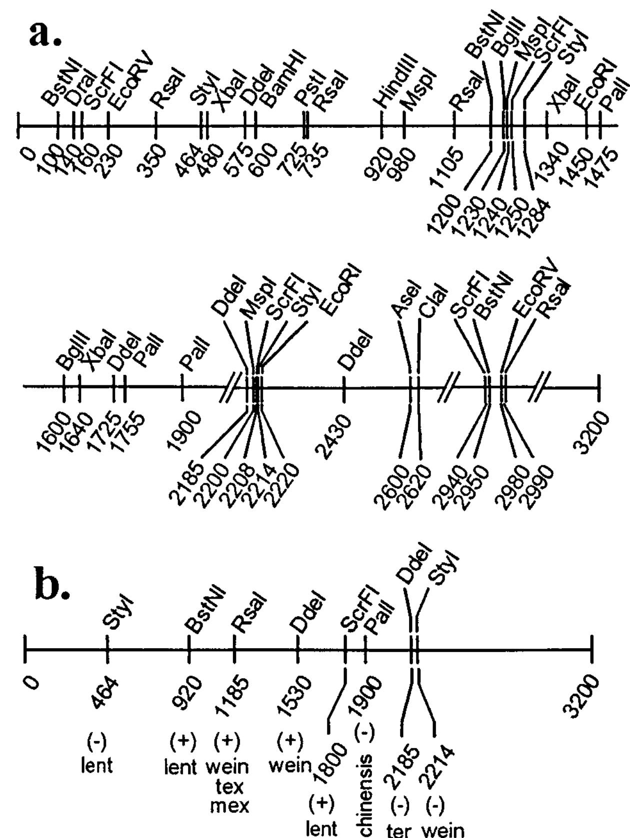

Fig. 3. Map of restriction sites within *Pistacia* cpDNA 3.2-kb amplified region. (*a*) Location of all identified restriction sites in *P. vera*. (*b*) Location of mutations *Dde*I-1, *Dde*I-2, *Bst*NI-4, *Sty*I-5, *Sty*I-6, *Rsa*I-7, and *Pal*I-8; +, site gain, −, site loss referenced to *P. vera*. Mutations *Hin*fI-10 and *Hin*fI-11 were not mapped. lent, *P. lentiscus*; wien, *P. weinmannifolia*; tex, *P. texana*; mex, *P. mexicana*.

UV cross-linking. The clones from the cpDNA library constructed from *Nicotiana tabacum* (9) were used as probes. The cpDNA probes were labeled with digoxigenin dUTP by random priming. Labeled probe was reused three times. Hybridization was carried out overnight at 65°C as described in the Genius system (Boehringer Mannheim). Hybridized membranes were washed under high stringency conditions (0.1% SDS/0.5X SSC at 65°C) and detected with Lumiphos 530′ (Boehringer Mannheim) by exposure for 30 min to x-ray film. Membranes were reprobed up to 5 times after stripping with 0.2 M NaOH at 40°C for 30 min.

Data Analysis. Unweighted pair group mean analysis cluster analysis using Neis and Rogers distances were performed with BIOSYS-1, Ver. 1.7 (17) using average linkage. Dollo, mixed Dollo (sequence rearrangements were considered to be unordered), and Wagner (unordered) parsimony analyses were performed with PAUP, ver. 3.0 (18) using the branch and bound search algorithm with simple addition to find the most parsimonious unrooted tree. Duplicate bootstrap trees, 500 replicates each, were performed with simple addition and tree-bisection-reconnection branch swapping for each of the PAUP procedures. A distance Wagner tree was derived with BIOSYS-1 using the Rogers distance matrix as well as neighbor joining trees with NTSYS-PC, Ver. 1.80 (19) using the index of Nei and Li (20).

RESULTS

PCR-Amplified cpDNA. Twenty seven restriction endonucleases were surveyed, of which 26% revealed polymorphisms; 29.5% of the 95 scored fragments were polymorphic among *Pistacia* sp. They were mapped to nine mutation sites (Fig. 3; Table 1). Seven additional mutation sites were identified in the outgroup, *Schinus molle*. No cpDNA polymorphisms were observed within species or between male and female genotypes within species. This suggests that within species cpDNA variation and especially cpDNA sexual dimorphism may not be present in *Pistacia.*

Assuming one nucleotide change per mutational event, a per nucleotide mutation frequency of $1.4 \times 10^{-2} \pm 2.0 \times 10^{-5}$ (21), was observed for the 708 nt surveyed. Possible length mutations or DNA rearrangements are not considered in this estimate. Ogihara *et al.* (13) found a 1:4 ratio of deletions-to-base substitutions.

Liston *et al.* (11) identified several length mutations during cpDNA PCR analysis of the genus *Datisca*. Length differences also were observed for our 4- and 5-base enzyme digestions (*Hin*fI, *Mbo*I, *Rsa*I). However, the missing low molecular weight fragments were recovered in the Metaphor agarose gels. Several mutation sites associated with low molecular weight fragments were resolvable only with the Metaphor agarose (Fig. 2). All of the DNA from the original 3.2-kb sequence for all species and restriction enzyme combinations was accounted for using both standard and Metaphor agarose gels. A length mutation detected with *Rsa*I appeared to be a small inversion, but confirmation is needed.

Restriction Analysis of Chloroplast Genome. cpDNA restricted with five restriction enzymes in combination with 40 probes provided a total of 183 bands for analysis within *Pistacia*. Six mutational events were detected among *Pistacia* species, of which two were site mutations (Table 2). The other four were apparent rearrangements characterized by the net gain or loss of a band, suggesting a gain or loss of a DNA sequence (fragment sizes for polymorphic fragments available from D.E.P. on request). Eleven additional mutations were scored for *Schinus molle*. The mutation frequency for the chloroplast genome was $2.5 \times 10^{-4} \pm 1.3 \times 10^{-7}$ using only the two site mutations for the 1972 nt surveyed.

The 15 mutations from the combined data were sufficient to differentiate the 10 species into eight groups with six species characterized uniquely (22). However, two pairs of species, *P. mexicana*:*P. texana* and *P. khinjuk*:*P. vera*, were not polymorphic for any of the mutations and could not be separated.

Statistical Analysis. Dollo and mixed Dollo parsimony analyses (Fig. 4*a*) produced identical single maximally parsimonious trees, both with a tree length of 33 and a consistency index of 0.97. This tree is consistent with the tree from the NTSYS-PC neighbor joining analysis using the coefficient of Nei and Li (1979). The *Pistacia* species are divided into two monophyletic groups, one ("*Lentiscus*") containing *P. lentiscus*, *P. weinmannifolia*, *P. mexicana*, and *P. texana* and the other containing the remaining six species ("*Terebinthus*"). *P. chinensis* and *P. integerrima* formed a group that is monophyletic with *P. terebinthus*, *P. atlantica*, and *P. khinjuk*:*P. vera*. *P. weinmannifolia* and *P.mexicana*:*P.texana* formed a group that was monophyletic with *P. lentiscus*. Wagner (unordered) parsimony produced three trees, two of which were identical (due to a zero length arm) with the *P. chinensis* and *P. integerrima* furcation point ancestral to *P. terebinthus*, *P. atlantica*, and *P. khinjuk*:*P. vera* (Fig. 4*b*) and one of which was identical to the tree in Fig. 4*a*. The trees in Fig. 4 *b* and *c* also differed from Fig. 4*a* with respect to the placement of *P. lentiscus*, which shares a common furcation with group *Terebinthus* and the other group *Lentiscus* members. The distance Wagner tree was identical to the tree in Fig. 4*b*. The SD was 2.7%, $F = 1.923$, and the cophenetic correlation was 0.99 for the distance Wagner tree. The 50% majority rule bootstrap consensus tree for unordered parsimony was the same as for the tree in Fig.

Table 1. Character states for site mutations in the 3.2-kb cpDNA region bounded by rbcL and ORF 106 for 12 *Pistacia* sp. and *Schinus molle* outgroup

	Site mutations														
Species	*Dde*I 1	*Dde*I 2	*Bst*NI 4	*Sty*I 5	*Sty*I 6	*Rsa*I 7	*Pal*I 8	*Hin*fI 9	*Hin*fI 10	*Hin*fI 11	*Msp*I 12	*Ase*I 13	*Xba*I 14	*Xba*I 15	*Eco*RI 16
P. lentiscus	0	1	1	1	0	0	1	0	1	0	0	0	0	0	0
P. weinmannifolia	1	1	0	0	1	1	1	0	1	0	0	0	0	0	0
P. chinensis	0	1	0	1	1	0	0	0	1	0	0	0	0	0	0
P. integerrima	0	1	0	1	1	0	1	0	1	0	0	0	0	0	0
P. terebinthus	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0
P. atlantica	0	1	0	1	1	0	1	1	0	0	0	0	0	0	0
P. mexicana	0	1	0	1	1	1	1	0	1	0	0	0	0	0	0
P. texana	0	1	0	1	1	1	1	0	1	0	0	0	0	0	0
P. khinjuk	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0
P. vera	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0
Schinus molle	0	1	0	1	1	0	1	0	0	1	1	1	1	1	1

Letters designate restriction endonuclease used; mutation sites are designated by numbers below. 0, loss of mutation site; 1, gain of mutation site.

Table 2. Character states for site and length mutations in the chloroplast genome as determined from Southern analysis

	Site mutations					Length mutations											
	*Bam*HI	*Eco*RI	*Eco*RI	*Pal*I	*Bam*HI	*Eco*RI	*Hin*dIII	*Hin*dIII	*Hin*dIII	*Xba*I	*Bam*HI	*Eco*RI	*Pal*I	*Eco*RI	*Bam*HI	*Eco*RI	*Pal*I
Species	5,6	3+4	13	20	38	7+8+9	11	13	20b,21	3+4	5	10	11	13	21	25	27
P. lentiscus	0	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0
P. weinmannifolia	1	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0
P. chinensis	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0
P. integerrima	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0
P. terebinthus	0	1	1	1	1	1	0	1	1	0	0	0	1	0	0	0	0
P. atlantica	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0
P. mexicana	1	0	1	1	1	0	1	0	0	0	0	0	1	0	0	0	0
P. texana	1	0	1	1	1	0	1	0	0	0	0	0	1	0	0	0	0
P. khinjuk	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0
P. vera	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0
Schinus molle	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1

Letters designate restriction enzymes used; numbers denote probe number as described in Olmstead and Palmer (9). Commas indicate probes were used in separate analyses; + indicates that probes were combined and used in a single hybridization. 0, loss of mutation site; 1, gain of site.

4*b*, with a consistency index of 0.97 and a length of 33. The 50% majority rule consensus bootstrap tree for the mixed Dollo analysis (Fig. 4*c*) with a consistency index of 0.94 and length of 34 differed from the maximally parsimonious mixed Dollo tree of Fig. 4*a* with a common furcation for *P. lentiscus*, the *P. weinmannifolia*:*P.menixana* group, and the *Terebinthus* group as occurred with the tree in Fig. 4*b*. All figures are shown with zero length branches collapsed. Cladograms for the two bootstrap trees, which did not show branch length, were identical. The trees shown in Figs. 4 *b* and *c* suggest that *P. lentiscus* is distinct from groups *Lentiscus* and *Terebinthus*. The outgroup is correctly positioned in all of the analyses. The members of two tertiary monophyletic groups, *P. mexicana*:*P. texana* and *P. vera*:*P. khinjuk*, were inseparable—they had no mutational differences. Unweighted pair group mean analyses from Nei's and Rogers' distance data for BIOSYS-1 and NTSYS-PC differ from the preceeding analyses by placing *P. terebinthus* in a position ancestral to the remaining group *Terebinthus* members. The unweighted pair group mean analysis (Nei's distance, BIOSYS-1) SD was 26% with a cophenetic correlation of 0.98 and $F = 12.3$.

DISCUSSION

DNA Fragment Separation and Detection Procedures. Metaphor agarose enhances analysis of small DNA fragments from 40 to 200-bp, complementing the size range (200 bp to 3 kb) separated by electrophoresis through conventional agarose and separating small DNA fragments that are not seen in standard agarose gels. Small DNA fragments may not always be detected during Southern blotting because the small pore size of the gel matrix can inhibit DNA transfer from the gel to the membrane. The use of nonradioactive labeling of the tobacco probes was very successful and was functionally equivalent to ^{32}P-labeling of cpDNA restriction fragment length polymorphisms because of the strong signal from the multiple copies of DNA on the membrane.

Phylogeny of *Pistacia*. *Species*. The mixed Dollo tree (Fig. 4*a*) was considered to be the most correct phylogeny because a unique tree was obtained, and the assumptions used in the analysis were most appropriate for the data. Dollo parsimony is an appropriate analytical method for restriction site analysis (23) whereas the mixed Dollo analysis recognizes that length mutations could result from either gain or loss of restriction sites. The result from the mixed Dollo analysis was generally consistent with both unordered parsimony analysis, neighbor joining, and distance Wagner analysis (Fig. 4*b*). Although not the only realistic arrangement, it is probably the best description of the cpDNA data and is very similar to the most likely alternative arrangement shown in Fig. 4*b*. All of the phylograms shown in Fig. 4 support the following conclusions.

Two pairs of species could not be separated with the 16 mutations: *P. mexicana*:*P. texana* and *P. khinjuk*:*P. vera*. In each case, Zohary clearly defined the members as distinct species, and our samples of *P. mexicana* vs. *P. texana* and *P. khinjuk* vs. *P. vera* differed in leaf morphology as described by Zohary (2). *P. mexicana* and *P. texana* are the only New World species represented in the genus. There are significant mountain and desert geographic barriers between the reported ranges of *P. mexicana* and *P. texana*. Although they are indigenous to the same general area, the two species have distinct morphologies. In addition to distinctly different leaf morphology (fewer leaflets with sharp points), *P. mexicana* is unique because it is briefly semideciduous and because it sheds leaves in the spring and is the only pistachio species to do so. Thus, although not separable by this cpDNA analysis, further research will be needed to resolve their statuses as distinct species.

Zohary considered *P. mexicana* and *P. texana* to be derived from *P. chinensis* on the basis of morphology (2). *P. weinmannifolia* and the *P. mexicana*:*P. texana* group are monophyletic, suggesting that the New World species probably were derived from an Asian ancestor common to *P. weinmannifolia* rather than *P. chinensis*. All phylograms also support a common ancestor for *P. weinmannifolia* and *P. lentiscus*, as suggested by Zohary. The association of *P. weinmannifolia* with the *P. mexicana* and *P. texana* groups in all of the models would not be possible using his proposal of evolution via *P. chinensis* in Asia followed by the disappearance of *P. chinensis* from Asia. Our findings indicate that both *P. mexicana* and *P. texana* should be included in section *Eu Lentiscus* with *P. lentiscus* and *P. weinmannifolia* rather than in section *Lentiscella*, which could then be eliminated.

P. integerrima, an important rootstock for commercial pistachio production, was clearly different from *P. chinensis*, which is somewhat more derived. This result differs from the classification of Zohary that suggests that *P. integerrima* is a more recently diverged subspecies of *P. chinensis* (2). However, our classification is consistent with the flowering behavior of the two species. When grown in the same location in California, *P. integerrima* is the first of the 10 species to flower and *P. chinensis* is the last to flower. There is no overlap in flowering period. If both species were present at the same location in Asia, it is unlikely that hybridization could occur. The distributions of these species are also distinct, separated by the Himalaya and Karakoram mountain ranges. *P. integerrima* occurs on the south and west sides and *P. chinensis* occurs to the east side (2). Thus, both geographic and reproductive barriers are likely to maintain speciation. That *P. chinensis*

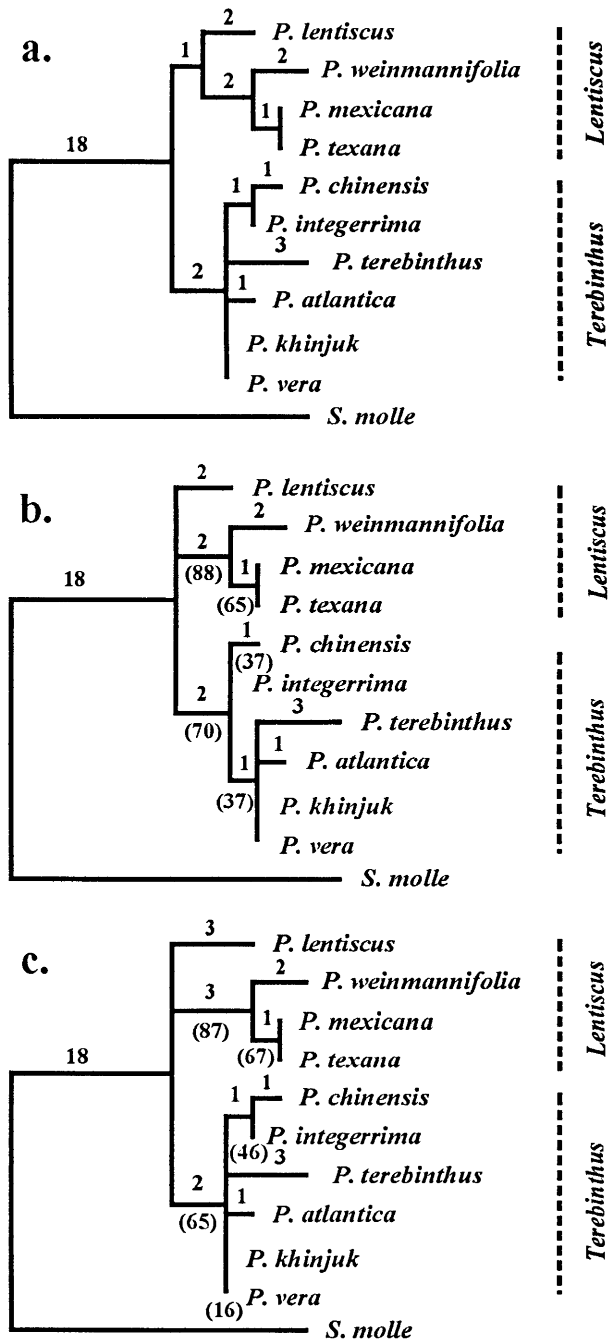

FIG. 4. Phylograms from (*a*) mixed Dollo parsimony, (*b*) unordered parsimony and 50% majority rule bootstrap consensus tree from unordered parsimony, and (*c*) 50% majority rule bootstrap consensus tree from mixed Dollo parsimony analyses of 10 *Pistacia* species and *Schinus molle* outgroup. Arm lengths represent relative extent of divergence from nodes with number of mutations/arm. Bootstrap confidence levels are shown in parentheses. Proposed sectional groupings are shown with dotted lines at *Right*.

apparently is more diverged also is consistent with the postulated center of diversity for *Pistacia* in the Eastern Mediterranean to Central Asian region. Zohary considered *P. integerrima* to be the bridge from *P. khinjuk* to *P. chinensis*, which supported the inclusion of *P. chinensis* in the section *Eu Terebinthus*. Although *P. integerrima* should be included in section *Eu Terebinthus* along with *P. chinensis*, it appears to have developed separately from *P. khinjuk* and *P. vera*.

All of the parsimony analyses place *P. atlantica* in the monophyletic group *P. atlantica*:*P. terebinthus*:*P. khinjuk*:*P. vera*, suggesting that it be placed in section *Eu Terebinthus* and that the section *Butmela* be eliminated. Both cpDNA statistical analysis and leaf and seed morphologies support the division of *Pistacia* into two sections, *Lentiscus* and *Terebinthus*. All of the species in *Lentiscus* are evergreen with a paripinnate leaflet arrangement whereas *Terebinthus* species lose their leaves in the autumn and have imparipinnate leaves. Seed size in *Lentiscus* species is much reduced compared with that of the *Terebinthus* species.

This study suggests that *P. khinjuk* and *P. vera* are primitive *Pistacia* species. This is consistent with a Central Asian center of diversity for the genus because the natural range for *P. vera* spans this region. Zohary postulated *P. khinjuk* to be directly descended from *P. vera*, a hypothesis that could not be supported or rejected with the cpDNA analysis. Zohary also considered *P. khinjuk* and *P. vera* to be the most primitive *Pistacia* species based on nine characters of primitive species: simple leaves, imparipinnate leaves, small number of leaflets per leaf, symmetrical leaflets, rounded leaflet apex, simple petiole (no wings), highly branched panicles, deciduous character, and large fruit. *P. khinjuk* and *P. vera* are the only *Pistacia* species with large edible nuts. Both have a similar somewhat unique three- leaflet imparipinnate leaf. Possibly the evolution toward a smaller seed with a hard endocarp parallels a change in reproductive strategy from distribution by ground squirrels (burying the seed, as with walnuts and oaks) to bird-mediated distribution, which would require a seed capable of passing through a bird's gut (24). Species in both *Lentiscus* and *Terebinthus*, which diverged relatively early, have evolved smaller leaves with more leaflets and smaller hard seeds although these monophyletic groups probably evolved independently. Smaller elongated leaflets with pointed shoot apices also are more efficient for water removal from the leaf surface compared with simple rounded leaves. This would be a useful adaptation as the genus moved into higher rainfall regions.

Rate of Evolutionary Change in *Pistacia*. Wolfe *et al.* (25) cited a level of $1.0–3.0 \times 10^{-9}$ number of substitutions/synonymous sites/year as a typical rate of chloroplast DNA change, which would be a $\approx 2.7–9.1 \times 10^{-10}$ total substitution rate based on the results of Bousquet *et al.* (26) and Frascaria *et al.* (27). Bousquet *et al.* (26) noted that substitution rates can vary substatially among taxa, and, in a later paper, Frascaria *et al.* (27) cited overall substitution rates of 0.71×10^{-10} substitutions/site/year for Fagaceae and 0.86×10^{-10} for Arecaceae vs. 2.75×10^{-10} for Solanaceae and 4.45×10^{-10} for Poaceae, suggesting that annual species are evolving at a rate four to six times that of trees. Zurawski *et al.* (28) and Wilson *et al.* (29) reported similar rates of 6.7×10^{-10} substitutions/site/year for Poaceae and $0.58–1.3 \times 10^{-10}$ in the Arecaceae, respectively.

Because three site mutations were observed for 2680 sites between *P. vera*, the most ancestral *Pistacia* species, and *P. terebinthus*, the most recently evolved species, a rough approximation of the age of *P. vera* is between 1.1 and 3.7 million years. This assumes no change in *P. vera* since divergence. Using the branch lengths of the mixed Dollo tree, an average of four mutations occurred between the bifurcation of monophyletic groups *Lentiscus* and *Terebinthus* and modern species. This would imply that the common ancestor of *Pistacia* is between 1.5 and 5.0 million years old. However, fossil evidence suggests that *Pistacia* evolved ≈80 million years ago so that the average mutation rate for *Pistacia* is 16 to as much as 50 times less than expected from the typical mutation rate cited in Wolfe *et al.* (25). Ogihara *et al.* (13) found substitution percentages per nucleotide for the 3.2-kb region among grass species of more than five times the level found in the present study with *Pistacia*, suggesting that *Pistacia* has evolved slowly. If Frascaria *et al.* (27) and Wilson *et al.* (29) are correct, then mutation rates for trees could be as much 6–13 times less than for annuals; however, higher mutation rates may exist for the

noncoding regions (30) sampled in this study. Ogihara *et al.* (13) note that mutation rates in noncoding cpDNA are 10 times greater than for coding regions, similar to the results for *Pistacia*, with 1.4×10^{-2} mutations/nucleotide for the 3.2-kb fragment (which includes a large noncoding region) and 10^{-3} mutations/nucleotide for the entire chloroplast genome.

Zohary notes that, based on fossil evidence, *P. lentiscus* originated ≈40 million years ago, and the genus as a whole probably developed more than 80 million years ago. Using 40 million years as a rough estimate of the divergence time between *P. lentiscus* and the *Lentiscus*:*Terebinthus* bifurcation point (1 nt change per 13 million years), the rate of cpDNA change in *Pistacia* is $\approx 4.9 \times 10^{-11}$ mutations/site/year. This is from 6–19 times less than typically reported for plant cpDNA, even though 59% of the 17 observed mutations (including DNA rearrangements) were found in the 3.2-kb hypervariable region, and this supports Zohary's contention that *Pistacia* has evolved very slowly. This mutation rate is similar (≈1.4 times less) to the rate reported for Fagaceae (27), another dicot tree family, or 6–14 times less than reported for Solanaceae and Poaceae (27, 28). However, the mutation rate for *Pistacia* includes data from a relatively variable cpDNA region so that the actual *Pistacia* mutation rate comparison vs. annual plants may be larger than the observed values suggest, perhaps similar to estimates based on cpDNA fossil evidence (6–19 times less than annuals). Li and Tanimura (31) suggest that differences in mutation rates among organisms may be more a function of generation time than DNA repair rates. *Pistacia* species have a long generation cycle, at least 10 years to first flowering, and a life-span estimated to be as much as 400 years in some cases. The average replacement cycle for pistachio is probably between 50 and 200 years in the wild, so it is not surprising that *Pistacia* has evolved much more slowly than the annual species used to derive standard mutation rate estimates.

We gratefully acknowlege the assistance of R. K. Jansen and J. D. Palmer for supplying the tobacco cpDNA library used in this study. This study was supported in part by the California Pistachio Commission, Ministry of Education of Spain, and the Agricultural Experiment Station of the University of California.

1. Wannan, B. S. & Quinn, C. J. (1991) *Bot. J. Linn. Soc.* **107,** 349–385.
2. Zohary, M. (1952) *Palestine J. Bot. Jerusalem Ser.* **5,** 187–228.
3. Kokwaro, J. O. & Gillett, J. B. (1980) *Kew Bull.* **34,** 745–760.
4. Grundwag, M. & Werker, E. (1977) *Israel J. Bot.* **25,** 152–167.
5. Dowling, T. E., Moritz, C. & Palmer, J. D. (1990) in *Molecular Systematics*, eds., Hillis D. M. & Moritz, C. (Sinauer, Sunderland, MA), pp. 250–315.
6. Palmer, J. D. (1987) *Am. Nat.* **130,** S6–S29.
7. Rajora, O. P. and Dancik, B. P. (1992) *Theor. Appl. Genet.* **84,** 280–285.
8. Waugh, R., van de Ven, W. T. G., Phillips, M. S. & Powell, W. *Pl. Syst. Evol.* **172,** 65–75.
9. Olmstead, R. G. & Palmer, J. D. (1992) *Ann. Missouri Bot. Gard.* **79,** 346–360.
10. Arnold, M. L., Buckner, C. M. & Robinson, J. J. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1398–1402.
11. Liston, A., Rieseberg, L. H. & Hanson, M. A. (1992) *Plant Syst. Evol.* **181,** 121–132.
12. Rieseberg, L. H., Hanson, M. A. & Philbrick, C. T. (1992) *Syst. Bot.* **17,** 324–336.
13. Ogihara, Y., Terachi, T. & Sasakuma, T. (1991) *Genetics* **129,** 873–884.
14. Doyle, J. J. & Doyle, J. L. (1987) *Phytochem. Bull.* **19,** 11–15.
15. Shinozaki, K., Ohme, M., Tanake, M., Wakasugi, T., Hayshida, N., *et al.* (1986) *Plant Mol. Biol. Rep.* **4,** 110–147.
16. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
17. Swofford, D. L. (1989) BIOSYS-1: A Computer Program for the Analysis of Allelic Variation in Population Genetics and Biochemical Systematics, Release 1.7 (Illinois Natural History Survey, Champaign, IL).
18. Swofford, D. L. (1990) PAUP: Phylogenetic Analysis Using Parsimony, Ver. 3.0 (Illinois Natural History Survey, Champaign, IL).
19. Rohlf, F. J. (1993) NTSYS-PC Numerical Taxonomy and Multivariate Analysis System (Exeter Software, Setauket, NY).
20. Nei, M. & Li, W. H. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 5269–5273.
21. Nei, M. (1987) *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York), pp. 64–66.
22. Sneath, P. R. & Sokal, R. R. (1973) *Numerical Taxonomy: The Principles and Practice of Numerical Classification* (Freeman, San Francisco), p. 351.
23. Swofford, D. L. & Olsen, G. J. (1990) in *Molecular Systematics*, eds. Hillis, D. M. & Moritz, C. (Sinauer, Sunderland, MA), pp. 411–501.
24. Jordano, P. (1989) *Oikos* **55,** 357–386.
25. Wolfe, K. H., Li, W. H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 9054–9058.
26. Bousquet, J., Strauss, S. H., Doersen. A. H. & Price, R. A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7844–7848.
27. Frascaria, N., Maggia, L., Michaud, M. & Bousquet, J. (1993) *Genome* **36,** 668–671.
28. Zurawski, G, Clegg, M. T. & Brown, H. D. (1984) *Genetics* **106,** 735–749.
29. Wilson, M. A., Gaut, B. & Clegg, M. T. (1990) *Mol. Biol. Evol.* **7,** 303–314.
30. Clegg, M. T. & Zurawski, G. (1992) in *Molecular Systematics of Plants*, eds. Soltis, P. S., Soltis, D. E. & Doyle, J. J. (Chapman and Hall, New York), pp. 1–13.
31. Li, W. H. & Tanimura, M. (1987) *Nature (London)* **326,** 93–96.