Internal Transcribed Spacer Sequence Phylogeny of *Crambe* L. (Brassicaceae): Molecular Data Reveal Two Old World Disjunctions

Javier Francisco-Ortega,*,[1] Javier Fuertes-Aguilar,† César Gómez-Campo,‡
Arnoldo Santos-Guerra,§ and Robert K. Jansen‖

*Department of Biological Sciences and Fairchild Tropical Garden, Florida International University, University Park, Miami, FL 33199;
†Real Jardín Botánico, Consejo Superior de Investigaciones Científicas, Plaza de Murillo 2, Madrid, 28014, Spain;
‡Departamento de Biología Vegetal, Escuela Técnica Superior de Ingenieros Agrónomos, Ciudad Universitaria, Madrid, 28040, Spain;
§Jardín Botánico de La Orotava, Calle Retama 2, Puerto de La Cruz, 38400, Tenerife, Canary Islands, Spain;
and ‖Department of Botany and Institute of Cell and Molecular Biology, University of Texas, Austin, Texas 78713

Crambe L. (Brassicaceae) is an Old World genus with a disjunct distribution among four major centers of species diversity. A phylogenetic analysis of nucleotide sequences of the internal transcribed spacers (ITS) of the nuclear ribosomal repeat was conducted with 27 species of *Crambe* and 18 related genera. Cladistic analyses using weighted and unweighted parsimony support *Crambe* as a monophyletic genus with three major lineages. The first comprises those taxa endemic to the Macaronesian archipelagos. Taxa with a predominant Mediterranean distribution form the second assemblage, and a disjunction between east Africa (*C. abyssinica*) and the Mediterranean (*C. hispanica*) occurs in this clade. The third lineage includes all Eurosiberian–Asian taxa and *C. kilimandscharica*, a species from the highlands of east Africa. A basal biogeographic split between east Africa and Eurasia is present in the third clade. The patterns of relationships in the ITS tree are concordant with known climatic events in northern Africa and southwestern Asia since the middle Miocene. The ITS trees are congruent with the current sectional classification except for a few members of sections *Crambe, Leptocrambe,* and *Orientecrambe* (*C. cordifolia, C. endentula, C. kilimandscharica,* and *C. kotschyana*). Low levels of support in the basal branches do not allow resolution of which genera of the subtribes Raphaniae or Brassicinae are sister to *Crambe.* Both subtribes appear to be highly polyphyletic in the ITS trees. © 1999 Academic Press

INTRODUCTION

Plant groups with intercontinental disjunctions are relatively common in most floristic regions of the world

(Thorne, 1972). Remarkable examples of these major disjunctions can be found between eastern Asia and eastern North America (Boufford and Spongberg, 1983) and between Africa and South America (Bremer, 1993). Other cases of congeneric disjunctions are limited to a single continent, such as to northern and southern Africa (Quézel, 1985; Coetzee, 1993) and to the Pacific coasts of North and South America (Constance *et al.,* 1963). Molecular phylogenies can document the presence of major biogeographical disjunctions and distinguish whether these phytogeographical patterns are due to morphological convergence among closely related groups or to historical circumstances, such as long distance dispersal or vicariance (e.g., Parks and Wendel, 1990; Wen and Zimmer, 1996; Sang *et al.,* 1997; Vargas *et al.,* 1997).

One flowering plant group that exhibits major geographical disjunctions is *Crambe* (Brassicaceae). This genus comprises approximately 38 Old World species (Bramwell, 1969; Santos-Guerra, 1983, 1996; Khalilov, 1991a,b) and is the second largest member of the tribe Brassiceae, surpassed only in numbers of species by *Brassica* (Gómez-Campo, 1980; Warwick and Black, 1997a). *Crambe* has disjunct distributions among four major geographical regions: Macaronesian (12 species), Mediterranean (4 species), east African (3 species), and Eurosiberian–southwest Asian (ca. 20 species) (Bramwell, 1969, 1995; Gómez-Campo, 1980; Santos-Guerra, 1983, 1996; Khalivov, 1991a,b).

The Macaronesian species of *Crambe* occur in both the Canarian and Madeiran archipelagos. This group is similar to many other insular endemics in exhibiting woodiness (Carlquist, 1965). The island taxa have been suggested to represent an ancient lineage that remained in Macaronesia subsequent to major climatic and geologic events in the Mediterranean [6.5 million years ago (Mya)] (Gómez-Campo, 1980, 1993; Aguinagalde and Gómez-Campo, 1984).

The Mediterranean group includes four species. Three

[1] To whom correspondence should be addressed. Fax: (305) 348-3094. E-mail: ortega@fiu.edu.

of these (*C. glabrata, C. kralikii,* and *C. filiformis*) are restricted to the western Mediterranean, whereas *C. hispanica* has a wider distribution in Albania, Croatia, Cyprus, Greece, Israel, Italy, Lebanon, Spain, and Syria (Post, 1932; Ball, 1964; Zohary, 1966; Meikle, 1977; Mouterde, 1986). Two species (*C. hispanica* and *C. glabrata*) are annual herbs, whereas *C. kralikii* and *C. filiformis* are suffructescent perennials.

The third group, which includes the three annual species *C. abyssinica, C. kilimandscharica,* and *C. sinuato-dentata,* is restricted to the Congo, Ethiopia, Kenya, Tanzania, and Uganda (Robyns and Boutique, 1951; Jonsell, 1982). *C. abyssinica* is cultivated for its erucic acid-rich oils and high potential as food for livestock (Nieschlag and Wolff, 1971; Lessman and Meier, 1972; Van Dyne and Blase, 1989; Anderson *et al.,* 1993; Lazzeri *et al.,* 1994; Mulder and Mastebroek, 1996).

The remaining species occur from western China to eastern Europe, Bulgaria, and Turkey, mainly in the Caucasus and southwestern republics of the former Soviet Union (Kitamura, 1960; Ball, 1964; Chernyakovskaya, 1970; Czerepanov, 1995; Jalas *et al.,* 1996). Furthermore, one species (*C. maritima*) is found in coastal zones of the Black, Baltic, and North Seas. All species of this group are herbaceous perennials with swollen roots, which in some cases (i.e., *C. edentula, C. gordjaginii, C. schugnana*) are woody (Chernyakovskaya, 1970).

Crambe is placed within the subtribe Raphaninae in which it forms a distinct genus with unique biogeographical and morphological features (reviewed by Warwick and Black, 1997b). The genus is so distinctive within the Brassiceae that Gómez-Campo (1980) suggested that it could be recognized as a monotypic subtribe. The most recent infrageneric classification of *Crambe* recognizes seven sections (Kahlilov, 1991a,b). Two of these (*Dendrocrambe* and *Rhipocrambe*) include only the Macaronesian taxa. The seven species from the Mediterranean basin and east Africa are placed in sect. *Leptocrambe.* The remaining four sections (*Crambe, Astrocrambe, Flavocrambe,* and *Orientecrambe*) include species with an Eurasian distribution. Most European species are placed within sect. *Crambe* while most Asian taxa are in sects. *Astrocrambe, Flavocrambe,* and *Orientecrambe.*

Palynological studies provide some support for this sectional classification because unique pollen ornamentation is found in sects. *Astrocrambe, Crambe–Orientecrambe, Dendrocrambe, Flavocrambe,* and *Leptocrambe* (Khalilov and Archangelsky, 1991). However, differences in pollen morphology are not present between sects. *Crambe* and *Orientecrambe.*

The basic chromosome number of *Crambe* is x = 15 [reviewed by Gómez-Campo and Hinata (1980); Magulaev (1986); Fig. 1]. There are reports of species with n = 30 (*C. hispanica*), n = 45 (*C. abyssinica*), n = 60 (*C. cordifolia, C. grandifolia*), and n = 75 (*C. gordjaginii*).

A recent chloroplast DNA (cpDNA) restriction site analysis (Warwick and Black, 1997b) of the subtribes Raphaninae and Cakilinae included 15 species from three of the seven sections of *Crambe* recognized by Khalilov. The two major conclusions of this study were that: [1] three lineages corresponding to the sections recognized in an earlier classification by Schulz (1919) (i.e., sects. *Crambe, Dendrocrambe,* and *Leptocrambe*) are supported within *Crambe* and [2] the Macaronesian taxa are sister to the other members of the genus. However, limited sampling and poor support (30% bootstrap value) for the monophyly of *Crambe* weaken these conclusions.

In this paper, we present a phylogenetic analysis of nucleotide sequences of the Internal Transcribed Spacers (ITS) of the nuclear ribosomal DNA of *Crambe* and related genera. The three objectives are to: [1] test the monophyly of *Crambe* and its placement within the subtribe Raphaninae, [2] compare the groups in the ITS tree with the current sectional classification and with those groups identified in the cpDNA restriction site trees, and [3] relate the ITS phylogeny to the major climatic and geologic events since the middle Miocene (20 Mya) when the African and Eurasian plates collided.

MATERIALS AND METHODS

Plant Material

Our sampling included 27 species of *Crambe,* with at least 1 species from six of the seven sections of Khalilov (1991a,b) (Table 1). Section *Rhipocrambe* was not included because ITS trees for all 12 Macaronesian species indicate that this section should be merged with sect. *Dendrocrambe* (Francisco-Ortega *et al.,* unpubl.).

Twenty-four of the 27 currently recognized continental species were examined (Table 1). These species represent the entire geographic distribution of the genus on the continent. The ingroup also included the five genera (*Calepina* Adans., *Crambella* Maire, *Hemicrambe* Webb, *Kremeriella* Maire, and *Muricaria* Desv.) which have been proposed as closely related to *Crambe* (Clemente and Hernández-Bermejo, 1978a,b, 1980a,b; Gómez-Campo and Tortosa, 1974; Gómez-Campo, 1980). Three additional genera considered part of the core of the subtribe Raphaninae (*Guiraoa* Coss., *Raphanus* L., and *Rapistrum* Crantz) (Gómez-Campo, 1980) were also included. Furthermore, representatives of the core of the subtribes Brassicinae (*Brassica* L., *Eruca* Mill., *Hirschfeldia* Moench, *Sinapidendron* Lowe, and *Sinapis* L.) and Zillinae (*Physorrhinchus* Hook.) were sampled.

Four genera of the tribes (as defined by Al-Shehbaz, 1984, 1988) Alysseae (*Camelina* Crantz), Anchonieae (*Erysimum* L.), and Sisymbreae (*Cardaminopsis* (C. A. Mey.) Hayek, and Sisymbrium L.) were selected as outgroups.

DNA Isolation, PCR Amplification, and Sequencing

Genomic DNAs were isolated from both fresh material and herbarium specimens. Fresh material was obtained from plants cultivated in the greenhouses at The University of Texas at Austin and the Botanic Garden of La Orotava from seeds provided by several germplasm institutes (Table 1). The CTAB technique of Doyle and Doyle (1987) was used for fresh material. Isolations from herbarium specimens were performed as described in Loockerman and Jansen (1996). Primers ITS7A (Panero and Plovanich-Jones, unpubl.) and ITS4 (White *et al.,* 1990) were used for double-stranded DNA amplification of the ITS region. Amplification and purification of the PCR product was conducted following Kim and Jansen (1994).

Direct sequencing of double-stranded DNA was performed using Sequenase version 2.0 (United States Biochemical Corp.) and the snap-chill method (Winship, 1989). Two forward (ITS3 and ITS5) and two reverse (ITS2 and ITS4) primers (White *et al.,* 1990; Downie and Katz-Downie, 1996) were utilized to sequence both strands.

Sequence Analysis

Boundaries of the ITS1 and ITS2 regions were determined by comparisons with the published sequences of *Sinapis alba* (Rathgeberg and Capesius, 1989). CLUSTAL W (Higgins *et al.,* 1991) was used for multiple sequence alignment with minor manual corrections. Sequencing results were ambiguous and difficult to read due to compressions for the six nucleotides situated between positions 111 and 116 (Appendix 1). Therefore this region was excluded from all data analyses.

Transition/transversion ratios were calculated from the most parsimonious trees using MacClade (Maddison and Maddison, 1992, version 3.04). The software package MEGA (Kumar *et al.,* 1993) was used to estimate sequence divergence based on Kimura's two-parameter model (Kimura, 1980) and the G + C content of both ITS1 and ITS2.

Relative rate tests (Sarich and Wilson, 1973; Wu and Li, 1985; Li and Tinamura, 1987; Li, 1997) were used to test for a molecular clock in *Crambe* based on nucleotide divergence values estimated from Kimura's two-parameter model. Standard errors of rate differences between lineages were computed using the Li and Tanimura (1987) and Kimura (1980) methods. Eight representative species of the major clades of *Crambe* (i.e., *C. edentula, C. gordjaginii, C. hispanica, C. kilimandscharica, C. maritima, C. orientalis, C. schugnana, C. sventenii*) were used for tests with *Raphanus sativus* serving as the reference taxon.

Phylogenetic Analysis

Parsimony analyses were conducted using PAUP 3.1.1 (Swofford, 1993) with ACCTRAN, MULPARS, and TBR options. All searches involved 100 random entries to search for multiple islands of equally parsimonious trees (Maddison, 1991). Gaps were treated in four different ways: (1) missing data, (2) eliminated, (3) considered as missing data and each coded as a binary character (absent or present), or (4) eliminated from the matrix and each coded as a binary character (absent or present). Transversions were weighted over transitions by a 1.4:1 ratio using the USERTYPE STEPMATRIX command of PAUP. A weight of 1.4 was selected because this was the transition/transversion ratio obtained from one of the most parsimonious trees in the initial unweighted analysis.

The amount of support for monophyletic groups was evaluated by 100 bootstrap replicates (Felsenstein, 1985). The ACCTRAN, MULPARS, and TBR options were used in a heuristic search with simple addition sequence of taxa. Limitations of computer memory required constraining the maximum number of trees to 1000 and 500 in the unweighted and weighted analyses, respectively.

Consistency Index (CI) (Kluge and Farris, 1969), Retention Index (RI) (Farris, 1989), and the $g1$ statistic (Hillis and Huelsenbeck, 1992) were also computed. The latter was obtained after computing the tree-length distribution of 100,000 random parsimony trees by means of the RANDOM TREES command.

RESULTS

Length Variation, Base Composition, and Nucleotide Divergence

Although some species of *Crambe* are polyploids (i.e., *C. cordifolia,* n = 60; *C. gordjaginii,* n = 75), only one ITS type was detected in each individual of all 45 species examined. The length of the aligned ITS region, excluding the 5.8s gene, was 488 bp (Table 2) (ITS1 was 292 bp and ITS2 was 196 bp). Two of the outgroup species had the shortest (*Sisymbrium,* 440 bp) and longest (*Erysimum,* 459 bp) ITS sequences. *Crambe* varied between 448 bp (*C. santosii*) and 456 bp (*C. amabilis*) (Appendix 1). The number of indels after alignment was 41, and most of these were 1, 2, and 3 bp. Two deletions of 10 and 16 bp were detected in the ITS1 of *Brassica oleracea* and *Sisymbrium altissimum,* respectively. The G + C content varied from 50.2% (*Crambe hispanica* and *Muricaria prostrata*) to 57.9% (*Erysimum wittmanii*) with a mean value of 52.3%. The *Crambe* species with the highest G + C content was *C. aspera* (52.3%).

Nucleotide divergence values among the 45 species examined ranged between 0 and 28.1% (Table 2). The highest divergence within *Crambe* (10.64%) was between the Canarian endemic *C. sventenii* and the east African endemic *C. abyssinica.* Several divergence values within *Crambe* were zero, and most of these involved comparisons among the 9 species of the "*orien-*

TABLE 1

Taxa Included in the ITS Study and Their Provenance

Taxon	Taxon abbreviation	Voucher and germplasm origin[a]	Geographical origin	GenBank Accession No. (ITS1/ITS2)
Ingroup				
Brassica oleracea L.	BROLE	Franc.-Ort., TEX	Red cabbage from market, Austin	AF03994/AF040037
Crambe L. sect. *Crambe*				
C. aspera M. Bieb.	CRASP	FOM-1994/594, ORT	Ukraine	AF039994/AF040026
C. cordifolia Steven	CRCOR	ETSIA-0511/68, TEX	Unknown	AF039966/AF040009
C. gibberosa Rupr.	CRGIB	Korotyaev, 23-04-1990, LE	Dagestan	AF039974/AF040017
C. kotschyana Boiss.	CRKOT	GAT-CRA12/93, TEX	Russia	AF039967/AF040010
C. maritima L.	CRMAR	ETSIA-0510/67, TEX	France	AF039970/AF040013
C. pinnatifida R. Br.	CRPIN	LE-2211, TEX	Russia	AF039965/AF040008
C. pontica Rupr.	CRPON	KRI-1993/1995/902, TEX	Ukraine	AF039973/AF040016
C. tataria Sebeók	CRTAT	ETSIA-0578/68, TEX	Hungary	AF039964/AF040007
Crambe sect. *Astrocrambe* I. Khalilov				
C. schugnana Korsh.	CRSCH	Sultanov, 28-08-1970, LE	Tadzhikistan	AF039980/AF040023
Crambe sect. *Dendrocrambe* DC.				
C. gomerae H. Christ	CRGOM[f]	Franc.-Ort & Fdez.-Galván, 10-07-1995, TEX	Canary Islands	AF039960/AF040003
C. santosii Bramwell	CRSAN	ETSIA-4054/76, TEX	Canary Islands	AF039959/AF040002
C. sventenii Bramwell & Sunding	CRSVE	ETSIA-5304/79, TEX	Canary Islands	AF039976/AF040019
Crambe sect. *Flavocrambe* I. Khalilov				
C. gordjaginii Spryg. & M. Pop.	CRGOR	Bochantsev & Mikhailova, 19-08-1978, LE	Turkmenistan	AF039981/AF040024
Crambe sect. *Leptocrambe* DC.				
C. abyssinica Hochst.	CRABY[f]	PI-392326, TEX	Turkey	AF039958/AF040001
		PI-384522, TEX	Ethiopia	
		PI-279346, TEX	Ethiopia	
		PI-372925, TEX	Spain	
		ETSIA-1397/67, TEX	Spain	
C. filiformis Jacq.	CRFIL	ETSIA-1182/67, TEX	Spain	AF039963/AF040006
C. glabrata DC.	CRGLA[f]	PI-388857, TEX	Portugal	AF039962/AF040005
		PI-388853, TEX	Portugal	
		PI-388783, TEX	Morocco	
		PI-388786, TEX	Morocco	
C. hispanica L.	CRHIS[f]	ETSIA-4004/75, TEX	Israel	AF036209/AF036210
		PI-378590, TEX	Spain	
		PI-337996, TEX	Israel	
		PI-388835, TEX	Italy	
		PI-388822, TEX	Cyprus	
C. kilimandscharica O. E. Schulz	CRKIL[f]	R. Gutzwiller, Num. 661, K	Zaire	AF039975/AF040018
		C. Mutiso s.n., ORT	Kenya	
C. kralikii Coss.	CRKRA	ETSIA-1104/67, TEX	Morocco	AF039961/AF040004
Crambe sect. *Orientecrambe* I. Khalilov				
C. aculeolata (N. Busch) Czerniak.	CRACU	Khalilov, 24-06-1987, LE	Nakhichevan	AF039978/AF040021
C. amabilis Butk. & Majlun	CRAMA	SO-22, TEX		AF039971/AF040014
C. edentula Korsh.	CREDE	Averianov, 9-05-1979, LE	Turkmenia	AF039968/AF040011
C. grossheimii I. Khalilov	CRGRO	Khalilov, 22-06-1987, LE	Nakhichevan	AF039982/AF040025
C. juncea M. Bieb.	CRJUN	PI-325274, TEX	Russia	AF039979/AF040022
C. koktebelica (Junge) N. Busch	CRKOK	ETSIA-0189/67, TEX	Unknown	AF039972/AF040015
C. orientalis L.	CRORI	ETSIA-6243/83, TEX	Turkey	AF039969/AF040012
C. persica Boiss.	CRPER	M. Jafari, 07-1996, TEX	Iran	AF039977/AF040020
Crambella teretifolia (Batt.) Maire	CMTER	ETSIA-1971/71, TEX	Morocco	AF039986/AF040029
Eruca sativa Mill.[b]	ERSAT			AF039996/AF040039
Guiraoa arvensis Coss.	GUARV	ETSIA-1550/68, TEX	Spain	AF039987/AF040030
Hemicrambe fruticulosa Webb	HEFRU	ETSIA-2232/73, TEX	Morocco	AF039984/AF040027
Hirschfeldia incana (L.) Lagr.-Foss.	HIINC	Santos-Guerra & Franc.-Ort. 2-02-1995, ORT	Canary Islands	AF039989/AF040032
Kremeriella cordylocarpus (Coss. & Durieu) Maire	KRCOR	ETSIA-1142/67, TEX	Morocco	AF039985/AF040028
Muricaria prostrata (Desf.) Desv.	MUPOS	ETSIA-1855/70, TEX	Algeria	AF039992/AF040035
Physorrhynchus chamaerapistrum Boiss.	PHCHA	ETSIA-3749/75, TEX	Iran	AF039990/AF040033
Raphanus sativus L.	RASAT	ETSIA-1279/67, TEX	Spain	AF039988/AF040031

TABLE 1—*Continued*

Taxon	Taxon abbreviation	Voucher and germplasm origin[a]	Geographical origin	GenBank Accession No. (ITS1/ITS2)
Rapistrum rugosum (L.) All.	RPRUG	J. Franc.-Ort., 5-05-1994, TEX	Texas	AF039991/AF040034
Sinapidendron frutescens (Aiton) Lowe	SNFRU	ETSIA-4971/79, TEX	Madeira	AF039993/AF040036
Sinapis alba L.[c]	SIALB			X15915[e]
Outgroup				
Calepina irregularis (Asson) Thell.	CLIRR	ETSIA-2158/72, TEX	Spain	AF039995/AF040038
Camelina microcarpa Andrz.[d]	CMMIC	O'Kane & Dihoru 3596, MO	Romania	AF039999/AF040042
Cardaminopsis arenosa (L.) Hayek[d]	CAARE	O'Kane & Dihouru 3600, MO	Romania	AF039998/AF040041
Erysimum wittmanii Zawadzki[d]	ERWIT	O'Kane & Dihoru 3596, MO	Romania	AF040000/AF040043
Sisymbrium altissimum L.[b]	SMALT	R. Price 1117, GA	California	AF039997/AF040041

[a] Institute codes follow Holmgren *et al.* (1990) except the following germplasm seed banks: ETSIA = Escuela Tecnica Superior de Ingenieros Agronomos, Madrid, Spain; FOM = Hortus Botanicus Fominianus, Ukraine; KRI = Botanical Garden of Krivoy, Ukraine; PI = North West Plant Introduction Station of the USDA, Ames, Iowa.

[b] Sequence provided by Robert Price (University of Georgia).

[c] Sequence published by Rathgeber and Capesius (1989).

[d] Sequence provided by Steve O'Kane (University of Northern Iowa).

[e] DNA accession number deposited at the GenBank/EMBL.

[f] Multiple accessions were sequenced but only one was submitted to GenBank and included in the phylogenetic analyses because sequences of all accessions were identical.

TABLE 2

Main Features of the ITS Region of the 45 Taxa Included in This Study

	ITS1	ITS2	ITS1 and ITS2
Length range (bp)	257–274	179–187	440–459
Aligned length (bp)	292	196	488
G + C content range (%)	47.8–58.0	52.1–57.9	50.2–57.9
Sequence divergence (%)[a,b]			
Complete deletion of gaps and missing data	0.00–23.9	0.00–22.8	0.00–22.5
Pairwise deletion of gaps and missing data	0.00–33.9	0.00–24.4	0.00–28.1
Number of indels[b]			
1 bp	14	11	25
2 bp	3	4	7
3 bp	2	2	4
4 bp	1	0	1
5 bp	2	0	2
10 bp	1	0	1
16 bp	1	0	1
Number of variable sites[b,c]	154	89	243
Number of informative changes[b,c]	109	61	170
Number of constant sites[b,c]	132	107	239
Number of autapomorphic sites[b,c]	45	28	73
Transitions (unambiguous)[b,c]	188	102	290
Transversions (unambiguous)[b,c]	136	69	205
Transitions/transversions[b,c]	1.38	1.47	1.41
Skewness of tree-length distribution (g_1)[b,c,d]	−0.900	−0.421	−0.799

[a] Nucleotide divergence based on Kimura's two-parameter model.

[b] Calculations made after excluding the six uncertain sites.

[c] Calculations made after considering gaps as missing data.

[d] $g1$ value calculated for 100,000 random trees.

talis" group and the 6 species of the "*maritima*" group (see discussion on these two groups). The two least divergent taxa from *Crambe* were *Guiraoa arvensis* (7.21%, compared with *C. koktebelica*) and *Raphanus sativus* (7.52%, compared with *C. koktebelica*).

Phylogenetic Analysis

A total of 243 variable sites was detected in the ITS region (Table 2). Approximately 70% of the sites (170 sites) were potentially phylogenetically informative, with most of these (109) in ITS1. The $g1$ value was −0.799 (100,000 random trees), suggesting that the ITS sequences have a substantial amount of phylogenetic signal (Hillis and Huelsenbeck, 1992). None of the initial weighted or unweighted analyses (trees not shown) included *Calepina* within the ingroup; therefore it was treated as an outgroup in subsequent phylogenetic analyses.

The number of equally parsimonious trees from the four unweighted analyses ranged from 468 (gaps removed from the data matrix) to 4674 (gaps removed from the data matrix and coded as binary characters) (Figs. 1 and 2). The strict consensus trees from these four analyses yielded *Crambe* as a monophyletic group supported by bootstrap values between 63 and 73% (Figs. 1 and 2).

All four unweighted analyses identified three major monophyletic groups within *Crambe*. The first, which was strongly supported by an average bootstrap value of 99.75%, comprised the Macaronesian endemic section *Dendrocrambe* (Macaronesian clade). The second clade included all Mediterranean species and the east African endemic *C. abyssinica* (Mediterranean clade). The monophyly of this clade was also strongly supported by an average bootstrap value of 91.25%. All
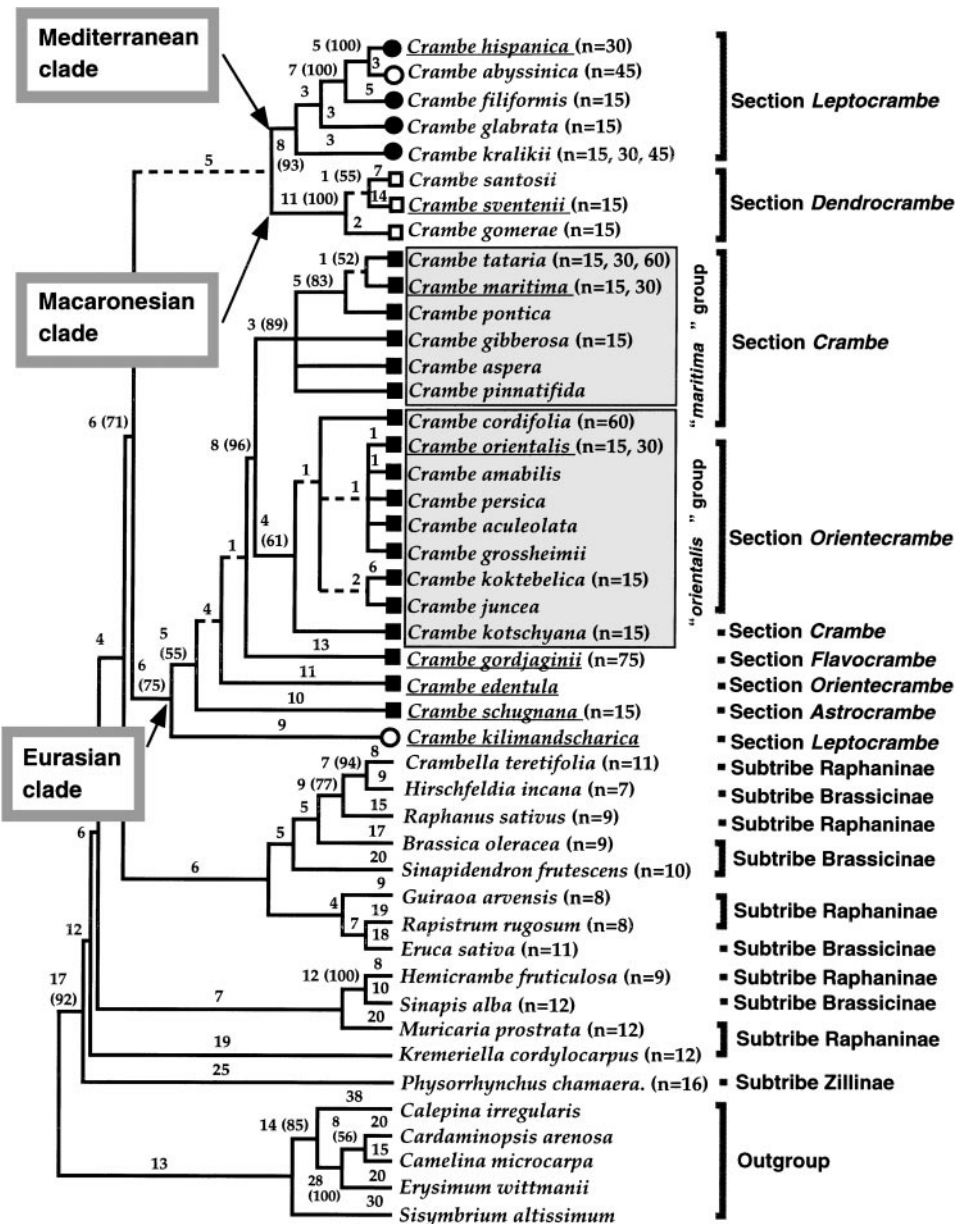
FIG. 1. One of the 1491 shortest trees from the unweighted parsimony analysis that treated indels as missing data (637 steps; CI = 0.478, without autapomorphies; RI = 0.681). Branches which collapse in the strict consensus tree are indicated by dashed lines. Number of changes are indicated along each branch. Bootstrap values higher than 50% are indicated in parenthesis. Distribution of *Crambe* species are: closed circles, Mediterranean; open circles, east African; open squares, Macaronesian; closed squares, Eurasian. Haploid chromosome numbers are indicated in parentheses when known. Species examined to test the molecular clock hypothesis are underlined.

Eurosiberian–Asian taxa together with the east African endemic *C. kilimandscharica* were in the third clade (Eurasian clade) which had an average bootstrap support of 73%. *C. kilimandscharica* was sister to the Eurosiberian–Asian taxa, which was supported by an average bootstrap value of 59%. These three major clades formed an unresolved trichotomy in two of the analyses. However, when gaps were excluded (Figs. 2a and 2b) the Mediterranean clade was sister to the rest of the genus.

Five major lineages were identified in the Eurasian clade. Four of these included only 1 species (*C. kilimandscharica, C. gordjaginii, C. edentula,* and *C. schugnana*). The fifth lineage comprised the remaining 15 species of the Eurasian clade, and it was strongly supported by an average bootstrap value of 95.25%. Two major groups could be distinguished within this lineage. The first ("*maritima*") was strongly supported by an average bootstrap value of 82.75%. This group included 6 species of sect. *Crambe* (*C. aspera, C. gibber-*
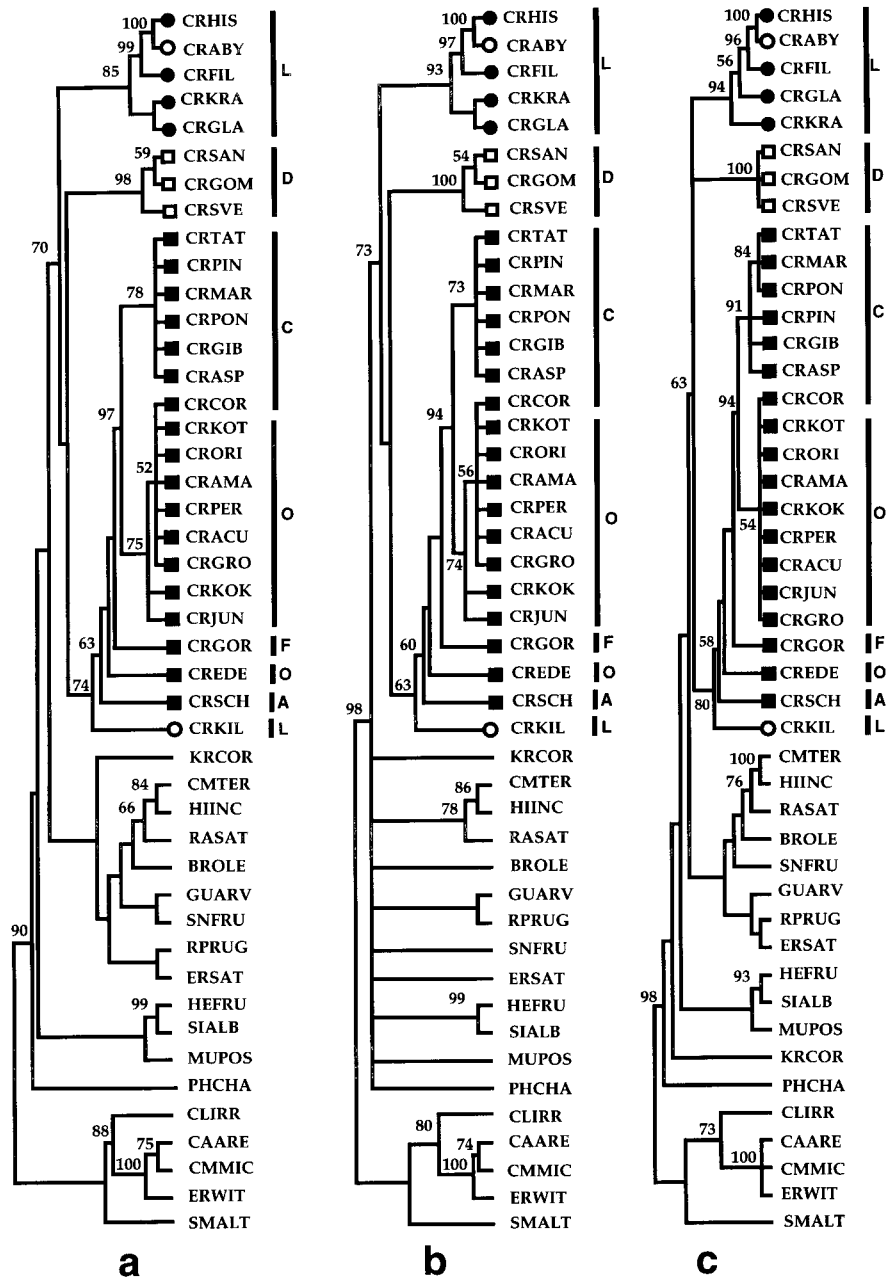
FIG. 2. Three strict consensus trees obtained after unweighted parsimony. (a) Indels removed from the analysis (468 most parsimonious trees, 566 steps each; CI = 0.481, without autapomorphies; RI = 0.682). (b) Indels removed from the analysis and coded as binary characters (4674 most parsimonious trees, 619 steps each; CI = 0.574, without autapomorphies; RI = 0.686). (c) Indels considered as missing data and coded as binary characters (2982 most parsimonious trees, 699 steps each; CI = 0.481, without autapomorphies; RI = 0.680). Bootstrap values higher than 50% are indicated along branches. Taxon abbreviations are given in Table 1. Distributions of *Crambe* species are: closed circles, Mediterranean; open circles, east African; open squares, Macaronesian; closed squares, Eurasian. Sectional abbreviations are: A = *Astrocrambe;* C = *Crambe;* D = *Dendrocrambe;* F = *Flavocrambe;* L = *Leptocrambe;* and O = *Orientecrambe.*

osa, C. maritima, C. pinnatifida, C. pontica, C. tataria). The second group ("*orientalis*") was supported by an average bootstrap value of 66%. It included species of both sects. *Crambe* (*C. cordifolia* and *C. kotschyana*) and *Orientecrambe* (*C. aculeolata, C. amabilis, C. cordifolia, C. grosheimii, C. koktebelica, C. juncea, C. orientalis, C. persica).*

No single genus consistently appeared sister to *Crambe.* Three of the four strict consensus trees (Figs. 1, 2a, and 2c) identified an assemblage of eight genera (*Brassica, Crambella, Eruca, Guiraoa, Hirschfeldia, Raphanus, Rapistrum, Sinapidendron*) as the sister group of *Crambe.* However, bootstrap support for this relationship was less than 50%. Although the main goal

of this paper was not to examine phylogenetic relationships of the subtribes Raphaninae and Brassicinae, the ITS tree does not support the monophyly of either subtribe. However, there were two generic groups, *Crambella–Hirschfeldia–Raphanus* and *Hemicrambe–Sinapis,* which were strongly supported in all analyses.

Weighted and unweighted parsimony produced identical strict consensus trees when gaps were not coded as binary characters. Differences between weighted and unweighted trees were detected when gaps were coded as binary characters. These differences, however, were minor and did not change dramatically the main results of the unweighted parsimony (Fig. 3).

Relative Rate Tests

Relative rate tests showed that only 1 of the 28 pairwise comparisons (*C. gordjaginii* vs *C. orientalis*) was significantly different at the 5% level (Table 3). This indicates that there is very little rate heterogeneity between the different lineages. Therefore, a molecular clock for ITS sequences cannot be rejected in most comparisons in *Crambe.*

DISCUSSION

Monophyly and Origin of Crambe

Brassica, Crambe, Diplotaxis, and *Erucastrum* are the only genera of the Brassiceae with more than 10 species. Among these genera, *Crambe* is the only genus that has been shown to be monophyletic using molecular data (Warwick *et al.,* 1992; Warwick and Black, 1993, 1997a,b). Chloroplast DNA restriction site data provided only weak support (30% bootstrap value) for the monophyly of *Crambe* (Warwick and Black, 1997b). Our analysis of ITS variation provides stronger support for the monophyly of the genus (bootstrap values ranged between 63 and 71%).

Both cpDNA restriction site and morphological data suggest that *Crambe* may represent a distinct subtribe within the Brassicaceae (Gómez-Campo, 1980; Warwick and Black, 1997b). The ITS data do not resolve this issue because relationships among genera of the Brassiceae are weakly supported in most cases.

No single genus is sister to *Crambe* in the ITS trees. Previous suggestions of a close phylogenetic relationship between *Crambe* and *Calepina* (Clemente and Hernández-Bermejo, 1978a,b, 1980a,b) or *Hemicrambe* (Gómez-Campo and Tortosa, 1974) are not supported by the ITS data. Our phylogeny agrees with other hypotheses (Gómez-Campo, 1980; Warwick and Black, 1997c), which suggested that *Calepina* should not be considered part of the Brassiceae.

Sectional Classification

The main incongruence between the ITS phylogeny and the current sectional classification concerns *C.*

kilimandscharica. This annual species from east Africa has been traditionally considered part of sect. *Leptocrambe.* In addition, this species has small, rugose fruits that are virtually identical to most of the Macaronesian species. *C. kilimandscharica,* however, is the basal member of a clade which includes the four Eurosiberian–Asian sections of Khalilov (1991a,b). This placement may indicate that the morphological similarities between *C. kilimandscharica* and sect. *Leptocrambe* may be homologous and plesiomorphic for the Eurasian clade.

Another disagreement between the ITS trees and the current sectional classification concerns the monophyly of the four Eurosiberian–Asian sections (Khalilov 1991a,b). These sections are clearly not monophyletic in the ITS tree (Figs. 1–3). However, the ITS phylogeny is congruent with the earlier classification of Schulz (1919), which recognized one section for all taxa of the Eurasian clade with the exception of *C. kilimandscharica.* Within the Eurasian clade, the ITS trees identify two major lineages, here designated "*maritima*" and "*orientalis*". These two groups are also geographically distinct. All species in the "*maritima*" group have European distributions and do not occur east of the Caucasus region. In contrast, species in the "*orientalis*" group have a predominant Asian distribution.

The major disagreement between all previous treatments of the Eurasian taxa and the ITS trees concerns *C. edentula.* This species has been suggested to be closely related to *C. orientalis* by Khalilov (1991a,b). However, the ITS phylogenies do not place it in the "*orientalis*" group.

The ITS trees and previous taxonomic treatments (Candolle, 1821, 1824; Schulz, 1919; Gómez-Campo, 1980; Khalilov, 1991a,b) concur that the Macaronesian taxa are monophyletic. They also indicate a close phylogenetic relationship among the Mediterranean species (including *C. abyssinica*). The sister relationship of *C. hispanica* and *C. abyssinica* has strong support, confirming previous hypotheses of the taxonomic proximity of these species (White, 1975; Jonsell, 1982).

Comparison of Nuclear and Plastid Phylogenies

Comparisons between our ITS phylogeny and the cpDNA restriction site tree of Warwick and Black (1997b) are complicated by the limited sampling in the cpDNA study. For example, *C. kilimandscharica, C. edentula,* or representatives of sects. *Astrocrambe* and *Flavocrambe* were not included in the cpDNA comparisons. Both DNA phylogenies agree in strongly supporting three major clades within the genus. The main disagreement between the ITS and cpDNA trees concerns the relationships among these three groups.
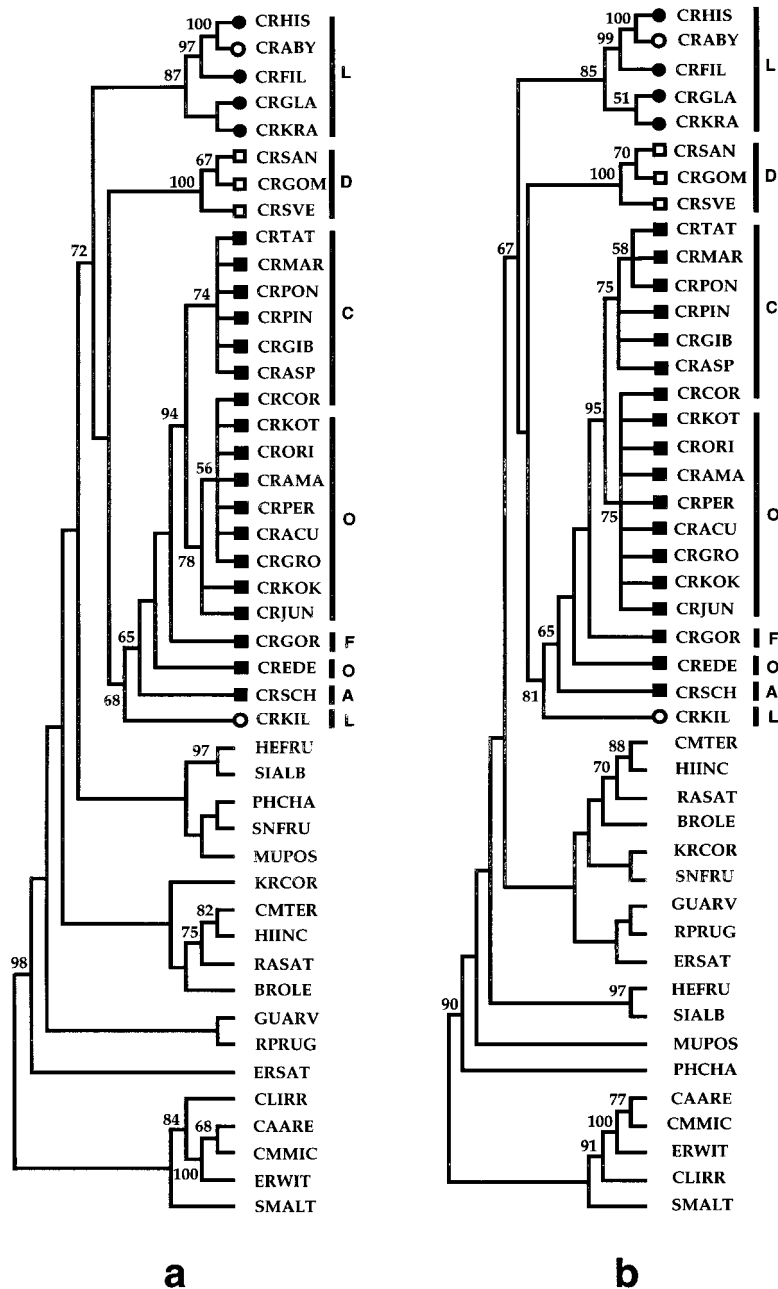
FIG. 3. Two strict consensus trees obtained after weighted parsimony (transversions were weighted over transitions by a 1.4:1 factor). (a) Indels removed from the analysis and coded as binary characters (468 most parsimonious trees, 7074 steps each; CI = 0.57; RI = 0.69). (b) Indels considered as missing data and coded as binary characters (99 most parsimonious trees, 8024 steps each; CI = 0.56; RI = 0.68). Bootstrap values higher than 50% are indicated along branches. Taxon abbreviations are given in Table 1. Distributions of *Crambe* species are: closed circles, Mediterranean; open circles, east African; open squares, Macaronesian; closed squares, Eurasian. Sectional abbreviations are: A = *Astrocrambe;* C = *Crambe;* D = *Dendrocrambe;* F = *Flavocrambe;* L = *Leptocrambe;* and O = *Orientecrambe.*

The ITS phylogeny did not resolve these relationships, whereas the cpDNA tree strongly supported (100% bootstrap value) a sister relationship between the Eurasian and the Mediterranean clades. The Macaronesian taxa were sister to this clade. However, support for this relationship was weak (30% bootstrap value).

One conflict between the plastid and the nuclear phylogenies concerns relationships within the Eurasian clade. *C. tataria* is sister to *C. orientalis* in the cpDNA tree, whereas this species is a member of the "*maritima*" group in the ITS phylogeny. Furthermore, the three species that appear closely related to *C. maritima* in the cpDNA tree (*C. cordifolia, C. kotschyana, C. koktebelica*) are part of the "*orientalis*" group in the

TABLE 3

Evaluation of Molecular Clock Hypothesis between Eight Selected Species of *Crambe* (Underlined in Fig. 1)

	CREDE[a]	CRGOR[a]	CRHIS[b]	CRKIL[b]	CRMAR[a]	CRORI[a]	CRSCH[a]	CRSVE[c]
CREDE		0.019 (0.01)	0.011 (0.01)	0.007 (0.01)	0.013 (0.01)	0.014 (0.01)	0.003 (0.01)	0.016 (0.01)
CRGOR	ND		0.007 (0.02)	0.025 (0.02)	0.004 (0.01)	0.033 (0.01)	0.022 (0.02)	0.002 (0.01)
CRHIS	ND	ND		0.018 (0.01)	0.002 (0.01)	0.026 (0.02)	0.015 (0.01)	0.005 (0.02)
CRKIL	ND	ND	ND		0.020 (0.01)	0.007 (0.01)	0.003 (0.01)	0.023 (0.02)
CRMAR	ND	ND	ND	ND		0.027 (0.02)	0.016 (0.02)	0.003 (0.02)
CRORI	ND	**	ND	ND	ND		0.011 (0.01)	0.031 (0.02)
CRSCH	ND	ND	ND	ND	ND	ND		
CRSVE	ND	ND	ND	ND	ND	ND	ND	

Note. Relative rate test values and variances (parentheses) are shown above diagonal. Significance levels of evolutionary rate differences are shown below the diagonal; not significant differences (ND), significant at 5% level (**). Taxon abbreviations are in Table 1.

[a] Perennial herbs.
[b] Annual herbs.
[c] Perennial shrubs.

ITS tree. The "*maritima*" and "*orientalis*" groups in the ITS tree clearly conform with geographical and morphological data. Thus, this incongruence may be due to hybridization, a phenomenon that has been implicated in many plant groups (Rieseberg and Soltis, 1991; Rieseberg *et al.,* 1996). Support for this hypothesis comes from the polyploid chromosome numbers of some Eurosiberian–Asian species and the sympatric distribution of these two groups in southern Russia, Ukraine, and the Caucasus.

Biogeography

There is a correlation between the ITS phylogeny and the geographic distribution of *Crambe.* One of the three major clades is restricted to Macaronesia, whereas the other two are distributed mainly in the Mediterranean and the rest of Eurasia. In addition, the "*maritima*" and "*orientalis*" groups are confined primarily to Europe and Asia, respectively.

Previous biogeographic hypotheses in *Crambe* suggested that the Macaronesian taxa represent remnants of a Tertiary flora that existed in the Mediterranean basin (Gómez-Campo, 1980; Warwick and Black, 1997b). Two lines of evidence were used to support this hypothesis: (1) the Macaronesian taxa are woody and (2) there is a basal split between the Macaronesian taxa and the two other lineages in the cpDNA phylogeny, although this relationship is weakly supported by a 30% bootstrap value. Low levels of clade support in the ITS trees does not enable resolution of this question. However, trees from five of the eight parsimony analyses do not agree with the cpDNA phylogeny because they place the Macaronesian species as sister to the Eurasian clade. This result suggests a biogeographical connection between Macaronesia, east Africa (*C. kilimandscharica*), and Eurasia.

There are several examples of biogeographical disjunctions between east Africa and Macaronesia–northwestern Africa [e.g., *Aeonium* Webb & Berth. (Crassulaceae), *Campylanthus* Roth (Scrophulariaceae),

Canarina L. (Campanulaceae), *Erucastrum* C. Presl. (Brassicaceae), *Hemicrambe fruticulosa, Euphorbia balsamifera* Aiton (Euphorbiaceae), and *Dracaena* L. (Dracaenaceae)]. These phytogeographical links led Bramwell (1985) to suggest that the Macaronesian taxa of *Crambe* are phylogenetically related to *C. kilimandscharica.* A connection between Macaronesian and east Africa has been confirmed by a phylogenetic analysis of cpDNA restriction site data of the Canarian endemic *Solanum vespertilio* Aiton (Solanaceae) (Olmstead and Palmer, 1997). These biogeographical patterns are considered vestiges of an extinct flora that existed in northern Africa prior to desertification (Axelrod, 1975; Axelrod and Raven, 1978; Sunding, 1979; Bramwell, 1985; Potts and Behrensmeyer, 1992). The Sahara desert was well established approximately 2.5 Mya (Potts and Behrensmeyer, 1992) and the effects of Quaternary glacial aridity were so dramatic that they even made an impact in tropical Africa (Street, 1981; Elenga *et al.,* 1994).

Phenetic studies of flavonoid data (Aguinagalde and Gómez-Campo, 1984) and floral traits (Clemente and Hernández-Bermejo, 1978a,b, 1980a,b) also suggested a close relationship between the Macaronesian *Crambe* and taxa from the Eurasian clade. Unfortunately, these studies did not include *C. kilimandscharica.* Furthermore, all Macaronesian taxa, except *C. sventenii* and *C. fruticosa,* have fruits that are virtually identical to those of *C. kilimandscharica.* These data, combined with geologic and climatic information, suggest that the Macaronesian clade may be related to the Eurasian clade, as suggested by the strict consensus ITS trees from five of the eight analyses.

Two strongly supported groups of taxa with disjunct geographical distributions are revealed in all ITS trees (Fig. 4). The first involves *C. abyssinica* and *C. hispanica.* The former species is restricted to east Africa, whereas the latter occurs in the Mediterranean basin. These two species occur in the most derived position in
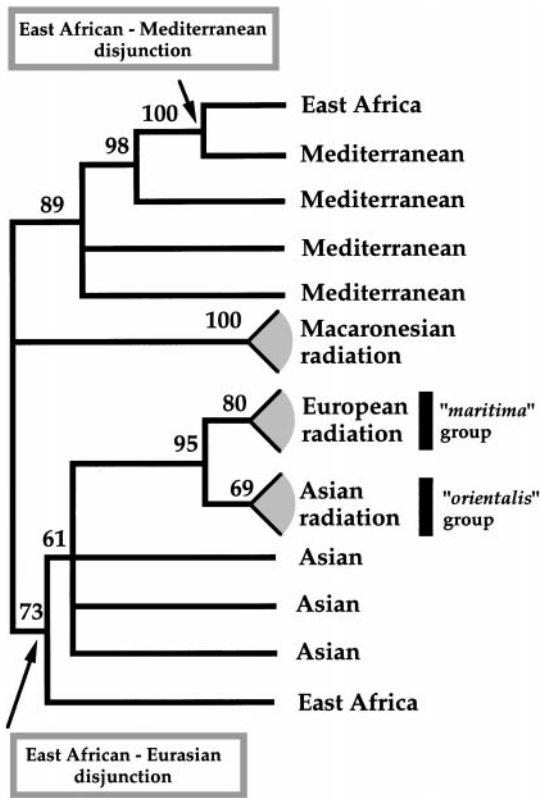
FIG. 4. Strict consensus tree obtained after the eight weighted and unweighted cladistic analyses. The geographic distribution of *Crambe* species is indicated at each terminal branch. Average bootstrap values from these eight analyses are shown along each node. The two major geographical disjunctions identified in the ITS tree are indicated. Geographical distribution of the *Crambe* species included in the analysis is also indicated at each terminal branch.

the Mediterranean clade and differ by only three nucleotide substitutions (Fig. 1). This derived position and the low level of divergence suggests that these species differentiated very recently, probably long after the desertification of northern Africa. This would be consistent with a hypothesis that this geographical disjunction is due to long distance dispersal. Populations of *C. hispanica* occur in the Middle East; therefore, it is possible that long-distance dispersal took place between this region and east Africa.

The second major biogeographic disjunction in the ITS trees concerns *C. kilimandscharica* and the taxa endemic to the Eurosiberian region and southwestern Asia. This disjunction is in a basal position in the Eurasian clade and a large number of changes separate the Eurosiberian–Asian taxa from *C. kilimandscharica* (Fig. 1). Thus, this split probably occurred much earlier than the disjunction involving *C. abyssinica* and *C. hispanica.*

The possibility of an ancient long distance dispersal of *Crambe* between the Eurosiberian–Asian region and east Africa cannot be ruled out. However, it is possible that prior to desertification of northern Africa and glaciation of the northern hemisphere, *Crambe* had a fairly continuous distribution between east Africa and

southwestern Asia. Axelrod (1973) and Coetzee (1993) argued that there was a continuous vegetation of savanna and sclerophyll thorn scrub in most of the Saharan region until the late Pliocene (3 Mya). Furthermore, there was a massive intercontinental interchange of the biota following the collision of the African (including Arabia) and the Asian plates at Iran in the middle Miocene (Rögl and Steininger, 1984). This led to the establishment of a north–south corridor for plants and animals along east Africa (Hedberg, 1965; Axelrod and Raven, 1978). An increase in the aridity of northeastern Africa since the end of the Miocene (6 Mya) and a severe decrease of winter temperatures in most of Europe and Asia in the early Pleistocene (1.6 Mya) may have led to the extinction of many species along this corridor and the establishment of genera with disjunct distributions in some regions of Eurasia and Africa.

Therefore, many of the species of *Crambe* currently in the postglacial plains of Central Europe and the Baltic region would have had a post-Pliocene origin. If this hypothesis is correct it would mean that *Crambe* in Eurasia originated from an area situated between the northern hemisphere ice shield and the deserts of northern Africa and southwestern Asia. *C. schugnana, C. edentula,* and *C. gordjaginii* are basal to the rest of the Eurosiberian–Asian taxa in all the ITS phylogenies. These three species occur in an area just north of the arid regions of southwestern Asia and the Middle East but south of the ice shield limit.

There are several examples of biogeographical disjunctions between eastern Africa and western Africa, northern Africa, temperate Eurasia, or the Mediterranean basin (Burtt, 1971; Thorne, 1972; Hamilton, 1974; Quézel, 1978; Demissew, 1993; Lovett, 1993). However, we have been unable to find any examples of disjunct distributions between east Africa and the Eurosiberian–Asian region. The ITS phylogeny of *Crambe* documents the first molecular evidence for such a connection and appears to support an old biogeographical link between these two regions.

Fossils of *Crambe* are not available to calibrate times of divergence. Therefore, any estimates of divergence times for the two major disjunctions in *Crambe* must be based on rates of ITS sequence evolution from other plant groups [i.e., Betulaceae (Savard *et al.,* 1993); Winteraceae (Suh *et al.,* 1993); *Dendroseris* D. Don (Asteraceae) (Sang *et al.,* 1994); *Gossypium* L. (Malvaceae) Wendel *et al.,* 1995); *Robinsonia* DC. (Asteraceae) (Sang *et al.,* 1995)]. As we would expect, across a broad range of taxa the rates among these groups show considerable variation, ranging between 11×10^{-10} substitutions per site per year for *Alnus* Mill. and *Betula* L. (Betulaceae) (Savard *et al.,* 1993) to 2.5×10^{-9} substitutions per site per year for ITS2 in *Gossypium* (Wendel *et al.,* 1995). This wide range of substitution rates makes it difficult to use ITS sequences to estimate divergence times in *Crambe.* An additional complication for estimation of divergence

times of *Crambe* based on molecular data is the possible existence of "generation-time effect" (Kohne, 1970; Li and Graur, 1991; Li, 1997). All Macaronesian and Eurosiberian–Asian taxa are perennial, whereas those from the Mediterranean and east Africa are annual. However, the molecular clock for *Crambe* was rejected in only one of the pairwise comparisons, suggesting that generation time has had minimal effect on rates of ITS evolution. We chose both the slowest and fastest rates for *Gossypium* (Wendel *et al.,* 1995) to estimate divergence times because this genus has similar life forms to those of *Crambe.*

The highest nucleotide divergence value between *C. kilimandscharica* and other member of the Eurasian clade was 7.31% (*C. kilimandscharica* vs *C. gordjaginii*). If we use the fastest published rate for sequence divergence [1.75% nucleotide divergence per million years for *Gossypium* (Wendel *et al.,* 1995)] this would date the current disjunction of *Crambe* between east Africa and Eurasia at approximately 4 Mya, before the establishment of the Sahara desert. This time of divergence corresponds with an arid interval of expansion of savannas and grasslands in most of northeastern Africa and the middle East. It appears that these extensive savannas were interrupted in these two regions only by isolated pockets of montane forests located in east Africa and southwestern Asia (Potts and Behrensmeyer, 1992). Some populations of *Crambe* could have found refuge in these mountains.

C. abyssinica and *C. hispanica* have a nucleotide divergence value of 0.67%. Using the fastest rate of 1.75% nucleotide divergence per million years indicates that these species separated 383,000 years ago. If we use the slowest rate of 0.49% reported in *Gossypium* (Wendel *et al.,* 1995) then the disjunction between *C. abyssinica* and *C. hispanica* would be approximately 1.4 Mya. Both estimates date the disjunction after the late Pliocene (2.5 Mya) when desert conditions were already established in extended areas of the present Sahara region. These divergence times support our hypothesis that this disjunction is due to long-distance dispersal between East Africa and the Mediterranean basin. Since populations of *C. hispanica* also occur in the Middle East, it is possible that long-distance dispersal took place between this region and east Africa. It has been established that post-Pliocene corridors existed, which could have connected the least arid zones of northeastern Africa (Koeniguer, 1992). However, it is not known to what extent these corridors could have linked the Ethiopian highlands to Mediterranean enclaves.

Crambe, together with *Brassica, Sinapis,* and *Raphanus,* is one of the four major crops of the subtribe Brassiceae (Gómez-Campo, 1980). In this study, we have demonstrated that the genus consists of three major lineages. We have also established that it is likely that the major climatic and geological changes that occurred since the Middle Miocene in Africa and Eurasia had a dramatic effect on the morphological differentiation and evolution of Tethyan elements of the Brassicaceae.

APPENDIX 1

Aligned ITS sequences from the taxa studied (left column). Gaps are indicated by "-." Uncertain sites are shown as "?." Polymorphic sites are as follows: M = A/C, R = A/G, S = C/G, W = A/T, Y = C/T. Taxon abbreviations are in Table 1. The six uncertain positions which were not included in the analysis are indicated by "*." Length of the sequence is indicated in brackets at the end of the sequence.

```
        →ITS1
              10            20            30            40            50            60
BROLE   TCG-TACCC-    -GGGAAA-CA    GAACGACCCG    AGAACGTTGA    AA-CATCACT    CT-CGGTGGG
CMTER   TCG-TATCC-    -TGGAAA-CA    GAACGACCTG    AGAACGTTGA    AA-CATCACT    CC-CGGTGGG
CRABY   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    TT-TGGTGGG
CRACU   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CATCACT    CT-CGGTGGG
CRAMA   TCG?TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CATCACT    CT-CGGTGGG
CRASP   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CAWCACT    CT-CGGTGGG
CRCOR   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CATCACT    CT-CGGTGGG
CREDE   TCG-?ATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    CT-CGGTGGG
CRFIL   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    CT-TGGTGGG
CRGIB   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CAWCACT    CT-CGGTGGG
CRGLA   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    CT-TGGTGGG
CRGOM   TCG-CATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATTA    AA-CATCACT    CT-TGGTGGG
CRGOR   TCG-?ATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CATCACT    CT-CGGTAGG
CRGRO   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    CT-CGGTGGG
CRHIS   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    TT-TGGTGGG
CRJUN   TCG-?ATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CATCACT    CT-CGGTGGG
CRKIL   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    CT-TGGTGGG
CRKOK   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    CT-CGGTGGG
CRKOT   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CATCACT    CT-CGGTGGG
CRKRA   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACGATGA    AA-CATCACT    CT-TGGTGGG
CRMAR   TCG-?ATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAACGA    AA-CATCACT    CT-CGGTGGG
CRORI   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CATCACT    CT-CGGTGGG
CRPER   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CATCACT    CT-CGGTGGG
CRPIN   TCG-TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAA?GA    AA-CATCACT    CT-CGGTGGG
CRPON   ????TATCC-    -TGGAAA-CA    GAACAACCCG    AGAACAATGA    AA-CA?CACT    CT-CGG?GGG
```

APPENDIX 1—*Continued*

```
                10              20              30              40              50              60
CRSAN  TCG-CATCC-  -TGGAAA-CA  GAACAACCCG  AGAACAATTA  AA-CATCACT  CT-TGGTGGG
CRSCH  TCG-?ATCC-  -TGGAAA-CA  GAACAACCCG  AGAACATTGA  AA-CATCACT  CT-CGGTAGG
CRSVE  TCG-CATCC-  -TGGAAA-CA  GAACAACCCG  AGAACAATTA  AA-CATCACT  CT-TGGTGGA
CRTAT  TCG-TATCC-  -TGGAAA-CA  GAACAACCCG  AGAACAA?GA  AA-CATCACT  CT-CGGTGGG
ERSAT  TCG-TACCC-  -TGGAAA-CA  GAACGACCCG  AGAACGATGA  AA-CATCACT  CT-CGGTGGG
GUARV  TCG-TACCC-  -TGGAAA-CA  GAACGACCCG  AGAACGATGA  AA-CATCACT  CT-CGGTAGG
HEFRU  TCG-TATCC-  -TGGAAA-CA  GAACAACCCG  ?GAACGTTGA  AA-CATCACT  CT-CGGTGGG
HIINC  TCG-TATCC-  -TGGAAA-CA  GAACGACCCG  AGAACGTTGA  AA-CATCACT  CC-CGGTGGG
KRCOR  TCG-TATCC-  -TGGAAA-CA  GAACAACCCG  TGAACGTTGA  AA-CATCACT  CT-CGGTGGG
MUPOS  TCG-TATCC-  -TGGAAA-CA  GAACGACCCG  AGAACGATGA  AA-CATCACT  CTTCGGTGGG
PHCHA  TCG-TACCC-  -TGGAAA-CA  GAACAACCCG  AGAAC?ATGA  AA-C?CCACT  CT-CGGTGGG
RASAT  TCG-TATCC-  -TGGAAA-CA  GAACGACCCG  AGAACGTTGA  AA-CATCACT  CT-CGGTGGG
RPRUG  TCG-?A?CC-  -TGGAAA-CA  GAATGACCCG  AGA?CGATGA  AA-CATCACT  CT-CGGTGGG
SIALB  TCG-TATC--  -TGGAAA-CA  GAACAACCCG  AGAACAATGA  AAACATCACT  CT-CGGTGGG
SNFRU  TCG-?A?CC-  -TGGAAA-CA  GAAC?ACCCG  AGAAC??TGA  AA-CATCACT  CT-CGGTGGG
CAARE  TCGATACCT-  GTCCAAAACA  GAACGACCCG  CGAACCAACG  AT-CACCACT  CT-CGGTGGG
CLIRR  TCGATACCTT  GACCAAAACA  GAACGACCCG  CGAACGATTG  AT-CATCACT  CT-CGGTGCG
CMMIC  TCGATACCT-  GTCCAAA-CA  GAACGACCCG  CGAACCAACG  AT-CACCACT  CT-CGGTAGG
ERWIT  TCGATACCT   GTCCAAAACA  GAACGACCCG  CGAACCAACG  AT-CACTACT  CG-?GGTGGG
SMALT  TCGATACCTA  GAC-----CA  GAACGACCCG  AGAACGATCG  AT-CACCACT  CT-CGGCAGG
                70              80              90             100             110             120
BROLE  CCGG-TA-TC  TTAGCTGATT  TCGT-GCCTA  CCGATTCCGT  GGTTATGCGT  TC---GTCAC
CMTER  CTGG-TC-TC  TTAGCCGATT  CCAT-GCCTG  CCGA?TCCGT  GGTTATGTGT  TC---GTCCT
CRABY  CCGG-TT-TC  TTAGTTGATT  CTGT-GCCTG  CCAATTCCGT  GGTTATGTGT  TTC--GTCCC
CRACU  CTGG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  TTCTCGTCCC
CRAMA  CTGG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGCGT  ????CGTCCC
CRASP  CYGG-TT-TC  TTTGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  ????CGTCCC
CRCOR  CYGG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CCRCTTCCGT  GGTTATGTGT  TC????TCCC
CREDE  CYGG-TT-TC  TTTRCTGATT  CYGT-GCCTG  CCATTTCCGT  GGTTATGTGT  TTT--GTCCC
CRFIL  CCGG-TT-TC  TTAGCTGATT  CTGT-GCCTG  CCAATTCCGT  GGTTATGTGT  TTC--GTCCC
CRGIB  CYGG-TT-TC  TTTGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  CTC--GTCCC
CRGLA  CCGG-TT-TC  TTAGTTGATT  CTGT-GCCTG  CCAATTCCGT  GGTTATGTGT  TTC--TTCCC
CRGOM  CCGG-TT-TC  TTAATTGATT  CTGT-GTCTG  CCAATTCCGT  GGTTATGTGT  TTT--GTCCC
CRGOR  CCGG-TT-TC  TTTGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  TTT--GTCCC
CRGRO  CTGG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CCRCTTCCGT  GGTTATGTGT  T????GTCCC
CRHIS  CCGG-TT-TC  TTAGTTGATT  CTGT-GCCTG  CCAATTCCGT  GGTTATGTGT  TTC--GTCCC
CRJUN  C?GG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CCGCTTCCGT  GGTTATGTGT  T????GTCCC
CRKIL  CCGG-TT-TC  TTAGATGATT  CTGT-GCCTA  CCAATTCCGT  GGTTATGTGT  TTC--GTCCC
CRKOK  CCGG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CCTATTCCGT  GGTTATGTGT  TTCTCGTCCT
CRKOT  C?GG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CC??TTCCGT  GGTTATGTGT  ???--GTCCC
CRKRA  CCGG-TT-TC  TTAGCTGATT  CTGT-GCCTG  CCAATTCCGT  GGTTATGTGT  TTC--GTCCT
CRMAR  CCGG-TT-TC  TTTGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  CTC--GTCCC
CRORI  CTGG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CCRCTTCCGT  GGTTATGTGT  TT?TCGT?CC
CRPER  CTGG-TT-TC  TTAGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  ????GTCCC
CRPIN  C?GG-TT-TC  TT?GCTGATT  CCGT-GCCTG  CC??TTCCGT  GGTTATGTGT  ????GTCCC
CRPON  CCGG-TT-TC  TTTGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  CTC--GTCC?
CRSAN  CCGG-TT-TC  TTAATTGATT  CTGT-TTCTG  CCAATTCCGT  GGTTATGTGT  TTC--?TCCC
CRSCH  CCGG-TT-TC  TWAGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  ATC--GTTCC
CRSVE  CCGG-TT-TC  TTGATTGATT  CTGT-GTCTG  CCAATTCCGT  GGTTATGCGT  TTT--GTCCC
CRTAT  CCGG-TT-TC  TTTGCTGATT  CCGT-GCCTG  CCACTTCCGT  GGTTATGTGT  CTC--GTCCC
ERSAT  CCGG-TC-TC  TTAGTTGATT  CCGT-GCCTG  CCGATTCCGT  GGTTATGTGT  TCC--GTCCC
GUARV  CCGG-TT-TC  GTAACTGTTT  CTGT-GCCTG  CCGATTCCGT  GGTTATGTGT  TCC--GTCCC
HEFRU  CCGG-T??TC  TTTGCTGATT  CCGT-GCCTG  CCGATTCCGT  GGTTWTGCGT  CAT--GTTCC
HIINC  CTGG-TC-TC  TTAACCGATT  CCAT-GCCTT  CCGGTTCCGT  GGTTATGTGT  TCC--GT?CC
KRCOR  CCGG-TT-TC  TTAGCTGATT  CCGT-GCCTA  CCGATTCCGT  GGTTATGGGT  CTC--GTCCC
MUPOS  CCGG-TT-TC  TTAACTGATT  CTGT-GCTCG  CCGATTCCGT  GGTTATGCGT  TAT??GTTCC
PHCHA  CATG-TC-TC  TTAGCTGATT  CTGT-GCCTG  CCGATTCCGT  GGTTATGCGT  GTC--GTCCA
RASAT  CTGGGTTCTC  TTACCGGAAT  CCAT-GCCTT  CCGGTTCCGT  GGTTATGTGT  TTC--GTCCT
RPRUG  CCGG-TC?TC  TTAACTGATT  CTGT-GCTCA  CCGATTCCGT  GGTTATGTGT  TTC--GTTCC
SIALB  CCGG-TT-TC  TTTGCTGATT  CTGT-GCCTG  CCGATTCCGT  GGTTATGCGT  TAA--GTTCC
SNFRU  CCGG-T?-TC  TTAATTGACT  CCGT-GCCTA  CCGATTCCGT  GGTTATG?GT  TCC--GTCCC
CAARE  CCGG-TT-TC  TTAACCGATC  CCGT-GCTCG  CCGGATCCGT  GGTTTCGCGT  ATC--GCCCT
CLIRR  CCGA-TT-TC  TTAACGGATC  TCGT-GCCTG  CCGATTCCGT  GGTTTCGCGT  AAG--TTCCT
CMMIC  CCGG-TT-TC  TTAGCCGATC  CCGTTGCCTG  CCGTCTCCGT  GGTTTCGCGT  ATC--TTCCC
ERWIT  CTGG-TT-TC  TTAGCCGATC  CCTT-G?CC?  CCGGATCCGT  GG?TTCGTGT  ACC--GTCCC
SMALT  CCGG-TT-TC  TTAGCCGATC  CCGT-GCTTG  CCGAGTCCGC  GGTTATGTGT  AT---GATCT
                                                                              ******
                130             140             150             160             170             180
BROLE  CGGCCTAG?-  ---------T  TTCGGTT-GG  AT--TGTA-C  GCATAGCTTC  CGGATATCAC
CMTER  CGGTCAAGAC  TTCAG----T  CTTGGCT-GG  TT--CGTG-C  ACATAGCTTC  CGGATATAAC
CRABY  T---CAAGGC  GTTAG----T  CTTGGTT-GG  TT--CGTG-C  ACATAGCTTC  CGGATATCAC
CRACU  CGGTCAAGGC  TTCAG----T  CTTGGTT-GG  GG--CGTG-C  ACATAGCTTC  CGGATATCAC
CRAMA  CGCTCAAGGC  TTCAG----T  CTTGGTT-GG  GG--CGTG-C  ACATAGCTTC  CGGATATCAC
CRASP  TGWTCAAGGC  TTCAG----T  CTCGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATATCAC
CRCOR  CG??CAAGGC  TTCAG----T  CTTGGTT-GG  G?--CGTG-C  ACATAGCTTC  CGGATATCAC
CREDE  YGGTCAAGGC  TTCAG----T  CTTGGTT-GG  GT--CATG-C  ACATAGCTTC  CGGATATCAC
CRFIL  TGGTCAAGGC  GTTAG----T  CTTGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATTTAAC
CRGIB  TGTTCAAGGC  TTCAG----T  CTCGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATATCAC
```

APPENDIX 1—*Continued*

```
              130         140         150         160         170         180
CRGLA    TGGTCAAGGC  GTTAG----T  CTTGGTT-GG  GT--CGTG-C  GCATAGCTTC  CGGATATCAC
CRGOM    CGGTCAAGAC  TTCAG----T  CTTGGTT-GG  GT--CACG-C  GCATAGCTTC  CGGATATCAC
CRGOR    TGGTCAAGGC  TTCAG----T  CTTGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATATAAC
CRGRO    CGGTCAAGGC  TTCAG----T  CTTGGTT-GG  GG--CGTG-C  AYATAGCTTC  CGGATATCAC
CRHIS    T----CAAGGC GTTAG----T  CTTGGTT-GG  TT--CGTG-C  ACATAGCTTC  CGGATATAAC
CRJUN    TGGTCAAGGC  TTCAG----T  CTTGGTT-GG  G?--CGTG-C  ACATAGCTTC  CGGATATCAC
CRKIL    ?AGTCAAGGC  TTCAG----T  CTTGGTT-GG  GT--CATG-C  ACATAGCTTC  CGGATATCAC
CRKOK    TGGTCAAGGC  TTCAG----T  CTTGGTT-GG  GG--CGTG-C  ACATAGCTTT  CGGATATCAC
CRKOT    CG?TCAAGGC  TTCAG----T  CTTGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATATCAC
CRKRA    TGGTCAAGGC  GTCAG----T  CTTGGTT-GG  GT--CGTG-C  GCATAGCTTC  -GGATATCAC
CRMAR    TGTTCAAGGC  TTCTG----T  CTCGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATATCAC
CRORI    CGGTCAAGGC  TTCAG----T  CTTGGTT-GG  GG--C?TG-C  ACATAGCTTC  CGGATATCAC
CRPER    CGGTCAAGGC  TTCAG----T  CTTGGTT-GG  GG--CGTG-C  ACATAGCTTC  CGGATATCAC
CRPIN    ?GTTCAAGGC  TTCAG----T  CTCGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATATCAC
CRPON    TGTTCAAGGC  TTCTG----T  CTCGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATATCAC
CRSAN    CRGTCAAGAC  TTCAG----T  CTTGGTT-GG  GT--CATG-C  GCATAGCTTC  CGGATATCAC
CRSCH    CGGTCAAGGC  TTCAG----T  CTTGGTTTCG  GT--CGGG-C  ACATAGCTTC  CGGATATCAC
CRSVE    CGATGAAGAC  TTCAG----T  CTTGGTT-GG  GT--CATG-C  GCATAGCTTC  CGGATATCAC
CRTAT    TGTTCAAGGC  TTCTG----T  CTCGGTT-GG  GT--CGTG-C  ACATAGCTTC  CGGATATCAC
ERSAT    CGGTCAAGAT  TTTTG----T  CTTGGTT-GG  TT--CGTG-C  GCATAGCTTC  CGGATTTCAC
GUARV    CGGTCAAGAC  TTCAG----T  CTCGGTC-GG  AT--CGTG-C  GCATAGCTTC  CGGATTTCAC
HEFRU    CAGCCAATAC  TTCAG----T  CTTGGTT-GG  GT--CGTG-C  GCATTGCTTC  CGGAT?TCAC
HIINC    CGGTTGAGAC  TTCACCGAGT  CTCGGCT-GG  AT--CGTG-C  ACATAGCTTC  CGGATTTAAC
KRCOR    CGGCCAGAT   TTCTT----T  CTCGGTT-GG  AT--CGTG-C  CCATGGCTTC  CGGATATCAC
MUPOS    CATCCAAGAT  TTCAA----T  CTCGCTT-GG  GT--CATG-C  GCATAGCTTC  CGGATATCAC
PHCHA    TGTCTGAGGC  GTTTG----T  CTCGATC-TG  GT--CGTC-C  GCATAGCTTC  CGGATATCAC
RASAT    CGGTCAAGGC  T?CAG----?  CTGGGCT-GG  GT--CGTG-C  --ACAGCTTC  CGGAGATCAC
RPRUG    CGGTCAAGAC  TTTAAG---T  CTTGGTT-GG  AT--CGTG-C  GCATAGCTTC  CGGATTTCAC
SIALB    CAGCCAGTAC  TTCAG----T  CTTGGTT-GG  GT--CTTG-C  GCATTGCTTC  CGGATATCAC
SNFRU    AGATCATGAC  TTCTG----T  CTCGAAC-TG  AT--CGTG-C  GCATAGCTTC  CGGATATCAC
CAARE    GGTCGGGAGC  TTTA-----T  CTCGGTC-TT  GT--CGTG-C  GCGTTGCTTC  CGGATATCAC
CLIRR    TGTCTTGAA?  TTCTT----G  GCTAGTA-GG  T---CTTG-C  GCGTTGCTTC  CGGATATCAC
CMMIC    GGTCGGGAGC  TCTA-----T  CTCGGTC-TG  GT--CGTG-C  GCGTTGCTTC  CGGATATCAC
ERWIT    GGTCGGGAGC  TCTA-----T  CTCGGTC-TG  GT--CGAGTC  ACGTTGCTTC  CGGATATCAC
SMALT    CAA-------  ---------T  CTCGGTT-GT  ATATCATG-C  GCATAACTTG  TGGATTTCAC
              190         200         210         220         230         240
BROLE    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATTCAACTA  AACAGCCT?-  -?TTTCGCCA
CMTER    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATACAACTA  AACAGTCTG-  -CTTTCGCCA
CRABY    CAAACCCCGG  CAC-GAGAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TTTTCGCCA
CRACU    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TGTTCGGCA
CRAMA    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TGTTCGGCA
CRASP    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -CGTTCGGCA
CRCOR    CAAACCCCGG  CAY-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -YGTTYGGCA
CREDE    CAAACCACGG  CAY-GAAAAG  TGTCAAGGAA  CATGCAATTA  AACAGYCTG-  -AGTTCGGCA
CRFIL    CAAACCCCGG  CAC-GAGAAG  TGTCAAGGAA  CATGCAACTA  AACAGCCTG-  -TTTTCGCCA
CRGIB    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -CGTTCGGCA
CRGLA    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCCTG-  -TGTTCGCCA
CRGOM    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCCTG-  -CATTCGCCA
CRGOR    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACCA  AACAGCCTG-  -CGTTCGGCA
CRGRO    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TGTTCGGCA
CRHIS    CAAACCCCGG  CAC-GAGAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TTTTCGCCA
CRJUN    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TGTTCGGCA
CRKIL    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCCTG-  -CATTCGGCA
CRKOK    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TGTTCGGCA
CRKOT    CAAACCCCGG  CACTGAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -YGTTYGGCA
CRKRA    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCCTG-  -TGTTCGCCA
CRMAR    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGTAACTA  AACAGCTTG-  -CGTTCGGCA
CRORI    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TGTTCGTCA
CRPER    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -TGTTCGGCA
CRPIN    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCTTG-  -CGTTCGGCA
CRPON    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGYAACTA  AACAGCTTG-  -CGTTCGGCA
CRSAN    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCCTG-  -CATTCGCCA
CRSCH    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCCTG-  -CATTCGGCA
CRSVE    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCCTG-  -CATTCACCA
CRTAT    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGYAACTA  AACAGCTTG-  -CGTTCGGCA
ERSAT    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAATTA  AACAGCCTG-  -CGTTCGCCA
GUARV    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGTCTG-  -CTTTCGTCA
HEFRU    CAAACCCCGG  CAC-AACAAG  TGTCAAGGAA  CATTCAACTA  GGTAGCCTS-  -GTT-CGCCA
HIINC    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATACAACTA  AACAGTCTG-  -CTTTCACCA
KRCOR    AAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGCC??-  -?TTTCGCCG
MUPOS    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  YATGCAACTA  AACAGCCT?-  -CTT-CGCCA
PHCHA    AAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  AATGCAACTA  AACAGCCTG-  -GGTTTCGCA
RASAT    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGTCTC-  -GTTTCGCCA
RPRUG    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACTA  AACAGTCTG-  -CTTTCCGCA
SIALB    CAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATTCAACTA  GGTAGCCTG-  -CTTTCGCCA
SNFRU    AAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  AATGCAACTA  AACAGTCTG-  -CTTTCACCA
CAARE    AAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CAAGCAAACG  AACGGCTGG-  -CATTCGCCT
CLIRR    AAAACCCCGG  CAC-AAATAG  TGTCAAGGAA  CATGCAACCA  AACAGTCTG?  ?TGTTCGCCT
```

APPENDIX 1—*Continued*

```
              190         200         210         220         230         240
CMMIC   AAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACCG  AACGGCTTTG  GCATTCGCCT
ERWIT   CAAACCCCGG  ?AC-GAAA?G  TGTC?AGGAA  CATGCAACCG  AACGGCCGG-  -CATTCGCCT
SMALT   AAAACCCCGG  CAC-GAAAAG  TGTCAAGGAA  CATGCAACAT  AACAGCCTG-  -CGTTCGCCT
                                                                    →ITS2
              250         260         270         280         290         300
BROLE   ACCCGGAGAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCYC
CMTER   ACCCGGAGAC  GGTGTTTGTT  CAGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRABY   ACCCGGAAAT  GGTGTTTGTT  CGGAAGTAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRACU   ACCCGGAAAC  GGTGTTTGTT  MGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRAMA   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  TAAAGTCTAT  CGTCGTCCCC
CRASP   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAWTGT  -AAAGTCTAT  CGTCGTCCCC
CRCOR   ACCCGGAAAC  GGTGTTTGTT  CGGAAGYAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CREDE   ACCCGGAGAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAWTGT  -AAAGTCTAT  CGTCGTCYCC
CRFIL   ACCCGGAAAC  GGTGTTTGGT  CGGAAGTTGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRGIB   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRGLA   ACCCGGAAAC  GGTGTTTGGT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRGOM   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCATTGT  -AAAGTCTAT  CGTCGTCCC-
CRGOR   ACCCGGAGAC  GGTGTTTGTT  TGAAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRGRO   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRHIS   ACCCGGAAAC  GGTGTTTGGT  CGGAAGTAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRJUN   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRKIL   GCCCGGAGAC  GGTGTTTGTT  CGGAAGYAGT  GCTGCAATGT  -AAAATCTAT  CGTCGTCCCC
CRKOK   ACCCGGAAAC  GGTGTTTGAT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRKOT   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCART  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRKRA   ACCCGGAAAC  GGTGTTTGGT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRMAR   ACCCGGAAAC  GGTGTTTGTT  CAGAAGTAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRORI   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRPER   ACCCGGAAAC  GGTGTTTGTT  MGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRPIN   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRPON   ACCCGGAAAC  GGTGTTTGTT  CRGAAGYAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRSAN   ACCCGGAAAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCATTGT  -AAAGTCTAT  CGTCGTCCC-
CRSCH   ACCTGGAAAC  GGTGTTTGTT  CGAAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CRSVE   ACCCGGAGAC  GGTGTTTGTT  CGAAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCC-
CRTAT   ACCCGGAAAC  GGTGTTTGTT  CRGAAGYAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
ERSAT   ACCCGGAGAC  GGTGTTTGTG  CGGAAGTTGT  GCTGCAATTT  -AAAGTCTAT  CGTCGTCCCC
GUARV   ACCCGGAGAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
HEFRU   ACCCGGAGAC  GGTGTTTGTG  CGGAAGCTGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
HIINC   ACCCGGAGAC  GGTGTTTGTT  CGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
KRCOR   ACCCGGAGAC  GGTGTTTG?G  CGRAAGC???  GCTGCAA??T  -AAAGTCTAT  CGTCGTCCCC
MUPOS   ACCCGGAAAC  GGTGTTTGTG  CGGAAGTTGT  GCTGCAATTT  -AAAGTCTAT  CGTCGTCCCC
PHCHA   TCCCGGAGAC  GGTGTTTGTG  CGGAAGCTGT  GCTGCGATAG  ?AAAGTCTAT  CGTCGTCCCC
RASAT   ACCCGGARAC  GGTGTTTGTT  CGGAAACAGT  GCTGAAATGT  -AAAGTCTAT  CGTCGTCCCC
RPRUG   ACCCGGAAAC  GGTGTTTGTG  TGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
SIALB   ACCCGGAGAC  GGTGTTTGTG  CGGAAGCTGT  GCTCCAATGT  -AAAGTCTAT  CGTCGTCCCC
SNFRU   ACCCGGAGAC  GGTGTTTGTA  AGGAAGCAGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CAARE   CCCCGGAGAC  GGTGTGTGCG  CGGATGCTGT  GCTGCGATAT  ?AAAGTCTAT  CGTCGTCCCC
CLIRR   CCCCGGAGAC  GGTGTGTGTG  CGGATGCGGT  GCTGCAATGT  -AAAGTCTAT  CGTCGTCCCC
CMMIC   CCCCGGAGAC  GGTGTGTGCG  CGGATGCTGA  GCTGCGATCT  -AAAGTCTAC  CGTCGTCCCC
ERWIT   CCC?GGAGAC  GGTGTGTGCG  TGGATGCTGA  GCTGCGATCT  -AAAGTCTAT  CATCGTCCCC
SMALT   ACCCGGAAAC  GGTGTTTGCG  TGAACGTTGT  GCTGGAATCT  -AAAGTCTAT  CTTCGTCCCC
              310         320         330         340         350         360
BROLE   --CAA-TCCT  CTC-GAGGAT  ATCGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CMTER   --CCA-TCCT  CTC-GAGGAT  ATTGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCATGCGG
CRABY   -TCCA-TCCT  CTT-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRACU   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRAMA   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRASP   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRCOR   --CCA-TCTT  CTCGGAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  C??CACGCGG
CREDE   --CCA-TCCT  CAC-CAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCACACGCGG
CRFIL   --CCA-TCCT  CAT-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRGIB   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRGLA   --CCA-TCCT  CTT-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCG?
CRGOM   --CCA-TCCT  CTT-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRGOR   --CCA-ACCT  CTT-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRGRO   --CCA-TCTT  CTC-GAGGAT  ATGGGACRGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRHIS   ?TCCA-TCCT  CTT-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRJUN   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRKIL   --CCA-TCCT  CTC-AAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRKOK   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  C-GCACGCGG
CRKOT   --CCA-TCTT  CT?-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRKRA   -TCCA-TCCT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  C-GCACGCGG
CRMAR   --CCA-TCTT  CTA-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRORI   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  C??CACGCGG
CRPER   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRPIN   --CCA-TCTT  CTC-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRPON   --CCA-TCTT  CTM-GAGGAT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRSAN   --CCA-TCCT  CTC-GAG-AT  ATGGGACGGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
CRSCH   --CCA-TCCW  CTC-TAGGAT  ATGGGACRGA  AGCTGGTCTC  CCGTGTGTTA  CCGCACGCGG
```

APPENDIX 1—*Continued*

```
              310          320          330          340          350          360
CRSVE   --CCA-TCCT   ATC-GAGGAT   AAGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
CRTAT   --CCA-TCTT   CTA-GAGGAT   ATGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   C?GCACGCGG
ERSAT   --CCA-TCCT   CTC-GAGGAT   AAGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
GUARV   --CCA-TCCT   CTA-GAGGAT   ATGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
HEFRU   --CCA-TCCT   CTC-GAGGAT   ACGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
HIINC   --CCA-TCCT   CTC-GAGGAT   ATCGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
KRCOR   --CCA-TCCT   CTC-GAGGAT   GTTGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
MUPOS   --CCA-TCCT   TTC-GAGGAT   ATGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
PHCHA   --CCA-TCCT   CTC-GAGGAT   ATTGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
RASAT   --CCA-TCCT   CTC-GAGGAT   ATAGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
RPRUG   --CCA-TCCT   CTC-CAGGAT   ATGGGACGGA   AGATGGTCTC   CCGTGTGTTA   CCGCACGCGG
SIALB   --CCA-TCCT   CTC-GAGGAT   ATGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
SNFRU   --CCA-TCCT   CTT-GAGGAT   ATGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
CAARE   -TATATTCCT   TTG--CGGAT   GCGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
CLIRR   --CCA-TCCT   CTT--TGGAT   AAGGGACGGA   AGTTGGTCTC   CCGTGTGTTA   CCGCACGCGG
CMMIC   -CAAA-TCCT   TTG--CGGAT   ACGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG
ERWIT   CCTCA-TCCT   CTA--AGGAT   TCGGGACGGA   A?CTGGTCTC   CCGTGTGTRTTA  CCGCACGCGG
SMALT   -CCCA-TCCT   TTC--AGGAT   AAGGGACGGA   AGCTGGTCTC   CCGTGTGTTA   CCGCACGCGG

              370          380          390          400          410          420
BROLE   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAG   -GAGCGTCTT   GACATGCGGT   GGTGAATT-C
CMTER   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRABY   TTGGCCAAAA   TCGGAGTGAA   GGAC-GCTAG   -GAGCGTCTC   GACATGCGGT   GGTTAATT-C
CRACU   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRAMA   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRASP   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRCOR   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CREDE   TTGGCCAAAA   TCTGAGCTAA   GGAT-GCCAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRFIL   TTGGCCAAAA   TCTGAGCAAA   GGAC-GCTAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRGIB   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRGLA   TTGGCCAAAA   TCTGAGCAAA   GGAC-GCCAG   -GTGCGTCTC   GACATGCGGT   GGTGAATT-A
CRGOM   TTGGCCAAAA   TCTGAGCGAT   GGAC-GTCAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRGOR   TTGGCCAAAA   TCTGAGCTAA   GGAT-GCCTG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRGRO   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCWT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRHIS   TTGGCCAAAA   TCTGAGTAAA   GGAC-GCTAG   -GAGCGTCTC   GACATGCGGT   GGTTAATT-C
CRJUN   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRKIL   TTGGCCAAAA   TCTGAGCTAA   GGAT-GCCAA   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRKOK   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRKOT   TTGGCCAAAA   TCYGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRKRA   TTGGCCAAAA   TCTGAGCAAT   GGAC-GCTTG   -GTGCGTCTC   GACATGCGGT   GGTGAATTTC
CRMAR   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRORI   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRPER   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRPIN   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRPON   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRSAN   TTGGCAAAAA   TCTGAGCGAT   GGAC-GTCAG   -GAGCGTCTC   GACATGCGGT   GGTGGATT-C
CRSCH   TTGGCCAAAA   TCTGAGCTAA   GGAT-GCCAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CRSVE   TTGGCCAAAA   TCTGAGCCAT   GGAC-GTCAG   -GAGCGTCTT   GACATGCGGT   GGTGAATT-C
CRTAT   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
ERSAT   TTGGCCAAAA   TCCGAGCAAA   GGAC-GCCTG   -GAGTGTCTC   GACATGCGGT   GGTGAATT-C
GUARV   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
HEFRU   TTGGCCAAAA   TCCGAGCTAA   GGAACGCAAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
HIINC   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
KRCOR   CTGGCCAAAA   TCCGAGCTAA   GAAC-GCCGT   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
MUPOS   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCAAG   -GAGCGTCTT   GACATGCGGT   GGTGAATT-C
PHCHA   TTGGCCGAAA   TCCGAGCCAA   GGAT-GCCAT   -GAGCGTCTT   GACATGCGGT   GGTGAATT-A
RASAT   TTGGCCAAAA   TCCGAGCTAA   GGAT-GCCAG   -GAGCGTCTT   GACATGCGGT   GGTGAATT-C
RPRUG   TTGGCCAAAA   TCCGAGCCAA   GGAC-GCCAG   -GAACGTCTC   GACATGCGGT   GGTGAATT-C
SIALB   TTGGCCAAAA   TCCGAGCTAA   GGAC-GTTTT   GGAGCGTCTC   GACATGCGGT   GGTGAATT-G
SNFRU   TTGGTCAAAA   TCCGAGCTAA   GGAT-GCAAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-C
CAARE   TTGGCCTAAA   TCCGAGCCAA   GGAC-GCCTG   -GAGCGTCCC   GACATGCGGT   GGTGAACA-T
CLIRR   TTGGCCAAAA   TCCGAGCTAA   GGAC-GCCAG   -GAGCGTCTC   GACATGCGGT   GGTGAATT-A
CMMIC   TTGACCTAAA   TCCGAGCCAA   GGAC-GCCTG   -GAGCGTTCC   GACATGCGGT   GGTGAACT-T
ERWIT   TTGGTCAAAA   TCCGAGCCAA   GGAC-AGCAG   GGAGCGTCCC   GACATACGGT   GGTGAAA?-T
SMALT   TTGGCCAAAA   TCCGAGCAAA   GGAT-GCAAG   -GAGCGTCCC   GACATGCGGT   GGTGAATT-C

              430          440          450          460          470          480
BROLE   AAT--TCTCG   T--CAAATCG   TCAGTCGTTT   CGGTCCGAAA   GCTCTTGAT-   GACCC-AAAG
CMTER   AATCTCCTCG   T--CATACAG   TCGGTCATTC   CGGTCCAAAA   GCTCTTGAT-   GACCC-AAAG
CRABY   AAA--CCTCG   T--CATTTTG   TCGGTCGTTC   CGGTCCAAAA   GCTCTTGAT-   GACCC-AAAG
CRACU   AAA--CCTCG   T--CATAGTG   TTGGTCGTTC   TGGTCCTAAA   GCTCTTGAT-   GACCC-AAAG
CRAMA   AAA--CCTCG   T--CATAGTG   TTGGTCGTTC   TGGTCCTAAA   GCTCTTGAT-   GACCC-AAAG
CRASP   AAA--CCTCG   T--CATAATG   TTGGTCGTTC   TGGTCCTAAA   GCTCTTGAT-   GACCC-AAAG
CRCOR   AAA--CCTCG   T--CATAGTG   TTGGTCGTTC   TGGTCCTAAA   GCTCTTGAT-   GACCC-AAAG
CREDE   TAA--CCTCG   T--CATAATG   TTGGTTGTT?   ?GATCCTAAA   GCTCTCGAT-   GACCC-AAAG
CRFIL   AAA--CCTCG   T--CATTTTG   TCGGTCGTTC   CAGTCCAAAA   GCTCTTGAT-   GACCC-AAAG
CRGIB   AAA--CCTCG   T--CATAATG   TTGGTCGTTC   TGGTCCTAAA   GCTCTTGAT-   GACCC-AAAG
CRGLA   AAA--CCTCG   T--CATATTG   TCGGTCGTTC   CGGTCCAAAA   GCTCTTGAT-   GACCC-AAAG
CRGOM   AAA--CCTCG   T--CATATCG   TCGGTCGTTC   CGGTCCAAAA   GCTCTCGAT-   GACCC-AAAG
CRGOR   AAA--CCTCG   T--CATAATG   TTGGTCGTTT   CGATCCTAAA   GTTCTTGAT-   GACCC-AAAG
```

APPENDIX 1—*Continued*

	430	440	450	460	470	480
CRGRO	AAA--CCTYG	T--CATAGTG	TTGGTCGTTC	TGGTYCTAAA	GCTCTTGAT-	GACCC-AAAG
CRHIS	AAA--CCTCG	T--CATTTTG	TCGGTCGTTC	CGGTCCAAAA	GCTCTTGAT-	GACCC-AAAG
CRJUN	AAA--CCTCG	T--CATAGTG	TTGGTYGTTC	TGGTCCTAAA	GCTCTTGAT-	GACCC-AAAG
CRKIL	AAT--CCTCG	T--CATAATG	TTGGTCGTTC	TGGTCCTAAA	GCTCTCGAT-	GACCC-AAAG
CRKOK	AAA--CCTCG	T--CATAGTG	TTGGTCGTTC	TGGTCCTAAA	GTTCTTGAT-	GACCC-AAAG
CRKOT	AAA--CCTCG	T--CATARTG	TTGGTCGTTC	TGGTCCTAAA	GCTCTTGAT-	GACCC-AAAG
CRKRA	AAA--CCTCG	T--CATATTG	TCGGTCGTTC	CGGTCCAAAA	GCTCTTGAT-	GACCC-AAAG
CRMAR	AAA--CCTCG	T--CATAGTG	TTGGTCGTTC	TGGTCCTAAA	GCTCTTGAT-	GACCC-AAAG
CRORI	AAA--CCTCG	T--CATAGTG	TTGGTCGTTC	TGGTCCTAAA	GCTCTTGAT-	GACCC-AAAG
CRPER	AAA--CCTCG	T--CATAGTG	TTGGTCGTTC	TGGTCCTAAA	GCTCTTGAT-	GACCC-AAAG
CRPIN	AAA--CCTCG	T--CATAATG	TTGGTCGTTC	TGGTCCTAAA	GCTCTTGAT-	GACCC-AAAG
CRPON	AAA--CCTCG	T--CATAATG	TTGGTCGTTC	TGGTCCTAAA	GCTCTTGAT-	GACCC-AAAG
CRSAN	AAA--CCTCG	T--CATTT-G	TTGGTCATTC	CGGTCCAAAA	GCTCTCG-T-	GACCC-AAAG
CRSCH	AAA--CCT?G	T--CATAATG	TTGGTCGTT?	?GATCCTAAA	GCTCTCGAT-	GACCC-AAAG
CRSVE	AAA--CCTCG	T--CACATCG	TCGGTCGTTC	CGGTCCAAAA	GCTCTCGAT-	GACCC-AAAG
CRTAT	AAA--CCTCG	T--CATAATG	TTGGTCGTTC	TGGTCCTAAA	GCTCTTGAT-	GACCC-AAAG
ERSAT	AAA--ACTCG	T--CATACTG	TCGATCATTC	CGGTCCAAAA	GCTCTTGAT-	GACCC-AAAG
GUARV	AATCTCCTCG	T--TATAAAG	TCGATCGTTC	CGGTCCAAAA	GCTCTCGAT-	GACCC-AAAG
HEFRU	TAA--CCTCG	T--CATATTG	TCGGTCGTTC	CGGTTCAAAA	GCTCTTGAT-	GACCC-AAAG
HIINC	AATCTCCTCG	T--GATACAG	TCGGTCGTTC	CGGTCCAAAA	GCTCTTGAT-	GACCC-AAAG
KRCOR	AAA--CCTCG	T--CATATTG	TCGGTCGTTC	CGGTTTAAAA	GCTCTCGATT	GACCC-AAAG
MUPOS	AAA--CCTCG	T--CATATTG	TCAGACGTTC	TGATCCGAAA	ACTCTAGAT-	GACCC-AAAG
PHCHA	AAA--ACTCG	T--CATACTG	TCG-TCGCTC	TGGTCCGAAA	GCTCTTGAT-	GACCC-AAAG
RASAT	AATCTCCTCG	T--CATATCG	TCGGTCGTTC	CGGTCCAAAA	GCTCTCGAT-	GACCC-AAAG
RPRUG	AA---CCTCG	T--CATATTG	TTGTTCGTTC	CGGTCCAGAA	GCTCTCGAT-	GACCC-AAAA
SIALB	TAA--CCTCG	T--CATATTG	TCGGTCGTTC	CGGTTCAAAA	GCTCTTGAT-	GACCC-AAAG
SNFRU	AAA--CCTCG	T--CATATTG	TCGGTCGTTT	TGGTCCAAAA	GCTTTTGAT-	GACCC-AAAG
CAARE	GAT--CCATT	T--GATATTG	TCGGTCGCTC	CTGTCCGGAA	GCTGTAGAT-	GACCC-AAAG
CLIRR	AAG--CATCT	T--CATATTG	TCGGATGCTC	CTGTCCGGAT	GCTTTCGAT-	GACCC-AAAG
CMMIC	GAT--CCACT	TTCCCTATCG	TCGGTCGCTC	TTGTCCGGAA	GCTCTAGAT-	GACCC-AAAG
ERWIT	GAT--CCACT	?T-CATATCG	TCGGCCGCTC	CTGTCCGGAG	GCTCTAGAT-	GACCCCAAAG
SMALT	AAG--CCTCG	T--CATATCG	TCGGTCGCTC	CTGTCTGGAA	GCTCTCGAT-	GACCC-AAAG

APPENDIX 1—*Continued*

	488	
BROLE	TCCTCAAC	[445]
CMTER	TCCTCAAC	[453]
CRABY	TCCTCAAC	[450]
CRACU	TCCTCAAA	[454]
CRAMA	TCCTCAAA	[456]
CRASP	TCCTCAAA	[454]
CRCOR	TCCTCAAA	[453]
CREDE	TCCTTAAC	[452]
CRFIL	TCCTCAAC	[452]
CRGIB	TCCTCAAA	[452]
CRGLA	TCCTCAAC	[452]
CRGOM	TCCTCAAC	[451]
CRGOR	TCCTCAAC	[452]
CRGRO	TCCWCAAA	[454]
CRHIS	TCCTCAAC	[451]
CRJUN	TCCTCAAA	[454]
CRKIL	TCCTCAAC	[452]
CRKOK	TCCTCAA?	[452]
CRKOT	TCCTCAAC	[453]
CRKRA	TCCTCAA?	[452]
CRMAR	TCCTCAAA	[452]
CRORI	TCCTCAAA	[454]
CRPER	TCCTCAAA	[454]
CRPIN	TCCTCAAA	[454]
CRPON	TCCTCAAA	[453]
CRSAN	TCCTCAAC	[448]
CRSCH	TCCTCAAC	[453]
CRSVE	TCCTCAAC	[451]
CRTAT	TCCTCAAC	[452]
ERSAT	TCCTCAA?	[452]
GUARV	TCCTCAAC	[454]
HEFRU	TCCTCAAC	[453]
HIINC	TCCTCAAC	[458]
KRCOR	TTCTCAAC	[453]

APPENDIX 1—*Continued*

	488	
MUPOS	TCCTCAAC	[454]
PHCHA	TCCTCAAC	[452]
RASAT	TCCTCAAC	[454]
RPRUG	TTCTCAAC	[453]
SIALB	TCCTCAAC	[453]
SNFRU	TCCTCAAC	[452]
CAARE	TCCTTTAA	[456]
CLIRR	TCCTCAAC	[456]
CMMIC	TCTTCAAA	[458]
ERWIT	TCCTCCAA	[459]
SMALT	TCCTCAA?	[440]

ACKNOWLEDGMENTS

REFERENCES

Aguinagalde, I., and Gómez-Campo, C. (1984). The phylogenetic significance of flavonoids in *Crambe* L. (Cruciferae). *Bot. J. Linn. Soc.* **89:** 277–288.

Al-Shehbaz, I. A. (1984). The tribes of Cruciferae (Brassicaceae) in southern United States. *J. Arnold Arbor.* **65:** 343–373.

Al-Shehbaz, I. A. (1988). The genera of Anchonieae (Hesperidae) (Cruciferae; Brassicaceae) in the southeastern United States. *J. Arnold Arbor.* **69:** 192–212.

Anderson, V. L., Slanger, W. D., Boyles, S. L., and Berg, P. T. (1993). *Crambe* meal is equivalent to soybean meal for backgrounding and finishing beef steers. *J. Anim. Sci.* **71:** 2608–2613.

Axelrod, D. I. (1973). History of the Mediterranean ecosystem in California. *In* "Mediterranean Type Ecosystems: Origin and Structure" (F. Di Castri and H. A. Mooney, Eds.), pp. 225–277. Springer-Verlag, Berlin.

Axelrod, D. I. (1975). Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. *Ann. Missouri Bot. Gard.* **62:** 280–334.

Axelrod, D. I., and Raven, P. H. (1978). Late Cretaceous and Tertiary vegetation history of Africa. *In* "Biogeography and Ecology of Southern Africa" (M. J. A. Werger, Ed.), pp. 77–130. Junk, The Hague.

Ball, P. W. (1964). *Crambe* L. *In* "Flora Europaea, Vol. 1" (T. G. Tutin, V. H. Heywood, N. A. Burges, D. H. Valentine, S. M. Walters, and D. A. Webb, Eds.), pp. 344–345. Cambridge Univ. Press, Cambridge, UK.

Boufford, D. E., and Spongberg, S. A. (1983). Eastern Asian–eastern North American phytogeographic relationships—A history from the time of Linnaeus to the twentieth century. *Ann. Missouri Bot. Gard.* **70:** 423–439.

Bramwell, D. (1969). The genus *Crambe* (Cruciferae) in the Canary Islands flora. *Cuad. Bot. Canaria* **6:** 5–12.

Bramwell, D. (1985). Contribución a la biogeografía de las Islas Canarias. *Bot. Macaroné* **14:** 3–34.

Bramwell, D. (1995). A new name for *Crambe gigantea* (Ceb. & Ort.) Bramwell. *Bot. Macaroné* **22:** 111.

Bremer, K. (1993). Intercontinental relationships of Africa and South American Asteraceae: A cladistic biogeographic analysis. *In* "Biological Relationships between Africa and South America" (P. Goldblatt, Ed.), pp. 105–135. Yale Univ. Press, New Haven, CT.

Burtt, B. L. (1971). From the south: An African view of the floras of western Asia. *In* "Plant Life of South-West Asia" (P. H. Davis, P. C. Harper, and I. C. Hedge, Eds.), pp. 135–149. Bot. Soc. Edinburgh, Edinburgh.

Candolle, A. P. de (1821). "Regni Vegetabilis Systema Naturale," Vol. 2, Treuttel and Wurtz, Paris.

Candolle, A. P. de (1824). "Prodromus Systematis Naturalis Regni Vegetabilis," Vol. 1, Treuttel and Wurtz, Strasbourg and London.

Carlquist, S. (1965). "Island Life: A Natural History of the Islands of the World," Natural History Press, New York.

Chernyakovskaya, E. G. (1970). *Crambe* (Tourn.) L. *In* "Flora of the U.S.S.R." (V. L. Komarov and N. A. Bush, Eds.), Vol. 8, pp. 354–366. Israel Program for Scientific Translations, Jerusalem. [Translated from Russian by R. Lavoott].

Clemente, M., and Hernández-Bermejo, E. (1978a). El aparato nectarígeno de la tribu Brassiceae (Cruciferae). *Anales Inst. Bot. Cavanilles* **35:** 279–296.

Clemente, M., and Hernández-Bermejo, E. (1978b). La corola de la tribu Brassiceae. *Anales Inst. Bot. Cavanilles* **35:** 297–334.

Clemente, M., and Hernández-Bermejo, E. (1980a). El cáliz de la tribu Brassiceae (Cruciferae). *Anales Jard. Bot. Madrid* **36:** 77–96.

Clemente, M., and Hernández-Bermejo, E. (1980b). Clasificación jerárquica de las Brasiceas según caracteres de las piezas estériles de su flor. *Anales Jard. Bot. Madrid* **36:** 97–113.

Coetzee, J. A. (1993). African flora since the terminal Jurassic. *In* "Biological Relationships between Africa and South America" (P. Goldblatt, Ed.), pp. 37–61. Yale Univ Press, New Haven, CT.

Constance, L., Heckard, R., Chambers, K. L., Ornduff, R., and Raven, P. H. (1963). Amphitropical relationships in the herbaceous flora of the Pacific coast of North and South America: A symposium. *Q. Rev. Biol.* **38:** 109–177.

Czerepanov, S. K. (1995). "Vascular Plants of Russia and Adjacent States (the Former USSR)," Cambridge Univ. Press, Cambridge, UK.

Demissew, S. (1993). The genus *Thymus* (Labiatae) in Ethiopia. *Opera Bot.* **121:** 57–60.

Downie, S. R., and Katz-Downie, D. S. (1996). A molecular phylogeny of Apiaceae subfamily Apioideae: Evidence from nuclear ribosomal DNA internal transcribed sequences. *Am. J. Bot.* **83:** 234–251.

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf material. *Phytochem. Bull.* **19:** 11–15.

Elenga, H., Schwartz, and Vincens, A. (1994). Pollen evidence of late Quaternary vegetation and inferred climate changes in Congo. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **109:** 345–356.

Farris, J. S. (1989). The retention index and homoplasy excess. *Syst. Zool.* **38:** 406–407.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using bootstrap. *Evolution* **39:** 783–791.

Felsenstein, J. (1993). "PHYLIP (Phylogeny Inference Package), Version 3.5," distributed by the author, Department of Genetics, Univ. of Washington, Seattle.

Gómez-Campo, C. (1980). Morphology and morpho-taxonomy of the tribe Brassiceae. *In* "*Brassica* Crops and Wild Allies. Biology and Breeding" (S. Tsunoda, K. Hinata, and C. Gómez-Campo, Eds.), pp. 3–30. Japan Sci. Press, Tokyo.

Gómez-Campo, C. (1993). Evolutionary trends in the tribe Brassiceae. *In* "Proceedings of the 5th OPTIMA Meeting" (H. Demiriz and N. Ozhatay, Eds.), pp. 573–578. OPTIMA, Istanbul.

Gómez-Campo, C., and Hinata, K. (1980). A check list of chromosome numbers in the tribe Brassiceae. *In* "*Brassica* Crops and Wild Allies. Biology and Breeding" (S. Tsunoda, K. Hinata, and C. Gómez-Campo, Eds.), pp. 51–63. Japan Sci. Press, Tokyo.

Gómez-Campo, C., and Tortosa, M. E. (1974). The taxonomic and evolutionary significance of some juvenile characters in the Brassiceae. *Bot. J. Linn. Soc.* **69:** 105–124.

Hamilton, A. (1974). The history of the vegetation. *In* "East African Vegetation" (E. M. Lind and M. E. S. Morrison, Eds.), pp. 188–209. Longman, London.

Hedberg, O. (1965). Afroalpine flora elements. *Webbia* **19:** 519–529.

Higgins, D. G., Bleasby, A. J., and Fuchs, R. (1991). CLUSTAL V: Improved software for multiple sequence alignment. *Comput. Appl. Biosci.* **8:** 189–191.

Hillis, D. M., and Huelsenbeck, J. P. (1992). Signal, noise, and reliability in molecular phylogenetic analyses. *J. Hered.* **83:** 189–195.

Holmgren, P. K., Holmgren, N. H., and Barnett, L. C. (1990). Index herbariorum. Part 1: The herbaria of the world, 8th ed. *Regnum Veg.* **120:** 1–693.

Jalas, J., Suominen, J., and Lampinen, R. (Eds). (1996). "Atlas Florae Europaeae, Vol. 11. Cruciferae (*Ricotia* to *Raphanus*)," Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.

Jonsell, B. (1982). Cruciferae. *In* "Flora of Tropical East Africa" (R. M. Polhill, Ed.), pp. 1–75. Balkema, Rotterdam.

Khalilov, I. I. (1991a). The system of the genus *Crambe* (Brassicaceae). *Bot. Zhurn. (St. Petersburg)* **76:** 1612–1613. [In Russian].

Khalilov, I. I. (1991b). Generis *Crambe* L. (Cruciferae) sectiones tres novae. *Novosti Sist. Vyssh. Rast.* **28:** 78–79. [In Russian].

Khalilov, I. I., and Archangelsky, D. B. (1991). Palynological studies in the species of the genus *Crambe* (Brassicaceae) in connection with its systematics. *Bot. Zhurn. (St. Petersburg)* **76:** 1582–1586. [In Russian].

Kim, K.-J., and Jansen, R. K. (1994). Comparisons of phylogenetic hypothesis among different data sets in dwarf dandelions (*Krigia*): Additional information from internal transcribed spacer sequences of nuclear ribosomal DNA. *Plant Syst. Evol.* **190:** 157–185.

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

Kitamura, S. (1960). "Flora of Afghanistan," Kyoto Univ. Kyoto.

Kluge, A. G., and Farris, J. S. (1969). Quantitative phyletics and the evolution of anurans. *Syst. Zool.* **18:** 1–32.

Koeniger, J. C. (1992). Gisements fossiliferes continentaux, Paleographie et apysages forestiers en Europe Occidentale et dans le Nord de l'Afrique au Cenozoique Superieur. *Biogeographica* **69:** 135–155.

Kohne, D. E. (1970). Evolution of higher organism DNA. *Q. Rev. Biophys.* **33:** 327–375.

Kumar, S., Koichiro, T., and Nei, M. (1993). "MEGA: Molecular Evolutionary Genetics Analysis, version 1.01," distributed by the Pennsylvania State University, University Park, PA.

Lazzeri, L., Leoni, O., Conte, L. S., and Palmieri, S. (1994). Some technological characteristics and potential uses of *Crambe abyssinica* products. *Industr. Crops Products* **3:** 103–112.

Lessman, K. J., and Meier, V. D. (1972). Agronomic evaluation of *Crambe* as a source of oil. *Crop Sci.* **12:** 224–227.

Li, W.-H. (1997). "Molecular Evolution," Sinauer, Sunderland, MA.

Li, W.-H., and Graur, D. (1991). "Fundamentals of Molecular Evolution," Sinauer, Sunderland, MA.

Li, W.-H., and Tanimura, M. (1987). The molecular clock runs more slowly in man than in apes and monkeys. *Nature* **326:** 93–96.

Loockerman, D. J., and Jansen, R. K. (1996). The use of herbarium material for DNA studies. *In* "Sampling the Green World" (T. F. Stuessy and S. H. Sohmer, Eds.), pp. 205–220. Columbia Univ. Press, New York.

Lovett, J. C. (1993). Temperate and tropical floras in the mountains of eastern Tanzania. *Opera Bot.* **121:** 217–227.

Maddison, D. R. (1991). The discovery and importance of multiple islands of most parsimonious trees. *Syst. Zool.* **40:** 315–328.

Maddison, W. P., and Maddison, D. R. (1992). "MacClade, version 3," Sinauer, Sunderland.

Magulaev, A. J. (1986). Chromosome numbers in some species of flowering plants of the Crimea and Caucasus floras. *Bot. Zhurn. (St. Petersburg)* **71:** 1575–1578. [In Russian].

Meikle, R. D. (1977). "Flora of Cyprus," Royal Botanic Gardens, Kew.

Mouterde, P. (1986). "Nouvelle Flore du Liban et de la Syrie," Vol. 2, Dar El-Machreq, Beirut.

Mulder, J. H., and Mastebroek, H. D. (1996). Variation for agronomic characteristics in *Crambe hispanica,* a wild relative of *Crambe abyssinica. Euphytica* **89:** 267–278.

Nieschlag, H. J., and Wolff, I. A. (1971). Industrial uses of high erucic oils. *J. Am. Oil Chem. Soc.* **48:** 723–727.

Olmstead, R. G., and Palmer, J. D. (1997). Implications for the phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. *Syst. Bot.* **22:** 19–29.

Parks, C. R., and Wendel, J. F. (1990). Molecular divergence between Asian and North American species of *Liriodendron* (Magnoliaceae) with implications for interpretations of fossil floras. *Am. J. Bot.* **77:** 1243–1256.

Post, G. E. (1932). "Flora of Syria, Palestine, and Sinai," Vol. 1, American Press, Beirut.

Potts, R., and Behrensmeyer, A. K. (1992). Late Cenozoic terrestrial ecosystems. *In* "Terrestrial Ecosystems through Time: Evolutionary Paleoecology of Terrestrial Plants and Animals" (A. K. Behrensmeyer, J. D. Damuth, W. A. DiMichele, R. Potts, H.-D. Sues, and S. L. Wing, Eds.), pp. 419–541. Univ. of Chicago Press, Chicago.

Quézel, P. (1978). Analysis of the flora of Mediterranean and Saharan Africa. *Ann. Missouri Bot. Gard.* **65:** 280–334.

Quézel, P. (1985). Definition of the Mediterranean region and the origin of its flora. *In* "Plant Conservation in the Mediterranean Area" (C. Gómez-Campo, Ed.), pp. 9–24. Junk, Dordrecht.

Rathgeberg, J., and Capesius, I. (1989). Nucleotide sequence of the 18S–25S spacer region from mustard DNA. *Nucleic Acids Res.* **17:** 7522.

Rieseberg, L. H., and Soltis, D. E. (1991). Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trends Pl.* **5:** 65–84.

Rieseberg, L. H., Whitton, J., and Linder, C. R. (1996). Molecular marker incongruence in plant hybrid zones and phylogenetic trees. *Acta Bot. Neerl.* **45:** 243–262.

Robyns, W., and Boutique, R. (1951). Cruciferae. *In* "Flore du Congo Belge et du Ruanda-Urundi," (Comité exécutif de la flore du Congo Belge et Jardin Botanique de l'état, Eds.), Vol. 2, pp. 522–543, Publications de L'Institut National pour L'Etude Agronomique du Congo Belge, Brussels.

Rögl, F., and Steininger, F. F. (1984). Neogene Paratethys, Mediterranean and Indo-Pacific seaways. *In* "Fossils and Climate" (P. Brenchley, Ed.), pp. 171–200. Wiley, Chichester, UK.

Sang, T., Crawford, D. J., Kim, S.-C., and Stuessy, T. F. (1994). Radiation of the endemic genus *Dendroseris* (Asteraceae) on the Juan Fernández Islands: Evidence from sequences of the ITS regions of nuclear ribosomal DNA. *Am. J. Bot.* **81:** 1494–1501.

Sang, T., Crawford, D. J., and Stuessy, T. F. (1995). ITS sequences and the phylogeny of the genus *Robinsonia* (Asteraceae). *Syst. Bot.* **20:** 55–64.

Sang, T., Crawford, D. J., and Stuessy, T. F. (1997). Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *Am. J. Bot.* **81:** 1494–1501.

Santos-Guerra, A. (1983). "Vegetación y Flora de La Palma," Interinsular Canaria, Santa Cruz de Tenerife.

Santos-Guerra, A. (1996). *Crambe feuilleii* (Brassicaceae) Santos sp. nova for the flora of Hierro Island (Canary Islands). *In* "2nd Symposium Fauna and Flora of the Atlantic Islands" (Anonymous, Ed.), p. 162. Departamento de Biología, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. [Abstract].

Sarich, V. M., and Wilson, A. C. (1973). Generation time and genomic evolution in primates. *Science* **179:** 1144–1147.

Savard, L., Michaud, M., and Bousquet, J. (1993). Genetic diversity and phylogenetic relationships between birches and alders using ITS, 18S rRNA, and *rbc*1 gene sequences. *Mol. Phylogenet. Evol.* **2:** 112–118.

Schulz, O. E. (1919). Cruciferae—Brassiceae. Pars prima. Subtribus I. Brassicinae et II. Raphaniae. *In* "Das Pflanzenreich Regni Vegetabilis Conspectus," (A. Engler, Ed.), Vol. 4, pp. 1–290. Verlag von Wilhelm Engelmann, Leipzig.

Street, F. A. (1981). Tropical palaeoenvironments. *Progr. Phys. Geogr.* **5:** 157–185.

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

Sunding, P. (1979). Origins of the Macaronesian flora. *In* "Plants and Islands" (D. Bramwell, Ed.), pp. 13–40. Academic Press, London.

Swofford, D. L. (1993). "PAUP: Phylogenetic Analysis Using Parsimony, version 3.1.1," distributed by the Illinois Natural History Survey, Champaign, IL.

Thorne, R. F. (1972). Major disjunctions in the geographic ranges of seed plants. *Q. Rev. Biol.* **47:** 365–411.

Van Dyne, D. L., and Blase, M. G. (1989). Commercializing high erucic acid oil crops. *Trans. Missouri Acad. Sci.* **23:** 13–22.

Vargas, P., Baldwin, B. G., and Constance, L. (1997). Nuclear ribosomal DNA evidence for a western North America origin of Hawaiian and South American species of *Sanicula* (Apiaceae). *Proc. Natl. Acad. Sci. USA* **95:** 235–240.

Warwick, S. I., Black, L. D., and Aguinagalde, I. (1992). Molecular systematics of *Brassica* and allied genera (subtribe Brassicinae, Brassiceae)—chloroplast DNA variation in the genus *Diplotaxis*. *Theor. Appl. Genet.* **83:** 839–850.

Warwick, S. I., and Black, L. D. (1993). Molecular relationships in tribe Brassicinae (Cruciferae, tribe Brassiceae). *Can. J. Bot.* **71:** 906–918.

Warwick, S. I., and Black, L. D. (1997a). Molecular phylogenies from theory to application in *Brassica* and allies (tribe Brassiceae, Brassicaceae). *Opera Bot.* **132:** 159–168.

Warwick, S. I., and Black, L. D. (1997b). Phylogenetic implications of chloroplast DNA restriction site variation in subtribes Raphaninae and Cakilinae (Brassicaceae, tribe Brassiceae). *Can. J. Bot.* **75:** 960–973.

Warwick, S. I., and Black, L. D. (1997c). Molecular phylogeny of tribe Brassiceae (Brassicaceae); tribal status of *Calepina, Conringia and Orychophragmus. Am. J. Bot.* **84(Suppl):** 243–244 [abstract, 48th annual meeting of the AIBS].

Wen, J., and Zimmer, E. A. (1996). Phylogeny and biogeography of *Panax* L. (the gingseng genus, Araliaceae): Inferences from ITS sequences of nuclear ribosomal DNA. *Mol. Phylogenet. Evol.* **6:** 167–177.

Wendel, J. F., Schnable, A., and Seelanan, T. (1995). An unusual ribosomal DNA sequence from *Gossypium gossypioides* reveals ancient, cryptic, intergenomic introgression. *Mol. Phylogenet. Evol.* **4:** 298–313.

White, G. A. (1975). Distinguishing characteristics of *Crambe abyssinica* and *C. hispanica. Crop Sci.* **15:** 91–93.

White, G. A., and Solt, M. (1978). Chromosome numbers in *Crambe, Crambella,* and *Hemicrambe. Crop Sci.* **18:** 160–161.

White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* "PCR Protocols. A Guide to Methods and Applications" (M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, Eds.), pp. 315–322. Academic Press, San Diego.

Winship, P. R. (1989). An improved method for direct sequencing of PCR amplified material using dimethyl sulphoxide. *Nucleic Acids Res.* **17:** 1266.

Wu, C.-I, and Li, W.-H. (1985). Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA* **82:** 1741–1745.

Zohary, M. (1966). "Flora Palestina," Vol. 1, The Israel Academy of Sciences and Humanities, Jerusalem.