

**Phylogenetic Relationships among the Genera of the Subtribe Sonchinae
(Asteraceae): Evidence from ITS Sequences**



Seung-Chul Kim; Daniel J. Crawford; Robert K. Jansen

Systematic Botany, Vol. 21, No. 3 (Jul. - Sep., 1996), 417-432.

Stable URL:

<http://links.jstor.org/sici?sici=0363-6445%28199607%2F09%2921%3A3%3C417%3APRATGO%3E2.0.CO%3B2-L>

Systematic Botany is currently published by American Society of Plant Taxonomists.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aspt.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

Phylogenetic Relationships among the Genera of the Subtribe Sonchinae (Asteraceae): Evidence from ITS Sequences

SEUNG-CHUL KIM AND DANIEL J. CRAWFORD¹

Department of Plant Biology, The Ohio State University, Columbus, Ohio 43210–1293

¹Author for Correspondence

ROBERT K. JANSEN

Department of Botany, University of Texas, Austin, Texas 78713–7640

Communicating Editor: Jeff J. Doyle

ABSTRACT. Sequences from the internal transcribed spacer region (ITS) of nuclear ribosomal DNA were used to assess relationships among genera of subtribe Sonchinae (Lactuceae). The data suggest that Sonchinae is paraphyletic, and that the Dendroseridinae should not be recognized as a distinct subtribe. Several Macaronesian genera, along with woody Macaronesian species of *Sonchus* (all diploid, $2n = 18$), appear to have been derived from a common ancestor. *Prenanthes pendula* is more closely related to this Macaronesian group than to other members of *Prenanthes*, suggesting that *Prenanthes*, like *Sonchus*, is polyphyletic. The Juan Fernandez Islands endemic genus *Dendroseris* ($2n = 36$) is monophyletic, but its phylogenetic relationships to other Sonchinae remain uncertain. *Reichardia*, which forms the sister genus of the Sonchinae, is also monophyletic. *Launaea*, by contrast, is considered to be paraphyletic. The monotypic genus *Aetheorhiza* ($2n = 18$) is sister to the annual weedy species of *Sonchus*, rather than being closely related to *Launaea*. Two monotypic endemic genera from New Zealand, *Embergeria* ($2n = 36$) and *Kirkianella* ($2n = 90, 126$), are sister genera, and together they form the sister clade to *Sonchus arvensis*-*S. maritimus*. This implies that these Pacific genera were derived either from section *Arvenses* or *Maritimi* of subgenus *Sonchus*, rather than from section *Apseri*.

Bremer (1993, 1994) proposed a revised subtribal classification of the Lactuceae in which he recognized a new subtribe Sonchinae from the *Launaea*-*Sonchus* group of Stebbins (1953) (Table 1). He also divided the subtribe Crepidinae of Stebbins (1953) into the four subtribes, i.e., Crepidinae s. str., Hieraciinae, Lactucinae, and Sonchinae (Table 1). Bremer (1993, 1994) suggested that *Launaea* and *Sonchus* are closely related and that the entire *Launaea*-*Sonchus* clade is monophyletic. In contrast, Jeffrey (1966) separated *Sonchus* and *Launaea* into two groups, i.e., the *Crepis* L. subgroup and *Sonchus* group (Table 1).

The Sonchinae as recognized by Bremer (1993) comprises 11 genera and approximately 130 species, and consists of two alliances: 1) *Launaea* and the related genera *Aetheorhiza* and *Reichardia*, and 2) *Sonchus* and a group of closely related genera, i.e., *Actites* Lander, *Babcockia*, *Embergeria*, *Kirkianella*, *Lactucosonchus*, *Sventenia*, and *Taeckholmia*. Most species occur in the Mediterranean region, Africa, and the Canary Islands, while several genera are found in the Pacific islands. Several genera of the *Sonchus* group, including *Babcockia*, *Lactucosonchus*, *Sventenia*, and *Taeckholmia*, are restricted to the Canary Islands. Two monotypic genera, *Kirkianella*

and *Embergeria*, are endemic to New Zealand, and *Actites* occurs exclusively in Australia. In contrast, *Sonchus* and *Launaea* are widely distributed in Africa, Macaronesia, and Eurasia. The 54 species of *Sonchus* are divided into three subgenera (Boulos 1972): *Dendrosonchus* (19 spp.); *Origosonchus* (14 spp.), and *Sonchus* (21 spp.). Subgenus *Dendrosonchus* consists of woody plants endemic to Macaronesia, subg. *Origosonchus* occurs exclusively in Africa, and subg. *Sonchus* includes widely distributed weedy species. The monotypic genus *Aetheorhiza* is distributed locally throughout the Mediterranean and western Europe (Babcock and Stebbins 1943). *Reichardia*, which is composed of eight species, occurs in the Mediterranean and Europe (Gallego et al. 1980).

In Jeffrey's (1966) informal classification of the Lactuceae, *Dendroseris* and *Thamnosseris* Phil., which are endemic to the Juan Fernandez and San Ambrosio Islands, respectively, were placed with *Sonchus* in his *Sonchus* group (Table 1). Stebbins (1953), however, placed *Dendroseris* in its own subtribe (Dendroseridinae) and *Thamnosseris* in the subtribe Stephanomeriinae. Bremer (1994) retained the Dendroseridinae with the addition of *Thamno-*

TABLE 1. Subtribal classifications of the Lactuceae.

Jeffrey (1966)	Stebbins (1953)	Bremer (1993, 1994)
<i>Scolymus</i> group	Scolyminae	<i>Scolymus</i>
<i>Tolpis</i> group	Cichoriinae	<i>Cichorium</i>
<i>Tolpis</i> subgroup	Microseridinae	Catananchinae
<i>Catananche</i> subgroup	Stephanomeridiinae	Microseridinae
<i>Stephanomeria</i> subgroup	<i>Thamnoseric</i>	Stephanomeridiinae
<i>Microseris</i> subgroup		Malacothricinae
<i>Sonchus</i> group	Dendroseridinae	Dendroseridinae
<i>Sonchus</i> :	<i>Dendroseris</i>	<i>Dendroseris</i>
Subg. <i>Sonchus</i>		<i>Thamnoseric</i>
Subg. <i>Origosonchus</i>		
Subg. <i>Dendrosonchus</i>		
<i>Thamnoseric</i> , <i>Dendroseris</i>		
<i>Hypochaeris</i> group	Hypochaeridinae	Hypochaeridinae
<i>Hypochaeris</i> subgroup	Scorzonerinae	Scorzonerinae
<i>Scorzonera</i> subgroup		
<i>Cichorium</i> group	Crepidinae	Crepidinae
<i>Cichorium</i> subgroup	<i>Dubyaea-Soroseric</i> line	Sonchinae: <i>Launaea</i> , <i>Sonchus</i> , <i>Reichardia</i> ,
<i>Crepis</i> subgroup	<i>Launaea-Sonchus</i> line: <i>Launaea</i> , <i>Son-</i>	<i>Aetheorhiza</i> , <i>Actites</i> , <i>Babcockia</i> ,
<i>Dubyaea</i> series	<i>chus</i> , <i>Reichardia</i> , <i>Aetheorhiza</i>	<i>Embergeria</i> , <i>Kirkianella</i> , <i>Lactucoson-</i>
<i>Prenanthes</i> series	<i>Hieracium-Andryala</i> line	<i>chus</i> , <i>Sventenia</i> , <i>Taechholmia</i>
<i>Crepis</i> series	<i>Prenanthes-Lactuca</i> line	
<i>Taraxacum</i> series	<i>Youngia-Ixeris</i> line	Hieraciinae
<i>Chondrilla</i> series	<i>Crepis</i> line	Lactucinae
<i>Launaea</i> series: <i>Launaea</i> ,		
<i>Reichardia</i> , <i>Aetheorhiza</i>		

seris. Furthermore, he indicated uncertainty with regard to the position of the two genera in Lactuceae (see also Sanders et al. 1987).

Chloroplast DNA (cpDNA) restriction site data provided several insights into phylogenetic relationships among genera of the Lactuceae (Whitton et al. 1995). For example, the Crepidinae of Stebbins (1953) was suggested to be highly polyphyletic, while *Dendroseris*, and thus his subtribe Dendroseridinae, appeared monophyletic. Monophyly of *Dendroseris* is concordant with the results of Sang et al. (1994). The cpDNA restriction site data of Whitton et al. (1995) also suggested that the Sonchinae of Bremer (1994) was paraphyletic because *Dendroseris* was nested within it. The cpDNA data indicated that *Dendroseris* was closely related to *Sonchus* and *Sventenia*. Restriction site data implicated *Reichardia* as the sister group to a monophyletic clade of *Dendroseris*, *Sonchus*, *Sventenia*, and *Aetheorhiza*. Although the cpDNA data of Whitton et al. (1995) resolved several phylogenetic issues, broader sampling within the Sonchinae is needed to assess relationships more rigorously.

The internal transcribed spacers (ITS) of nuclear ribosomal DNA (nrDNA) have proven useful for elucidating phylogenetic relationships among con-

generic species and closely related genera (see Baldwin et al. 1995). The primary purpose of this study was to use ITS sequences to assess phylogenetic relationships among genera within subtribes Sonchinae and Dendroseridinae. More specifically, we wished to: 1) assess the monophyly of Sonchinae; 2) test the monophyly of several genera within this subtribe, and 3) determine phylogenetic relationships among genera of the Sonchinae, especially of *Dendroseris* and other small Pacific genera, with the largely Macaronesian, African, and Mediterranean *Sonchus*.

MATERIALS AND METHODS

Complete sequences of the ITS region were generated for 36 accessions representing 11 genera and 31 species of subtribe Sonchinae (Bremer 1994), three species of *Prenanthes*, one species of *Taraxacum*, and one species of *Lactuca* (Table 2). Sequences from nine species of *Dendroseris* and the four outgroup genera *Krigia*, *Pyrrhopappus*, *Microseris*, and *Lactuca*, were obtained from Sang et al. (1994) and Kim and Jansen (1994), respectively. Four genera, *Krigia*, *Microseris*, *Pyrrhopappus*, and *Lactuca*, were chosen as outgroups based on the

TABLE 2. Sources of plants for ITS sequences. Voucher specimens are deposited in Ohio State University Herbarium (OS) and the Herbarium of University of Texas (TEX). ¶All sequences are deposited in GenBank; *sequences obtained from Sang et al. (1994); §sequences obtained from Kim and Jansen (1994).

Taxa	Voucher	¶Accession numbers (ITS 1, ITS 2)
<i>Aetheorhiza</i> Cass.		
<i>A. bulbosa</i> (L.) Cass.	Jansen 1105 (TEX)	L48135, L48136
<i>Babcockia</i> Boulos		
<i>B. platylepis</i> (Webb) Boulos	Kim S.-C. et al. 1028 (OS)	L48137, L48138
<i>Dendroseris</i> D. Don.*		
<i>D. litoralis</i> Skottsbo.	Stuessy et al. 11973 (OS)	L49508, L49509
<i>D. marginata</i> (Bert. & Dcne.) Hook. & Arn.	Stuessy et al. 11999 (OS)	L49510, L49511
<i>D. macrantha</i> (Bert. & Dcne.) Skottsbo.	Stuessy et al. 5149 (OS)	L49512, L49513
<i>D. micrantha</i> Hook. & Arn.	Stuessy et al. 11582 (OS)	L49514, L49515
<i>D. pruinata</i> (Johow) Skottsbo.	Stuessy et al. 11349 (OS)	L49516, L49517
<i>D. nerifolia</i> Hook. & Arn.	Stuessy et al. 11534 (OS)	L49518, L49519
<i>D. pinnata</i> (Bert. & Dcne.) Hook. & Arn.	Stuessy et al. 11334 (OS)	L49520, L49521
<i>D. berteriana</i> (Dcne.) Hook. & Arn.	Stuessy et al. 11589 (OS)	L49522, L49523
<i>D. regia</i> Skottsbo.	Landerø & Ruiz 9316 (OS)	L49524, L49525
<i>Embergeria</i> Boulos		
<i>E. grandifolia</i> (T. Kirk) Boulos	Atkinson 118/85 (OS)	L48139, L48140
<i>Kirkianella</i> Allan		
<i>K. novae-zelandiae</i> (Hook. f.) Allan	D. Glenney 4910 (OS)	L48141, L48142
<i>Krigia</i> Schreber.§		
<i>K. montana</i> (Michx.) Nutt.	Kim K.-J. 10141 (TEX)	L13946
<i>Lactuca</i> L.		
<i>L. perennis</i> L.	Bonn Bot. Gard. s.n.	L48143, L48144
<i>L. sativa</i> L.§	No Voucher (cultivated)	L13957
<i>Lactucosonchus</i> (Sch. Bip.) Svent.		
<i>L. webbii</i> (Sch. Bip.) Svent.	Kim S.-C. et al. 1033 (OS)	L48159, L48160
<i>Launaea</i> Cass.		
<i>L. arborescens</i> (Batt.) Murb.	Kim S.-C. et al. 1040 (OS)	L48145, L48146
<i>L. nudicaulis</i> (L.) Hook. f.	Kim S.-C. et al. 1053 (OS)	L48147, L48148
<i>Microseris</i> D. Don§		
<i>M. laciniata</i> Sch.-Bip.	Chambers 5369 (OSC)	L13954
<i>Prenanthes</i> L.		
<i>P. altissima</i> L.	Mehrhoff s.n. (TEX)	L48149, L48150
<i>P. pendula</i> Sch. Bip.	Kim S.-C. et al. 1051 (OS)	L48155, L48156
	Kim S.-C. et al. 1052 (OS)	L48157, L48158
	Kim S.-C. 1049 (OS)	L48151, L48152
<i>Pyrrhopappus</i> DC.§		
<i>P. grandifolus</i> Nutt.	Kim K.-J. 10508 (TEX)	L13953
<i>Reichardia</i> Roth		
<i>R. picroides</i> (L.) Roth	Belgium Bot. Gard. 2871	L48153, L48154
<i>R. tinginata</i> (L.) Roth	KEW 223-70-02 090	L48163, L48164
<i>R. ligulata</i> (Vent.) Kunkel & Sunding	Kim S.-C. et al. 1044 (OS)	L48165, L48166
<i>Sonchus</i> L.		
Subg. <i>Dendrosonchus</i> Sch. Bip. ex Boulos		
<i>S. canariensis</i> (Sch. Bip.) Boulos	Kim S.-C. et al. 1021 (OS)	L48291, L48292
<i>S. congestus</i> Willd.	Kim S.-C. et al. 1000 (OS)	L48173, L48174
<i>S. fruticosus</i> L. Fil.	Kim S.-C. et al. 1046 (OS)	L48125, L48126
<i>S. gonzalezpadroni</i> Svent.	Kim S.-C. et al. 1037 (OS)	L48127, L48128
<i>S. ortunoii</i> Svent.	Kim S.-C. et al. 1036 (OS)	L48129, L48130
Subg. <i>Origosonchus</i> Boulos		
<i>S. schweinfurthii</i> Oliv. et Hiern	Knox 2560 (OS)	L48295, L48296
<i>S. luxurians</i> (R. E. Fries) C. Jeffrey	Knox 2559 (OS)	L48297, L48298
Subg. <i>Sonchus</i>		
<i>S. kirkii</i> (T. Kirk) Allan	Silbury s.n. (OS)	L48299, L48300
<i>S. asper</i> L. Hill	Jansen 1109 (TEX)	L48301, L48302
<i>S. oleraceus</i> L.	I. Regk s.n. (OS)	L48303, L48304
<i>S. bourgeauii</i> Sch. Bip.	Kim S.-C. 1035 (OS)	L48305, L48306
<i>S. arvensis</i> L.	Jansen 1103 (TEX)	L48307, L48308
<i>S. maritimus</i> L.	L. Vilar s.n. (OS)	L48309, L48310
<i>S. palustris</i> L.	Kim S.-C. 1050 (OS)	L48311, L48312
<i>S. tuberosus</i> Svent.	Kim S.-C. et al. 1045 (OS)	L48313, L48314
<i>Sventenia</i> Font Quer		
<i>S. bupleuroides</i> Font Quer	Kim S.-C. et al. 1041 (OS)	L48315, L48316
<i>Taeckholmia</i> Boulos		
<i>T. pinnata</i> (L. Fil.) Boulos	Kim S.-C. et al. 1006 (OS)	L48319, L48320
<i>T. canariensis</i> Boulos	Kim S.-C. et al. 1043 (OS)	L48323, L48324
<i>T. heterophylla</i> Boulos	Kim S.-C. et al. 1037 (OS)	L48333, L48334
<i>T. aborea</i> (DC.) Boulos	Kim S.-C. et al. 1047 (OS)	L48325, L48326
<i>Taraxacum</i> Weber		
<i>T. officinale</i> Weber	Jansen 1107 (TEX)	L48337, L48338

cpDNA restriction site data of Whitton et al. (1995) and the advice of Charles Jeffrey (Royal Botanical Gardens, Kew) and Kåre Bremer (Uppsala University). Two additional species of *Prenanthes* were also included in this study to assess the phylogenetic relationships of the Canary Island endemic *P. pendula* (see Discussion). No material was available for *Actites* (subtribe Sonchinae) from Australia and *Thamnosseris* (subtribe Dendroseridinae) from the San Ambrosio Islands in the South Pacific off the coast of northern Chile.

Total genomic DNA was isolated from leaf tissue using the CTAB method of Doyle and Doyle (1987), and purified further by ultracentrifugation with CsCl/ethidium bromide gradients (Sambrook et al. 1989). Double-stranded DNA's of the entire ITS regions including the 5.8S coding region were amplified directly by symmetric PCR using universal primers (White et al. 1990). The initial PCR reaction was 3 min at 95°C for denaturation, 1 min at 50°C for annealing, and 1 min at 72°C for primer extension. The next 30 cycles consisted of 1 min at 95°C, 1 min at 50°C, and 45 sec at 72°C. A final extension of 5 min at 72°C was then performed. PCR products were purified by agarose gel (1X TAE buffer) electrophoresis and the concentrated DNA's were recovered using glass powder (U.S. Bioclean, U.S. Biochemical).

Double-stranded PCR products were directly sequenced using the Sequenase Version 2.0 (United States Biochemical Corp.) dideoxychain-termination method, employing two forward (ITS3 and ITS5) and two reverse (ITS2 and ITS4) primers (White et al. 1990). Modifications to the Sequenase protocol included denaturation of the double-stranded DNA by boiling the DNA/primer mix for 3 min, followed by snap-chilling the annealing mixture for 7 min in an ice water bath. In addition, 1 µl DMSO was added to both the labeling and termination reactions to reduce the effects of DNA secondary structure (Cosner et al. 1994).

DNA sequences were separated in 6% acrylamide gels using wedge-shaped spacers. Both short (3.5 hr at 1500 volts) and long (7.5 hr at 1500 volts) runs were conducted on the same gel, in order to read the entire ITS region. Gels were fixed for 30 min in 10% acetic acid, transferred to 3-MM Whatmann filter paper, dried under vacuum for 2.5 hr at 80°C, and exposed to Kodak XAR x-ray film for 12–72 hr.

The boundaries of the ITS and rDNA coding regions were identified by comparison to known sequences (Yokota et al. 1989; Ramon et al. 1990;

Baldwin 1992, 1993; Kim and Jansen 1994). Sequences were aligned initially using the Clustal W program (provided by D. Higgins). Several gap opening and extension penalties were used to align the entire sequences. The sequences aligned with Clustal W were then adjusted manually in order to align several regions of conserved sequences. Both the small size and number of indels made manual adjustments feasible.

Phylogenetic analyses using Fitch parsimony were performed employing PAUP (version 3.1.1; Swofford 1993) using the HEURISTIC search option with TBR branch swapping and MULPARS on. To search for multiple islands of trees (Maddison 1991), 100 replications of "random" taxa addition were conducted. Bootstrap analysis (Felsenstein 1985) was performed with 100 replications (maxtree = 100) to provide a measure of support for the clades. Decay analysis (Bremer 1988; Donoghue et al. 1992) was also performed to assess the robustness of the monophyletic groups. Trees up to five steps longer were examined.

RESULTS

Length Variation and Base Composition of the ITS Region. The length of ITS 1 varied from 233 bp to 255 bp, and ITS 2 varied from 220 bp to 226 bp. One 18 bp deletion in ITS 1 was detected in the monotypic genus *Aetheorrhiza*. The length of ITS 1 and ITS 2 of the Sonchinae is within the size range of other Asteraceae (Baldwin et al. 1995). The G+C content in ITS 1 varied from 42.0% (*Taeckholmia heterophylla*) to 48.2% (*Sonchus luxurians*), whereas ITS 2 varied from 46.8% (*Sonchus arvensis*) to 52.7% (*Sonchus palustris*).

No evidence of multiple rDNA repeat types in any of the taxa analyzed was observed. All of the double-stranded PCR products obtained appeared as sharp, single bands on 1.0% agarose gels. Furthermore, polymorphism at individual nucleotide sites was not commonly encountered (except for two taxa, *Taeckholmia heterophylla* and *Lactucosonchus*; Kim et al. 1996). Therefore, ITS sequence data in this study provide no evidence for different ITS length variants or major sequence variants within individual samples.

Phylogenetic Analyses. A total of 330 variable sites was found among all taxa examined, with 264 of them phylogenetically informative. There were 274 variable sites in the ingroup; 188 were cladistically informative. Only 64 of the 16,170 characters (0.4%) were scored as missing or ambiguous sites.

Unambiguous gap positions were treated as missing and they were not scored as separate characters in the phylogenetic analyses. The heuristic search identified 144 equally parsimonious trees with a length of 898, a consistency index (CI) of 0.580 (0.526 excluding uninformative changes), and a retention index (RI) of 0.743. The CI is close to the regression line of Sanderson and Donoghue (1989) for an analysis of 50 taxa.

The ITS tree (Fig. 1) suggests that *Reichardia* and *Launaea* are basal in the Sonchinae. The tree does not support the monophyly of the Sonchinae because *Dendroseris* and *Prenanthes pendula* are nested within the subtribe. The monophyly of the woody members of *Sonchus* and their close relatives in Macaronesia (Fig. 1; clade A) is very strongly supported (100% bootstrap value and a decay index of >5). *Prenanthes pendula*, which is endemic to the Canary Islands, is clearly grouped within this clade. Bootstrap analysis also shows strong support ($>85\%$) for the monophyly of several genera and groups of genera, including *Reichardia*, *Dendroseris*, subg. *Origosonchus*, and *Embergeria-Kirkianella*. Several genera, such as *Launaea* and *Taeckholmia*, appear to be paraphyletic. Both *Prenanthes* and *Sonchus* are polyphyletic.

DISCUSSION

Phylogenetic Relationships in the Sonchinae.

Phylogenetic analysis of ITS sequences provides several insights into evolutionary relationships within the Sonchinae (Fig. 1). The subtribe as delimited by Bremer (1993, 1994) is not monophyletic. The ITS data instead provide support for Jeffrey's view that *Dendroseris* is closely related to *Sonchus* and its relatives (Jeffrey 1966; Table 1). In the ITS tree, *Dendroseris* is nested within subtribe Sonchinae, which indicates that the Dendroseridinae of Stebbins (1953) should not be recognized as a distinct subtribe. This relationship is also concordant with cpDNA restriction site data (Whitton et al. 1995). The *Launaea-Sonchus* line (Table 1) recognized by Stebbins (1953) is not supported by the ITS data. The two major alliances of Bremer within the Sonchinae, i.e., *Launaea* with related genera, and *Sonchus* and a group of closely related genera, are not distinguished in the ITS phylogeny because *Aetheorhiza* is closer to certain members of *Sonchus* than to *Launaea* and *Reichardia* (Fig. 1). However, the two genera *Reichardia* and *Launaea* are not part of the main radiation of Sonchinae, which is

congruent with cpDNA restriction site data (Whitton et al. 1995).

The ITS sequence data also offer the opportunity to examine the monophyly of several genera. The ITS tree (Fig. 1) supports the monophyly of both *Reichardia* and *Dendroseris*, while *Taeckholmia* and *Launaea* appear paraphyletic. However, the paraphyly of *Launaea* (ca. 50 species) needs to be examined further with wider taxonomic sampling. In addition, the ITS phylogeny provides strong evidence that *Sonchus* and *Prenanthes* are polyphyletic. Within *Sonchus*, subg. *Origosonchus* is monophyletic, while the two subgenera *Dendrosonchus* and *Sonchus* are not.

Sonchus and Related Genera in Macaronesia.

Boulos (1972) recognized three subgenera in *Sonchus*. He suggested that all three subgenera are natural groups and proposed that subg. *Dendrosonchus* in Macaronesia evolved from subg. *Origosonchus* in Africa, which he considered the most primitive in the genus (Boulos 1967). Furthermore, he hypothesized that subg. *Sonchus* was derived from woody members of *Sonchus*, i.e., subg. *Dendrosonchus*, in Macaronesia. Neither monophyly of each of the three subgenera nor the phylogenetic relationships among them as proposed by Boulos (1967, 1972) is supported by the ITS data. For example, subg. *Dendrosonchus* is closer to several genera from Macaronesia than to either subg. *Sonchus* or subg. *Origosonchus* (Fig. 1; clade A). In addition, subg. *Dendrosonchus* is paraphyletic and members of subg. *Sonchus* were derived independently several times.

The ITS phylogeny strongly supports the monophyly of subg. *Dendrosonchus* and its close relatives, including the herbaceous tuberous perennial *Sonchus tuberifer* and the woody *Prenanthes pendula*, all of which occur in Macaronesia (Fig. 1; clade A). This suggests that members of subg. *Dendrosonchus*, along with *Sventenia*, *Babcockia*, *Taeckholmia*, *P. pendula*, *S. tuberifer*, and *Lactucosonchus*, were derived from a common ancestor in the Macaronesian Islands (Kim et al. 1996). Within this clade, *Lactucosonchus*, which is monotypic and endemic to La Palma in the Canaries, is sister to the rest of the taxa. The ITS tree also suggests that *Sonchus palustris*, a member of subg. *Sonchus* that occurs widely in Europe, is the sister taxon to the Macaronesian clade (Fig. 1). This relationship, however, is weakly supported.

The Juan Fernandez Islands Endemic *Dendroseris*. Our results, like those of Sang et al. (1994) and Whitton et al. (1995), support the monophyly

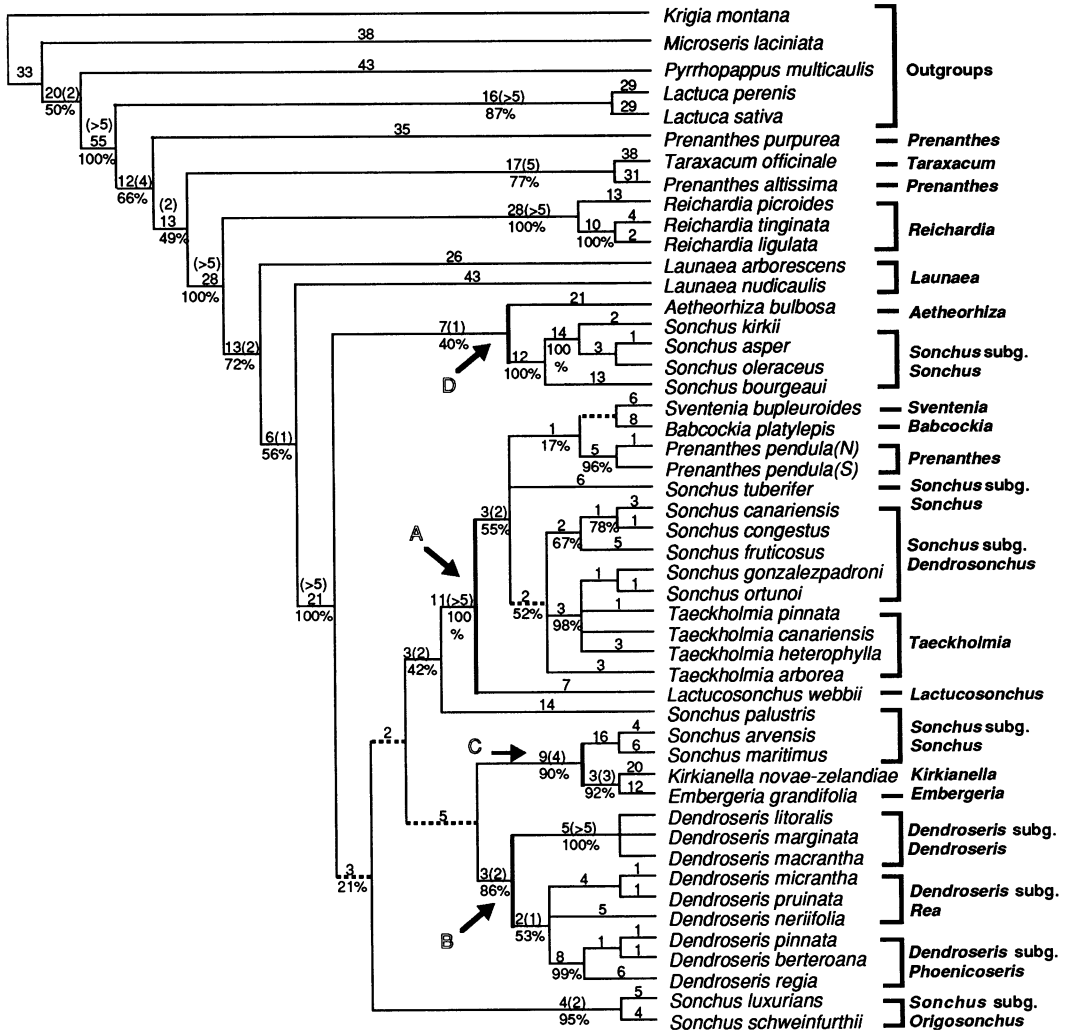


FIG. 1. ITS sequence phylogeny of subtribe Sonchinae. This tree is one of the 144 equally most parsimonious trees with a length of 898. Consistency index = 0.526 (excluding autapomorphies). Retention index = 0.743. Dashed lines indicate branches that collapse in the strict consensus tree. Numbers above lines represent the number of base substitutions followed by the decay values in parentheses. The bootstrap support (%) is shown below the nodes.

of *Dendroseris* (Fig. 1; clade B), and further suggest that, while two subgenera (*Dendroseris* and *Phoenicoseris*) are monophyletic, subg. *Rea* is paraphyletic. The subg. *Dendroseris* is weakly supported as sister to the remainder of the genus. The relationship of *Dendroseris* to other genera of Sonchinae remains uncertain beyond being part of the large clade that contains everything except *Launaea* and *Reichardia*.

The Pacific Genera *Embergeria* and *Kirkianella*.

The monotypic genus *Embergeria*, a tetraploid ($2n = 36$), is endemic to the Chatham Islands of

New Zealand, and has been viewed as a survivor from Pleistocene glaciation (Wardle 1963; Lander 1976; Webb et al. 1988). Lander (1976) hypothesized that during the late Pliocene subg. *Sonchus* migrated to New Zealand where *Embergeria* subsequently originated. Cytological, palynological, and morphological data suggested subg. *Sonchus*, especially sect. *Asperi* (*S. kirkii*), as the most probable ancestor of *Embergeria* (Boulos 1967; Pons and Boulos 1972; Roux and Boulos 1972). The monotypic genus *Kirkianella*, a pentaploid ($2n = 90$) and

heptaploid ($2n = 129$) (Beuzenberg and Hair 1984), is also endemic to New Zealand and is morphologically variable. This genus has been considered closely related to either *Sonchus* or *Launaea* (Allan 1961). Glenny (pers. comm.) also proposed a close relationship to *Youngia* Cass. and *Crepis* L., genera of Lactuceae outside Sonchinae (not sampled in this study). In the ITS tree, the two New Zealand endemics are sister to each other and form a monophyletic group with two species of *Sonchus* subg. *Sonchus* (Fig. 1; clade C). The presence of this strongly supported clade is somewhat unexpected because *Sonchus kirkii* has been considered ancestral to *Embergeria*, yet it is sister to the clade of *S. asper* and *S. oleraceus* (Fig. 1; clade D). Therefore, the ITS phylogeny does not support previous hypotheses about *Embergeria*. Rather, it seems more likely that *Embergeria* and *Kirkianella* are most closely related to *S. arvensis* (sect. *Arvenses*) and *S. maritimus* (sect. *Maritimi*).

Aetheorhiza. Babcock and Stebbins (1943) suggested that *Aetheorhiza* of the Mediterranean and western Europe is closely related to the *Launaea* and *Sonchus* groups. *Aetheorhiza* shares several features with *Launaea*, including four broad, round-ribbed, and obcompressed achenes, long corolla tubes with conspicuous white pubescence on their upper parts as well as on the base of the ligule, and a mixture of relatively broad and narrow setae in the pappus. Several other features of *Aetheorhiza*, especially those of the corolla and involucre, are shared with *Sonchus*. The distinctive habit of *Aetheorhiza* was the sole basis for Babcock and Stebbins' (1943) recognition of the genus; new rosettes of leaves form along elongated rhizomes in *Aetheorhiza*. The ITS tree shows that *Aetheorhiza* is clearly part of the lineage that includes no species of *Launaea* and all species of *Sonchus* (Fig. 1). It seems likely that *Aetheorhiza* diverged as a part of the radiation of the *Sonchus* group, which presumably occurred after the elements classified as *Launaea* had already diverged.

Prenanthes pendula. The phylogenetic relationship of the Canary Island endemic *P. pendula* to other species of *Prenanthes* is problematic. It is the only species of the genus from the Canary Islands that is locally common in mountain cliffs on the south and north sides of Gran Canaria. It has woody stems, small heads (about 2 mm across) with 5-6 yellow florets, and simple pappus hairs (Bramwell and Bramwell 1974). A close relationship between *P. pendula* and African *Prenanthes* species was suggested by Bramwell (1985). In the ITS tree

P. pendula is in a clade that is strongly supported by high bootstrap (100%) and decay values (>5) and that includes only plants from Macaronesia (Fig. 1; clade A). This finding is concordant with the results from Perez de Paz (1976), who suggested that *P. pendula* is closely related to *Sventenia* and *Sonchus*. Also, a naturally occurring intergeneric hybrid between *Sventenia bupleuroides* and *P. pendula* was reported by Sventenius (1960). Therefore, the ITS tree agrees with other results in suggesting a close relationship between *P. pendula* and woody *Sonchus* in Macaronesia. The ITS tree further suggests that *P. pendula* is not closely related to other species of *Prenanthes*. For example, *P. altissima* and *Taraxacum officinale* are sister taxa, and *P. purpurea* is a sister taxon to the clade that includes those taxa and the remainder of the subtribe. Thus, the ITS phylogeny suggests strongly that *Prenanthes* may be polyphyletic.

Although ITS sequence data have been useful in proposing phylogenetic hypotheses at and above the generic level in the subtribe Sonchinae, five issues remain unresolved: 1) relationships between subg. *Dendrosonchus* and its close relatives in the Macaronesian islands; 2) elucidation of the position of *Prenanthes pendula*; 3) confirmation of the paraphyly of *Launaea* and its relationships to other genera; 4) relationships of *Dendroseris* and *Thamnosseris* to each other and to other Sonchinae, and 5) investigation of the relationship of *Actites* (endemic to Australia) to other genera, especially *Embergeria*. Other data, especially from cpDNA, would provide additional data to resolve these issues.

ACKNOWLEDGMENTS. We are grateful to Javier Francisco-Ortega, Arnoldo Santos-Guerra, Aguedo Marrero, Pedro Ortega-Machín, and Francisco Jose González Artilles for assistance in the field. We also thank Eric Knox, Tom Myers, David Glenny, P. J. Garnock-Jones, Loutfy Boulos, J. L. S. Kessing, and Joan Pedrola i Monfort for providing plant and DNA materials. We are indebted to Charles Jeffrey and Kåre Bremer for their helpful suggestions about outgroups. We are extremely grateful to Jeff Doyle for his comments and suggestions during review of the manuscript. This paper represents a portion of a doctoral dissertation by S.-C. K. to The Ohio State University. This work was supported by American Society of Plant Taxonomists Grants for Graduate Student Research, Tinker Foundation (Latin American Studies Program, Ohio State University), Beatley Herbarium Award (Ohio State University), and Sigma Xi Grants-in-Aid of Research to S.-C. K., National Science Foundation Doctoral Disserta-

tion Improvement Grant DEB-9521017 to S.-C. K. and D. J. C., and NSF grant BSR 87-08246 to R. K. J. This study was also partially supported by the Centro de Investigación y Tecnología Agrarias de Canarias through the advice of Manuel Fernández-Galván. We dedicate this paper to the memory of our friend and colleague Mary Elizabeth Cosner.

LITERATURE CITED

- ALLAN, H. H. 1961. *Flora of New Zealand*. vol.1. Wellington, New Zealand: Government Printer.
- BABCOCK, E. B. and G. L. STEBBINS, JR. 1943. Systematic studies in the Cichorieae. University of California Publications in Botany 18: 227–240.
- BALDWIN, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. *Molecular Phylogenetics and Evolution* 1: 3–16.
- . 1993. Molecular phylogenetics of *Calycadenia* (Compositae) based on ITS sequences of nuclear ribosomal DNA: Chromosomal and morphological evolution reexamined. *American Journal of Botany* 80: 222–238.
- , C. S. CAMPBELL, J. M. PORTER, M. J. SANDERSON, M. F. WOJCIECHOWSKI, and M. J. DONOGHUE. 1995. Utility of nuclear ribosomal DNA internal transcribed spacer sequences in phylogenetic analyses of angiosperms. *Annals of the Missouri Botanical Garden* 82: 247–277.
- BEUZENBERG, E. J. and J. B. HAIR. 1984. Contributions to a chromosome atlas of the New Zealand flora – 27 Compositae. *New Zealand Journal of Botany* 22: 353–356.
- BOULOS, L. 1967. *Taeckholmia*, a new genus of Compositae from Canary Islands. *Botaniska Notiser* 120: 95–108.
- . 1972. Révision systématique du genre *Sonchus* L. s.l. I. Introduction et classification. *Botaniska Notiser* 125: 287–305.
- BRAMWELL, D. 1985. Contribución a la biogeografía de las islas Canarias. *Botanica Macaronésica* 14: 3–34.
- and Z. BRAMWELL. 1974. *Wild flowers of the Canary Islands*. London: Stanley Thornes (Publishers) Ltd.
- BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.
- . 1993. New subtribes of the Lactuceae. *Novon* 3: 328–330.
- . 1994. *Asteraceae—Cladistics and classification*. Portland, Oregon: Timber Press.
- COSNER, M. B., R. K. JANSEN, and T. G. LAMMERS. 1994. Phylogenetic relationships in the Campanulales based on *rbcl* sequences. *Plant Systematics and Evolution* 190: 79–95.
- DONOGHUE, M. J., R. G. OLMSTEAD, J. F. SMITH, and J. D. PALMER. 1992. Phylogenetic relationships of Dipsacales based on *rbcl* sequences. *Annals of the Missouri Botanical Garden* 79: 333–345.
- DOYLE, J. J. and J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf material. *Phytochemical Bulletin* 19: 11–15.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783–791.
- GALLEGO, M. J., S. TALAVERA, and S. SILVESTRE. 1980. Revisión del genero *Reichardia* Roth (Compositae). *Lagasalia* 9: 159–218.
- JEFFREY, C. 1966. Notes on Compositae, I. The Cichorieae in East Tropical Africa. *Kew Bulletin* 18: 427–486.
- KIM, K.-J. and R. K. JANSEN. 1994. Comparisons of phylogenetic hypotheses among different data sets in dwarf dandelions (*Krigia*): additional information from internal transcribed spacer sequences of nuclear ribosomal DNA. *Plant Systematics and Evolution* 190: 157–185.
- KIM, S.-C., D. J. CRAWFORD, J. FRANCISCO-ORTEGA, and A. SANTOS-GUERRA. 1996. A common origin for woody *Sonchus* and five related genera in the Macaronesian Islands: Molecular evidence for extensive radiation. *Proceedings of the National Academy of Sciences U.S.A.* 93: 7743–7748.
- LANDER, N. S. 1976. *Actites*, a new genus of Compositae from Australia. *Telopea* 1: 129–135.
- MADDISON, D. R. 1991. The discovery and importance of multiple islands of most-parsimonious trees. *Systematic Zoology* 40: 315–328.
- PÉREZ DE PAZ, J. 1976. Observaciones sobre la biología y relaciones de *Sventenia bupleuroides* F. Q. *Botanica Macaronésica* 1: 51–65.
- PONS, A. and L. BOULOS. 1972. Révision systématique du genre *Sonchus* L. s.l. III. Étude palynologique. *Botaniska Notiser* 125: 310–319.
- RAMON, A. T., M. GANAL, and V. HEMLEBEN. 1990. GC balance in the internal transcribed spacers ITS 1 and ITS 2 of nuclear ribosomal RNA genes. *Journal of Molecular Evolution* 30: 170–181.
- ROUX, J. and L. BOULOS. 1972. Révision systématique du genre *Sonchus* L. s.l. II. Étude caryologique. *Botaniska Notiser* 125: 306–309.
- SAMBROOK, J., E. F. FRITSCH, and R. MANIATIS. 1989. *Molecular cloning, a laboratory manual*, Second Edition. Plainview, New York: Cold Spring Harbor Laboratory Press.
- SANDERS, R. W., T. F. STUESSY, C. MARTICORENA, and M. SILVA O. 1987. Phytogeography and evolution of *Dendroseris* and *Robinsonia*, tree-Compositae of the Juan Fernandez Islands. *Oper Botanica* 92: 195–215.
- SANDERSON, M. J. and M. J. DONOGHUE. 1989. Patterns of variation in levels of homoplasy. *Evolution* 43: 1781–1795.
- SANG, T., D. J. CRAWFORD, S.-C. KIM, and T. F. STUESSY. 1994. Radiation of the endemic genus *Dendroseris* (Asteraceae) on the Juan Fernandez Islands: Evidence from sequences of the ITS regions of nuclear ribosomal DNA. *American Journal of Botany* 8: 1494–1501.

- STEBBINS, G. L. JR. 1953. A new classification of the tribe Cichorieae, family Compositae. *Madroño* 12: 65–81.
- SVENTENIUS, E. R. S. 1960. *Additamentum ad floram Canariensem*. Madrid, Spain: Instituto Nacional de Investigaciones Agronómicas Ministerio de Agricultura.
- SWOFFORD, D. L. 1993. *PAUP: Phylogenetic analysis using parsimony*, version 3.1.1. Champaign: Illinois Natural History Survey.
- WARDLE, P. 1963. Evolution and distribution of the New Zealand flora, as affected by Quaternary climates. *New Zealand Journal of Botany* 1: 3–17.
- WEBB, C. J., W. R. SYKES, and P. J. GARNOCK-JONES. 1988. *Flora of New Zealand*. vol. IV. Christchurch: Botany Division DSIR.
- WHITTON, J., R. S. WALLACE, and R. K. JANSEN. 1995. Phylogenetic relationships and patterns of character change in the tribe Lactuceae (Asteraceae) based on chloroplast DNA restriction site variation. *Canadian Journal of Botany* 73: 1058–1073.
- WHITE, T. J., T. BRUNS, S. LEE, and J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315–322 in *PCR protocols: A guide to methods and amplification*, eds. M. Innis, D. Gelfand, J. Sninsky, and T. White. San Diego: Academic Press.
- YOKOTA, Y., T. KAWATA, Y. IIDA, A. KATO, and S. TANIFUJI. 1989. Nucleotide sequences of the 5.8s rRNA gene and internal transcribed spacer regions in carrot and broad bean ribosomal DNA. *Journal of Molecular Evolution* 29: 294–301.

APPENDIX 1. Aligned ITS sequences from the subtribe Sonchinae and four outgroup genera. A dash represents a gap. Ambiguous data are indicated by "?". Polymorphic sites are represented by following symbols: R = A/G, Y = C/T, S = C/G, V = A/C/G. Species abbreviations are: *Krigia montana* (KRIG); *Pyrrhopappus multicaulis* (PYRR); *Microseris laciniata* (MICR); *Taraxacum officinale* (TARA); *Prenanthes altissima* (PRAL); *P. purpurea* (PRPU); *Lactuca perenis* (LAPE); *L. sativa* (LASA); *Reichardia picroides* (REPI); *R. tinginata* (RETI); *R. ligulata* (RELI); *Launaea arborescense* (LAAR); *L. nudicaulis* (LANU); *Aetheorhiza* (AETH); *Sventenia* (SVEN); *Lactucosonchus* (LACT); *P. pendula* (PRPE1); *P. pendula* (PRPE2); *Sonchus kirkii* (SOKI); *S. asper* (SOAS); *S. bourgeauii* (SOBO); *S. arvensis* (SOAR); *S. maritimus* (SOMA); *S. palustris* (SOPA); *S. oleraceus* (SOOL); *S. tuberosus* (SOTU); *S. luxurians* (SOLU); *S. schweinfurthii* (SOSC); *S. canariensis* (SOCA); *S. congestus* (SOCO); *S. fruticosus* (SOFR); *S. gonzalezpadroni* (SOGO); *S. ortunoii* (SOOR); *Babcockia* (BABC); *Taechholmia pinnata* (TAPI); *T. arborea* (TAAR); *T. canariensis* (TACA); *T. heterophylla* (TAHE); *Kirkianella* (KIRK); *Embergeria* (EMBE); *Dendroseris litoralis* (DELI); *D. marginata* (DEMA); *D. macrantha* (DEMC); *D. micrantha* (DEMI); *D. pruinata* (DEPR); *D. nerrifolia* (DENE); *D. pinnata* (DEPI); *D. berteroaana* (DEBE); *D. regia* (DERE).

	-ITS 1						
	10	20	30	40	50	60	70
KRIG	TCGAACCCCTG	CAAAGCAGA	GACGACCCGC	GAACCTGTAC	CCAT-AATCG	GGAGTCAGGG	ATA-TTGGCT
PYRR	TCGAACCCCTG	CAAAGC-GA	GACGACCCGC	GAACATGTAC	ATA-CAATCG	GGTTTGTATG	ATA-TTGACT
MICR	TCGAACCCCTG	CAAAGG-GA	GACTACCCGC	GAAGAGGTAC	CCA-AATATG	GGAGTTGGGG	ATA-TTGGCT
TARA	TCGAACCCCTG	CAAGG-CG	AACGACCTGT	GAACACGTAA	ATA-CAACTG	GGTGATGGGG	AGA-TGGATC
PRAL	TCGAACCCCTG	CAAGGC-AG	ACCGACCCGT	GAACACGTAA	ATA-CAACTT	GGTGACGGGG	AGA-TGGGCC
PRPU	TCGAACCCCTG	CAAGGC-AG	AACGACCTGT	GAACATGTAA	ATA-CAACCG	GGTGATGTGG	AGTC-GGGCC
LAPE	TCGAACCCCTG	CAA-GC-AG	AACGACCTGT	GAACATGTAA	ACA-CAACTG	GGTGACAGGG	AAA-TGGGAA
LASA	TCGAACCCCTG	CAA-GC-AG	AAACGCCCGT	GAACATGTAA	CCA-CAACGG	GGTGACCCGT	ATAA-GGGCC
REPI	TCGAACCCCTG	CAATGC-AG	AACGACCTGT	GAACATGTAA	AT-TCAACTC	GGTGTTGGTG	AAA-TGGGCC
RETI	TCGAACCCCTG	CAACGC-AG	AACGACCTGT	GAACATGTAA	AT-TCAACTC	GGTGTTGGTG	AAAA-TGGCC
RELI	TCGAACCCCTG	CAACGC-AG	AACGACCTGT	GAACATGTAA	AT-TCAACTC	GGTGCTGGTG	AAA-TGGCC
LAAR	TCGAACCCCTG	CAAGGC-AG	AACGACCCGT	GAACATGTAA	A-A-CAACTT	GGTGCTGGTG	AGA-TTGGCT
LANU	TCGAACCCCTG	CAAAG--	AACGACCTGT	GAACATGTAA	-T?CCA-CTT	GGTGCTGCTA	AGG-TGGGCT
AETH	TCGAACCC--	-----	-----	GA-C-TGTAA	-T--CA-GTT	GGTGCTGGTG	AGA-TGGGCC
SVEN	TCGAACCCCTG	CAAAGC-AG	AACGCCCTGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGACTGGGCC
LACT	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
PRPE1	TCGAACCCCTG	CAAAGC-?	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
PRPE2	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
SOKI	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	CAACATGTAA	ATA-CAAGTT	GGTGCTGGTG	TGA-TGGGCT
SOAS	TCGAACCCCTG	CAAAGG-?	AACGACCCGT	GAACATGTAA	AY?-CAAGTT	GGTGCTGGTG	TGA-TGGGCT
SOBO	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAAGTT	GGTGCTGGTG	AGA-TGGGCC
SOAR	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTT	GGTGCTGGTG	TGA-TGGGCC
SOMA	TCGAACCCCTG	TAAAGC-AG	AACGACCTGT	GAACATGTAA	ATA-CAACTT	GGTGCTGGTG	TGA-TGGGCC
SOPA	TCGAACCCCTG	CAAAGC-A	AACGACCTGT	GAACATGTAA	ATT-CAACTC	GGTGTTGGTT	AGA-TGGGCC
SOOL	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	AYA-CAAGTT	GGTGCTGGTG	TGA-TGGGCT
SOTU	TCGAACCCCTG	CAAAGC-?	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
SOLU	TCGAACC-TG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGACC
SOSC	TCGAACC-TG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	TTA-CAACTC	GGTGATGGTG	AGA-TGGGCC
SOCA	TCGAACCCCTG	CAAAGC-AG	AACGACCTGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
SOCO	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
SOFR	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
SOGO	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
SOOR	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
BABC	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGG?CC
TAPI	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATATCAACTC	GGTGTTGGTG	AGA-TGGGCC
TAAR	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	A?A-TG?GCC
TACA	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
TAHE	TCGAACCCCTG	CAAAGC-?	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGTTGGTG	AGA-TGGGCC
KIRK	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTT	GGTGCTGGTG	AGA-TTGGCC
EMBE	TCGAACCCCTG	CAAAGC-AG	AACGACCTGT	GAACATGTAA	ATA-CAACTT	GGTGCTGGTG	AGA-TGGGCC
DELI	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGGCC
DEMA	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGGCC
DEMC	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGGCC
DEMI	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGGCC
DEPR	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGGCC
DENE	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-YAACTC	GGTGCTGGTG	AGA-TGGGCC
DEPI	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGGCC
DEBE	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGGCC
DERE	TCGAACCCCTG	CAAAGC-AG	AACGACCCGT	GAACATGTAA	ATA-CAACTC	GGTGCTGGTG	AGA-TGGGCC

APPENDIX 1. Continued

	160	170	180	190	200	210	220
KRIG	CATT--AACAA	A-CCCCG-CA	C-GCAATGTG	CCAAGAAAA	CAAAAAACTG	AGAAGGACGC	GTCC--AATT
PYRR	CCAT--AACAA	A-CCCCG-CA	C-GGAATGTG	CCAAGAAAA	C-GAAATATG	AGAAGGGCAT	GTCC--ATTA
MICR	CATT--AACAA	A-CCCCG-CA	C-GGACTGTG	CCAAGAAAA	TATTAAACTG	AGAAGGACGC	GTCC--AATA
TARA	TTTT--AACAA	A-CCCCGCA	C-GGCATGTG	CCAAGAAAA	C-AATAAACG	AGAAGGACTC	GACC--TGTT
PRAL	C-T--AACAA	A-CCCCGGCA	C--GAATGT-	C-AAGGAAAA	C-AAATAATG	AGAAGGACTC	GTC--TTGTT
PRPU	CAT--AACAA	AACCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-AAAA-TG	AGAAGGACTC	AAACC--GTG
LAPE	CAT--AACAA	A-CCCCGGCA	C--GCATGTS	--AAGAAAA	C-AAAAATG	AGAAGGACAC	TTAC--TGTA
LASA	CAT--AACAA	AACCCCGGCA	C-GCTATGTG	CCAAGAAAA	C-AAAA-TG	AGAAGGACAC	TACCA--GTT
REPI	T-T--AACAA	A-CCCCGGCA	C--GTATGT-	C-AAGAAAA	C-AAGAAATG	AGAAGGTATC	GACC--TGAT
RETI	T-T--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-AAAA-TG	AGAAGGTATC	GACGACTTGA
RELI	T-T--AACAA	A-CCCCGGCA	C--GCATGTG	C-AAGAAAA	C-AAAA-TG	AGAAGGTATC	GAC--TTGAA
LAAR	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	G-AAAAATG	AGAAGATATA	GACC--TGTT
LANU	AAT--AACAA	A-CCCCGGCA	C-GG-ATGT-	CGAAGAAAA	C-AAAACATT	GGAAGGTATC	GACC--TGTT
AETH	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAA	AGAAGGTATA	TACC--TYAT
SVEN	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
LACT	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
PRPE1	TAT--AACAA	A-CCCCGGCA	C--GVATGTG	C-AAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
PRPE2	TAT--AACAA	A-CCCCGGCA	C--GAATGTG	C-AAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
SOKI	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	TT-AAATAAA	AGATGCTATT	TAC--TTGAT
SOAS	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	TT-AAATAAA	AGATGCTATT	TAC--TTGAT
SOBO	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAA	AGATGCTATC	TAC--TTGAT
SOAR	CAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATACG	AGAAGGTATC	TAC--TTGAT
SOMA	CAT--AACAA	A-CCCCGGCA	C--GCATGTG	C-AAGAAAA	CTGAAATAAG	AGAAGGTGTC	TAC--TTGAT
SOPA	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	T-GAAATAAG	TGAAGGTATC	TAC--TTGAT
SOOL	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	TT-AAATAAA	AGATGCTATT	TAC--TTGAT
SOTU	TAT--AACAA	A-CCCCGGCA	CAGGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
SOLU	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGTAGGTATT	CAC--TTGAT
SOSC	TAA-TRACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGTAGGTATC	CAC--TTGAT
SOCA	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
SOCO	TAT--AACAA	A-CCCCGGCA	C--GCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
SOFR	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
SOGO	TA-C-AACAA	A-TCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
SOOR	TATC-AACAA	A-TCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
BABC	TAT--AACAA	A-CCCCGGCA	C--GCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
TAPI	TAT--AACAA	A-TCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
TAAR	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATA	AGAAGGTATC	TAC--TTGAT
TACA	TAT--AACAA	A-TCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAT?	AGAAGGTATC	TAC--TTGAT
TAHE	AAT--?CAA	A-TCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATATG	AGAAGGTATC	TAC--TTGAT
KIRK	TATATAACAA	A-CCCCGGCA	C--GCATGTC	G-AAGAAAA	A-AAAAAAG	AGAAGGTATC	TACC--TGAT
EMBE	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-AAAAATG	AGAAGGTATC	TAC--TTGAT
DELI	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGAAGGTATC	TAC--TTCAT
DEMA	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGAAGGTATC	TAC--TTCAT
DEMC	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGAAGGTATC	TAC--TTCAT
DEMI	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGAAGGTATC	TAC--TTCAT
DEPR	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGAAGGTATC	TAC--TTCAT
DENE	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-TAAATAAG	AGAAGGTATC	TAC--TTCAT
DEPI	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-?AAATAAG	AGAAGGTATC	TTC--TTCAT
DEBE	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGAAGGTATC	TAC--TTCAT
DERE	TAT--AACAA	A-CCCCGGCA	C-GGCATGTG	CCAAGAAAA	C-GAAATAAG	AGAAGGTATC	TAC--TTCAT

						→ITS 2			
230	240	250	260	270	280	290	300		
TCGCCCCG-T	ACGCGGTGTG	CGTG-TTGTG	TGACCTCC-T	CCTTGAAATC	ACATC-RC--	TCGCCCCC-A	ACCAAGCATC		
TCGCCCCG-T	TCGCGGTGTG	CTTG--TCGC	CGT-GTCC-T	CCTTGAAATC	ACATC-AC--	TTGCCCCCCA	-CCATGCATC		
ATGCCCCG-T	TTGTGG?GTG	CATTC-TGAG	CGT-GTCCTC	CCTTG-AATC	ACATC-GC--	TCGCCCCCCA	TC-ATACATTC		
ATGTCCCG-T	TCGCGGTGTG	CATAC-TGGT	CGC-GGCC-T	CCTTGGAATC	ACATC-GC--	TCGTCCCCCA	-CCATACTTC		
TTGCCCTG-T	TTGC?GTGTG	CATGC?-GTT	CGT-GGAC-T	CFTTGTAAAT	ACATC-GC--	TCGCCCCC-A	TC-ATACAAC		
TTGCCCCG-T	TTGCGGTGTG	C?TGC--GGT	TGT-GGCCTT	CCTTGGAATC	ACATC-GCG-	TCGCTCCTAA	ACCATGCTTC		
TCGCCCCG-T	TTGCGGTGTG	CGTACA-GGT	CGT-GGCC-T	CCTTGGAATC	ACATC-GC--	TCGCTCCCCA	-CCATACCTC		
TTGCCCCGAT	TTATGGTGTG	CATG-ATGGT	TGT-ATCC-T	CCTTGAATTT	AAATC-GCG-	TCGCCCCCCA	TC-AAACAAC		
ATGCCCCGAT	TTACGGTGTG	CATG-ATGGT	TGT-ATCC-T	ACTTTAATTT	AAATC-GC--	TCGCCCCC-A	TC-AAACAGC		
TTGCCCCGAT	TTACGGTGTG	CATGCATGGT	TGT-ATCC-T	ACTTTAATTT	AAATC-GTGC	TCGCCCCC-A	TCCAACAGC		
TTGTCCCGTT	TTTCGGTGTG	CACACA-GGC	TCT-AGC-T	CCTTGAAATC	ACATC-GTGT	TGCCCCCC-A	GG-AAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGA	TAT-AGCC-T	CCATGAAATC	ACATC-GCG-	TCGCCCCCTT	GCCAACAGTA		
TTVCCG?GTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCCCA	-CCGAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTAAATC	ACATC-GTG-	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTAAATC	ACATC-GCG-	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTAAATC	ACATC-GTG-	TCGCCCCCC-	GCCAACATC		
CTGCCCCGTT	TTACGGTGTG	CATACA-GGT	GGT-AGCC-T	TCTTTAAATC	ACATCTGCG-	TCGCCCCCT-	GCCACACAT-		
TTGCCCTVGT	TTACGGTGTG	CATACA-GGT	GGT-AGCC-T	TCTTTAAATC	ACATCT-CG-	TCGCCCCCT-	GCCACACAT-		
CTGCCCCGTT	TAGCGGTGTG	CATACATG-T	GGT-GGCA-T	TCTTTAAATC	ACATC-GCG-	TCGCCCCCC-	GCCACACAT-		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCC-T	CTTTAAATC	AAATC-GTCT	TGSCCCC-T-	GCCAACAT-		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTAAATC	AAATC-GCG-	TCGCCCCC-T-	GCCAACAT-		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCCC-	GCCAACATC		
TTGCCCYGTT	TTACGGTGTG	CATACA-GGT	GGT-AGCC-T	TCTTTAAATC	ACATCT-C?	TCGCCCCYT-	GCCACACAT-		
TTGCCCCGTT	TT?CGGTGTG	CATGCA-GGT	GGT-AGCA-?	TCTTTAAATC	ACATC-GC--	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCTCC-	GCCAACATC		
TTGCCCYGTT	TTACGGTGTG	CATGCA-GGT	GGT-AGCA-A	TCTTTAATTT	ACATC-GC--	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTACGGTGTG	CATGCA-GGT	GGT-AGCA-A	TCTTTAATTT	ACATC-GC--	TCGCCCCCC-	GCCAACATC		
TTGCCCYGTT	TTACGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTAAA-T	ACATC-GTG-	TGCCCCCCA	-CCAACATC		
TTGCCCCGTT	TTACGGTGTG	CATGCATG-T	GGT-AGCA-T	TCTTTAAATC	ACATC-GCG-	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTACGGTGTG	CATGCATG-T	GGT-AGCA-T	TCTTTAAATC	ACATC-GCG-	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTACGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTCAATC	ACATC-GTG-	T?SCCCCTC-	GCCAACATC		
TTGCCCCGTT	TTACGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTAAATC	ACATC-GC--	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTACGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTAAATC	ACATC-GS--	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTACGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTAAATC	ACATC-GCG-	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTACGGTGTG	CATGCA-GGT	GGT-AGCA-T	TCTTTAAATC	ACATC-GCG-	TCGCCCCCC-	GCCAACATC		
TTGCCCTGTT	TTGCGGTGTG	CATFCA-GGT	GGT-AGCC-T	CCTTAAATC	ACATC-ACG-	TCGCCCCCC-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-TGCC-T	CCCTAAAAAC	ACATC-GCG-	TCGCTCCCT-	GCCAACATC		
TTGCCCCGTT	TT-CGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCC-T-	GCCAACATC		
TTGCCCCGTT	TT-CGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TC?CCCC-T-	GCCAACATC		
TTGCCCCGTT	TT-CGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCC-T-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATG?A-GTT	GGT-AGCC-T	ACTTTAAATC	ACATC-GCG-	TCGCCCCC-T-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GTT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCC-T-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGCA-GGT	GGT-AGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCC-T-	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGAA-GAT	GGT-TGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCC-T?	GCCAACATC		
TTGCCCCGTT	TTGCGGTGTG	CATGAA-GAT	GGT-TGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCC-T-	GCCAACATC		
TTGCTCCGTT	TTGCGGTGTG	CATGAA-GGT	GGT-TGCC-T	CCTTTAAATC	ACATC-GCG-	TCGCCCCCT-	GCCAACATC		

APPENDIX 1. Continued

	310	320	330	340	350	360	370
KRIG	C-TCAT-GGG	ATG--CTTGG	CATCGGGG--	CGGAGATTGG	CCTCCCGTGC	CTTGTGTGTG	GTTGGCCTAA
PYRR	CTT-AC-GGG	ATG--CTTGG	CATTGGGG--	CGGAGATTGG	CCTCCCGTGC	TTTTGTGTGC	GTTGGCCTAA
MICR	CTATTT-GGG	ATG--ATTGG	CATCGGGG--	CGGATATTGG	ACTCCCGTGC	CTTTGTGTGTG	GTTGTCCCTAA
TARA	CCTTAA-GGG	TAGT-CGTGG	TGATTGGGAG	CGGAGATTGG	CTCCCGTGC	TTGTGTGTGC	GTTGGFCAAA
PRAL	CATGAT-GGT	TAGT-CATGG	TGTTTGGGGG	CGGAGATTGG	CCTCCCGTGC	TTGTGTGTGC	GTTGGCTTAA
PRPU	CCAACC--GG	TTGT-CATGG	TGAT--GGG-G	CGGAGATTGG	TCTCCCGTAC	TTGTT--CCG	GTTGGCC?AA
LAPE	CCTAAC-GGG	TTGT-GATGG	TGTTAGG--G	CGGATAGTGG	CCTCCCGTTC	TTATGTTTTCG	GTTGGCCTAA
LASA	CCT-AC--GG	TTGG-CATGG	TGTTGGGG--	CGGATAATGG	CCTCCCGTGC	TTGTGTTTTCG	GTTGGCCTAA
REPI	CC-CAT-GCG	TAAG-TATGG	TGATGGGG-G	CG-AAAATTGG	CCTCCCGTTC	TT--CGTGGC	GTTGGCCTAA
RETI	CCTCAT-GGG	TAAG-TTTGG	TGATGGGG--	CGGAAAATTGG	CCTCCTGTTC	TTG-CGTGGC	GTTGGCCTAA
RELI	CCTCAT-GGG	TAAG-TTTAG	TGAT--GGAG	CGGAAAATTGG	CCTCCCGTTC	TTG-CGTGGC	GTTGGCCTAA
LAAR	CCTAAC-GGG	TAAT-CCTGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-CGTGGC	GTTGGCCTAA
LANU	CC-ATC-GGG	TAAT-GCTGG	TGATCGGG-G	CGGAAACTGG	TCTCCTGTTC	TTA-CGTGCA	GTTGGCCTAA
AETH	CTTAAAGGGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGATCG	GTTGGCCCAA
SVEN	CCCTAA--GG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCCAA
LACT	CCCAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCCG	GTTGGCCTAA
PRPE1	CCCAAA-GGG	TAATTC?TGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
PRPE2	CCCAAA-GGG	TAAT-CGTGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
SOKI	CCTAAAGGTG	TAAT-AATGG	TGATGGTG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTTTG	GTTGGCCTAA
SOAS	CCTAAAGGTG	TAAT-AATGG	TCATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTTTG	GTTGGCCTAA
SOBO	CCTAAA-GGG	TAAT-CGTGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
SOAR	CCTTAATGGG	TAAG-CATGG	TGATGGGG--	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCCG	GTTGGCCTAA
SOMA	CCTTAATGGG	TAAG-CATGG	TGATGGGG--	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTTTCG	GTTGGCCTAA
SOPA	CCGAA--GG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-CGTCCG	GTTGGCCTAA
SOOL	-CTAAAGGTG	TAAT-AATGG	TCATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTTTG	GTTGGCCTAA
SOTU	CCCAAA--GG	--ATT-A--GG	TGATCGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
SOLU	CCGAAA--GG	TAGT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTTTCG	GTTGGCCTAA
SOSC	CCGAAA--GG	TAGT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCCG	GTTGGCCTAA
SOCA	CCCAAA--GG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
SOCO	CCCAAA--GG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	-TG-TGTCTG	GTTGGCCTAA
SOPR	CCCAAA--GG	TAAT-AATGG	TGAAGGGG-G	CGGAAAATTGG	--TCCCGTTC	TTG-TGTCCG	GTTGGCCTAA
SOGO	CCCAAA--GG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
SOOR	CCCAAA--GG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTGA
BABC	CCGAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
TAPI	CCCAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
TAAR	CCYAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
TACA	CCCAAA--GG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
TAHE	CCCAAA--GG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCTG	GTTGGCCTAA
KRIK	CTGAGA-GGG	TAAT-A--GG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCCG	GCTGGCCTAA
EMBE	CTGAGA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCCG	GTTGGCCTAA
DELI	CTGAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCG?TC	TTG-TGTTTCG	GTTGGCCTAA
DEMA	CTGAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCG?TC	TTG-TGTTTCG	GTTGGCCTAA
DEMC	CTGAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCG?TC	TTG-TGTTTCG	GTTGGCCTAA
DEMI	CTGAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCG?TC	TTG-TGTCCG	GTTGGCCTAA
DEPR	CTGAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCG?TC	TTG-TGTCCG	GTTGGCCTAA
DENE	CTGAAA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCG?TC	TTG-TGTCCG	GTTGGCCTAA
DEPI	CTGATA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCG?TC	TTG-TGTCCG	GTTGGCCTAA
DEBE	CTGATA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCG?TC	TTG-TGTCCG	GTTGGCCTAA
DERE	CTGCTA-GGG	TAAT-CATGG	TGATGGGG-G	CGGAAAATTGG	CCTCCCGTTC	TTG-TGTCCG	GTTGGCCTAA

APPENDIX 1. Continued

	460	470	480	490	500	510
KIRG	GTGAGGGAGG	CCCTT--CAT	GAAGACCCCA	AT-GTGGTGT	CTTGCGACGA	TGCTT-CGAC
PYRR	GTGAGGTGGG	CCCTT--GAT	GAAGACCCCT	AT-GTATCGT	CATGTGACGA	TGCTT-CGAC
MICR	GTAAGGGA-G	GCCTT---T	GAAGACCCCA	AC-GTGA-GT	CTTGCGACGA	TGCTT-CGAC
TARA	GCTAGGGAAA	CCCTC--AAA	AAAGAACCCTA	AT-GTATCGT	TCTAGGACGA	TGCTT-CGAC
PRAL	GCTAGGGAAA	CCCTC--ATC	AAAGACCCCA	AT-GTATT?-	CTTGCGACGA	TGCTT-CGAC
PRPU	GCTAGGGGAA	-CCTC--ATC	AA-GACCCCA	TC-GTATCGT	TTTAGGACGG	TGCTT-CGAC
LAPE	GTGAGGGGAA	CCCTC--ATC	AATGACCCCT	TT-GTATCGT	CTTCGGACGG	TGCTT-CGAC
LASA	GTAAGGGTAG	CCCTC--ATC	AAAGACCCCA	TT-GTATCGT	CTTCGGATGA	TGCTT-CGAC
REPI	ATTAGGGGAA	GCCTT--TTA	TAAGACCCCA	TT-GTATCGT	TATAAAAACGG	TATAT-CGAC
RETI	GTTTGGGAA	GCCTT--ATA	TATGACCCCA	TT-CTATCGT	TATAAAAACGG	TATAT-CGAC
RELI	GTTTGGGAA	GCCTT--ATA	TATGACCCCA	TT-CTATCGT	TATAAAAACGG	TATAT-CGAC
LAAR	GTGAGGGGAA	-CCTA-CTTT	TACGACCCCA	C?-GTATCGT	TAT?AGACGA	TATAT-CGAC
LANU	GTTAGGGGAA	-TCTC--ATT	TAAGACCC-A	TT-GTATCGT	TATAAGGCGA	TATAT-CGAC
AETH	GCGAGGGGAG	CTCTT--ATT	TTAGACCCCA	TT-GTATCTT	TAAAAAACGA	TACAT-CGAC
SVEN	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAAAGACGA	TATAT-CGAC
LACT	GTAAAGGAAA	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAAAGATGA	TATAT-CGAC
PRPE1	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAAAGACGG	TATAT-CGAC
PRPE2	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAAAGACGG	TATAT-CGAC
SOKI	GTGAGGGAAA	TTCTCTCAAT	TTAGACCCCA	CT-GTATCGT	TAAAAAACGA	TATAT-CGAC
SOAS	GTGAGGGAAA	TTCTC--AAY	TTAGACCCCA	CT-GTATCGT	TAAAAAACGA	TATAT-CGAC
SOBO	GTGAGGGAAA	TTCTC--ATT	TTAGACCCCA	CT-GTATCGT	TTGAAAACGA	TACAT-CGAC
SOAR	GTGAGGGGAA	TTCTC--ATT	TTAGACCCCTA	CT-GTATCGT	TAAAA-ATGA	TATAT-CGAC
SOMA	GTGAGGGGAA	CCCTC--ATT	TTAGACCCCTA	CT-GTATCGT	TAAAA-ATGA	TATATT-GAC
SOPA	GTGAGGGGAA	CTCTC--TTT	TCAGACCCCTA	AT-ATATCGT	TTACAGACGA	TATAT-CGAC
SOOL	GTGAGGGAAA	TTCTC--AAT	TTAGACCCCA	CT-GTATCGT	TAAAAAACGA	TATAT-CGAC
SOTU	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTT	TT-GTATCGT	TAAAAGACGA	TATAT-CGAC
SOLU	GCGAGGGGAG	CTCTC--ATT	TTAGACCCCTA	AT-GTATCGT	TAAAAGACGA	TATAT-CGAC
SOSC	GTGAGGGGAA	CTCTC--ATT	TTAGACCCCTA	AT-GTATCGT	TAAAAGACGA	TATAT-CGAC
SOCA	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	-GAAAGACGA	TATAT-CGAC
SOCO	GTGAAGGAAG	TTCTC--ATG	TCAGACCCCTA	CY-GTATCGT	TAAAAGAYGA	TATAT-CGAC
SOFR	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAAAGACGA	TATAT-CGAC
SOGO	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAATGACGA	TATAT-CGAC
SOOR	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAATGACGA	TATAT-CGAC
BABC	GTGAAGGAAG	TTGTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAAAGACGA	TATAT-CGAC
TAPI	GTGAAGGAAG	TTCTA--ATT	TCAGACCCCTA	CT-GTATCGT	TAAATGACGA	TATAT-CGAC
TAAR	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAAAGACGA	TATAT-CGAC
TACA	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAATGACGA	TATAT-CGAC
TAHE	GTGAAGGAAG	TTCTC--ATT	TCAGACCCCTA	CT-GTATCGT	TAAATGACGA	TATAT-CGAC
KIRK	GTGAGGGGAA	CTCTC--ATT	TTAGACCCCA	CT-GTATCGT	TTAAAAATGA	TATATTGATC
EMBE	GTGAGGGGAA	CTCTC--ATT	TTAGACCCCA	CTAGTATTGT	TAAAAATGA	TATAT-CGAC
DELI	GGGAGGGGAA	CTCTC--ATT	CTAGACCCCTA	CT-GTATCGT	TAAAAATGA	TATAT-CGAC
DEMA	GGGAGGGGAA	CTCTC--ATT	CTAGACCCCTA	CT-GTATCGT	TAAAAATGA	TATAT-CGAC
DEMC	GGGAGGGGAA	CTCTC--ATT	CTAGACCCCTA	CT-GTATCGT	TAAAAATGA	TATAT-CGAC
DEMI	GAGAGGTAAG	CTCTA--ATT	TTAGACCCCA	CT-GTATCGT	TAAAAATGA	TATAT-CGAC
DEPR	GAGAGGTAAG	CTCTA--ATT	TTAGACCCCA	CT-GTATCGT	TAAAAATGA	TATAT-CGAC
DENE	GGGCGGGGAA	CTCTC--ATT	TTAGAACCCCA	CT-GTATCGT	TAAAAATGA	TATAT-CGAC
DEPI	GGGAGGGGAA	TTCTC--ATT	TTAGACCCCA	CT-GTATCGT	TAAAAATGA	TATAT-CGAC
DEBE	GGGAGGGGAA	TTCTC--ATT	TTAGACCCCA	CT-GTATCGT	TAAAAATGA	TATAT-CGAC
DERE	GGTAGGGGAA	TTCTC--ATT	CTAGACCCCA	CT-GTATCGT	TATAAATGA	TATAT-CGAC