Pl. Syst. Evol. 190: 157–185 (1994)

—Plant—
Systematics
and
Evolution
© Springer-Verlag 1994
Printed in Austria

Comparisons of phylogenetic hypotheses among different data sets in dwarf dandelions (*Krigia*, *Asteraceae*): additional information from internal transcribed spacer sequences of nuclear ribosomal DNA

K.-J. Kim and R. K. Jansen

Received December 10, 1992; in revised version September 27, 1993

Key words: *Asteraceae, Lactuceae, Microseridinae, Krigia.* – Internal transcribed spacer (ITS), ITS region (ITS 1, 5.8 S rDNA, and ITS 2), intergenic spacer (IGS) region, nuclear ribosomal DNA (nrDNA), chloroplast DNA (cpDNA), sequence divergence, polymerase chain reaction (PCR), maximum parsimony (MP) tree, neighbor joining (NJ) tree.

Abstract: The internal transcribed spacer (ITS) region of the 18 S–25 S nuclear ribosomal DNA repeat was sequenced from 19 populations of the tribe *Lactuceae*, including all species of dwarf dandelion (*Krigia*) and five outgroup genera. The incidence of length changes and base substitutions was at least two times higher for ITS 1 than ITS 2. Interspecific sequence divergence within *Krigia* averaged 9.62% (1.61%–15.19%) and 4.26% (0%–6.64%) in ITS 1 and ITS 2, respectively. Intergeneric sequence divergence ranged from 15.6% to 44.5% in ITS 1 and from 8.0% to 28.6% in ITS 2. High sequence divergence and homoplasy among genera of tribe *Lactuceae* suggest that the phylogenetic utility of ITS sequence data is limited to interspecific studies or comparisons among closely related genera. Trees generated from ITS sequences are essentially identical to those from restriction site comparisons of the entire nuclear ribosomal (nr) DNA region. The degree of tree resolution differed depending on how gaps were treated in phylogenetic analyses. The ITS trees were congruent with the chloroplast DNA and morphological phylogenies in three major ways: 1) the sister group relationship between *Krigia* and *Pyrrhopappus*; 2) the recognition of two monophyletic sections, *Krigia* and *Cymbia*, in genus *Krigia*; and 3) the monophyly of the *K. occidentalis*-*K. cespitosa* clade in sect. *Cymbia*. However, the two nrDNA-based trees are not congruent with morphology/chloroplast DNA-based trees for the interspecific relationships in sect. *Krigia*. An average of 22.5% incongruence was observed among four *Krigia* data sets. The relatively high degree of incongruence among data sets is due primarily to conflict between trees based on nrDNA and morphological/cpDNA data. The incongruence is probably due to the concerted evolution of nrDNA repeating units. The results from *Krigia* and the *Lactuceae* suggest that nrDNA data may have limited utility in phylogenetic studies of plants, especially in groups which exhibit high levels of sequence divergence. Our combined phylogenetic analysis as a total evidence shows the least conflict to each of the individual data sets.

Congruence of phylogenetic trees generated from different data sets is one way to

examine the robustness of phylogenetic hypotheses. Despite the numerous examples of concordance between molecular and morphological data both in the plant and animal literature, complete congruence is often not achieved (DOYLE 1992, HARRISON 1991, HILLIS 1987, MIYAMOTO 1985, SWOFFORD 1991a). Phylogenetic hypotheses from additional data sets often disagree in minor details and sometimes show major conflicts. Therefore, the estimation of the true species phylogeny from a single data source or from a single gene tree is a dangerous practice in phylogeny reconstruction. If some degree of conflict is encountered among data sets, the possible causes of the discrepancies should be identified through rigorous analytical methods both in evolutionary process and in tree construction process.

Some plant genera have been subjected to extensive phylogenetic study using several independent lines of data. *Krigia* (dwarf dandelion) is one such genus. Three independent phylogenies were published recently using morphological and cytological data (KIM & TURNER 1992) and restriction site comparisons of chloroplast (cp) DNA (KIM & al. 1992a) and nuclear ribosomal (nr) DNA (KIM & MABRY 1991). Those independent data sets provided an opportunity to evaluate new phylogenetic markers and to construct a robust species phylogeny. As a fourth phylogenetic marker for the genus, we have examined sequence variation of internal transcribed (ITS) regions of nrDNA repeating units.

Tandemly repeated nrDNA encodes three ribosomal RNA (rRNA) genes (18S, 5.8S, and 25S) and each copy contains a transcribed region that is separated by a long nontranscribed intergenic spacer (IGS). The transcribed region contains three rRNA coding genes along with two internal transcribed spacers (ITSs), which occur in the 5′-18S-ITS1-5.8S-ITS2-25S-3′ order and are transcribed as a single precursor rRNA. The two ITSs (ITS1 and ITS2) are subsequently removed and three rRNA coding regions eventually mature into three rRNAs (reviewed in ROGERS & BENDICH 1987).

All three rRNA coding regions are highly conserved both in structure and in the level of sequence divergence (LAKE 1985, OLSEN 1987, FIELD & al. 1988, HAMBY & ZIMMER 1992). In contrast to the coding regions, numerous studies demonstrated that the IGS region experiences more rapid evolutionary change in structure and sequence (APPELS & DVORAK 1982, BARKER & al. 1988, UEKI & al. 1992, YAKURA & al. 1984, YAKURA & NISHIKAWA 1992).

The ITS region also exhibits a substantial amount of variation both in length and sequence (JORGENSEN & CLUSTER 1988, RIESEBERG & al. 1988, SYTSMA & SCHAAL 1990, VENKATESWARLU & NAZAR 1991). However, it has not been widely used for phylogenetic studies because the ITS is too short for restriction site comparisons. Several ITS sequences were reported from various crops and wild plants such as rice (TAKAIWA & al. 1985), mustard (RATHGEBER & CAPESIUS 1989), mung bean (SCHIBEL & HEMLEBEN 1989), tobacco (VENKATESWARLU & NAZAR 1991), tomato (KISS & al. 1988), melon (KAVANAGH & TIMMIS 1988), oats (CHATTERTON & al. 1992a), barley (CHATTERTON & al. 1992b), and *Populus deltoides* (D'OVIDIO 1992). Unfortunately, none of these studies focused on phylogenetic relationships. Three recent studies discussed the phylogenetic utility of ITS sequences in animals (GONZALEZ & al. 1990) and fungi (O'DONNELL 1992, LEE & TAYLOR 1992). In flowering plants, ITS sequences were utilized successfully for examining phylogenetic relationships among the genera of *Madiineae* group of the family *Asteraceae*

(BALDWIN 1992, 1993). The small size of the ITS regions (less than 800 bp) makes the marker particularly appropriate to direct sequencing of amplified DNA from polymerase chain reaction (PCR). In addition, the conserved nature of the surrounding coding regions of nrDNA (16 S, 5.8 S, and 25 S) regions allow the construction of universal primers that can be used for a wide diversity of organisms from fungi to flowering plants (WHITE & al. 1990).

In the present paper, we report on the interspecific and intergeneric relationships of *Krigia* using ITS sequences. The new ITS data is then compared and combined with the three previously published data sets. We also describe the detailed direct sequencing strategy from the double strand (ds) amplification products of the ITS region.

Material and methods

Source of plant material. Sources of plant material for the 14 populations of *Krigia* and species from five outgroup genera are given in Table 1. All voucher specimens of *Krigia* are deposited in either the Plant Resource Center, University of Texas at Austin (TEX) or the Herbarium of Oregon State University (OSC). Ploidy levels for all populations were determined from meiotic counts of bud materials (KIM & TURNER 1992).

Total DNA extraction and purification. Total DNAs were extracted from fresh leaf tissue pulverized in liquid nitrogen according to the methods of SAGHAI-MAROOF & al. (1984) as modified by DOYLE & DOYLE (1987). DNA was extracted from a single individual of *K. dandelion, K. montana, K. biflora*, and *K. cespitosa.* Because of the small size of each individual, up to 15 plants were combined for DNA extraction in *K. occidentalis, K. virginica*, and *K. wrightii.* DNAs were further purified by cesium chloride/ethidium bromide gradients. The concentrations of the DNAs ranged from 20 ng/μl to 80 ng/μl depending on samples. The purified DNAs were diluted 10 times with distilled water and 1 μl of each diluted DNA (2–8 ng) was used for 100 μl amplification reactions.

Amplification of ITS region. We amplified the entire ITS region using standard double-strand PCR reactions using the primers P 1 and P 4 (Fig. 1) and sequenced both strands directly from dsPCR products using four primers (P 1, P 2, P 3, and P 4 in Fig. 1). Direct sequencing of dsPCR products is superior to that of single-strand (asymmetric) PCR products for two reasons. First, ds amplification typically yields more DNA than asymmetric amplification. Second, only a single amplification is required for sequencing both strands. Two independent amplifications are necessary for sequencing both strands in asymmetric amplification (GYLLENSTEN & ERLICH 1988). Indeed, a single 100 μl ds amplification yields sufficient amounts of DNA for sequencing both strands using the four ITS primers (Fig. 1).

The highest yield of PCR product was achieved using the following conditions. The 100 μl PCR reactions contained (in order of addition) 70.5 μl sterile water, 10 μl of 10 × *Taq* polymerase reaction buffer, 2 mM (8 μl of 25 mM stock) magnesium chloride solution, 0.2 mM (2 μl of 10 mM stock) of each dNTPs (total 8 μl), 0.25 μM (1 μl of 25 μM stock) of each primer (total 2 μl, P 1 and P 4 in Fig. 1), 2.5 units (0.5 μl of 5 units/μl) of *Taq* DNA polymerase, and 2–8 ng (1 μl of 2–8 ng/μl) of purified total DNAs. We obtained identical amplification results using either *Taq* DNA polymerase (Perkin-Elmer or Promega) or *Tfl* DNA polymerase (Epicentre Tech.). For multiple sample amplifications, all of the reaction components were premixed, 99 μl of the premixed solution was aliquoted to each tube and DNA was added last. The premixed solution without template DNA was used as a negative control.

Reaction mixtures were sealed with one or two drops of mineral oil to prevent evaporation during thermal cycling. For complete denaturation of the template DNA, the first cycle used a longer denaturation time (3 min) than the normal cycle (1 min) at 95 °C. Each

Table 1. Species and populations used for ITS region sequencing. Voucher: *K* K.-J. KIM, *KC* KIM & E. CALVO, *KG* KIM & S. GINZBARG, *KW* KIM & L. WOODRUFF, *Ca* K. L. CHAMBERS, *BLT* B. L. TURNER, *JP* H. J. PRICE

	Abbreviation	Chromosome (n)	Voucher	Locality State: County	*Sequence accession
Krigia dandelion (L.) NUTT.	DAN 1	30	KC 10626 (TEX)	Alabama: Elmore	L 20483
	DNA 2	30	K 10116 (TEX)	Texas: Upshur	L 13945
K. montana (MICHX.) NUTT.	MONt	10	K 10141 (TEX)	North Carolina: Buncombe	L 13946
	MONh	15	Ca 2880 (OSC)	North Carolina: Buncombe	L 13947
K. biflora (WALT.) S. F. BLAKE	BIFd	5	K 10652 (TEX)	Arkansas: Garland	L 13948
	BIFt	10	Ca 1404 (OSC)	West Virginia: Nicholas	L 20484
K. virginica (L.) WILLD.	VIRd	5	KC 10609 (TEX)	Louisiana: St. Tammany	L 13949
	VIRt	10	KC 10622 (TEX)	Alabama: Lee	L 20478
K. occidentalis NUTT.	OCC	6	K 10529 (TEX)	Texas: Montgomery	L 13950
K. wrightii (A. GRAY) KIM	WRI 1	9	K 10115 (TEX)	Texas: Upshur	L 20480
	WRI 2	9	K 10519 (TEX)	Texas: Lee	L 13951
K. cespitosa (RAF.) CHAMBERS	CES 1	4	KG 10128 (TEX)	Texas: San Jacinto	L 20481
	CES 2	4	KW 10106 (TEX)	Texas: Burleson	L 13952
	CES 3	4	K 10109 (TEX)	Texas: Robertson	L 20482
Pyrrhopappus grandiflorus NUTT.	PYR	6	K 10508 (TEX)	Texas: Washington	L 13953
Microseris laciniata SCH. – BIP.	MIC	9	Ca 5369 (OSC)	Oregon: Denton	L 13954
Agoseris heterophylla GREENE	AGO	9	JP C61-0-11 (TAES)	California: Montery	L 13955
Stephanomeria pauciflora A. NELS.	STE	9	BLT 15802 (TEX)	California:	L 13956
Lactuca sativa L.	LAC	9	Cultivated	–	L 13957

* All sequences are deposited in GenBank and also available through EMBL and DDBJ under same accession number.

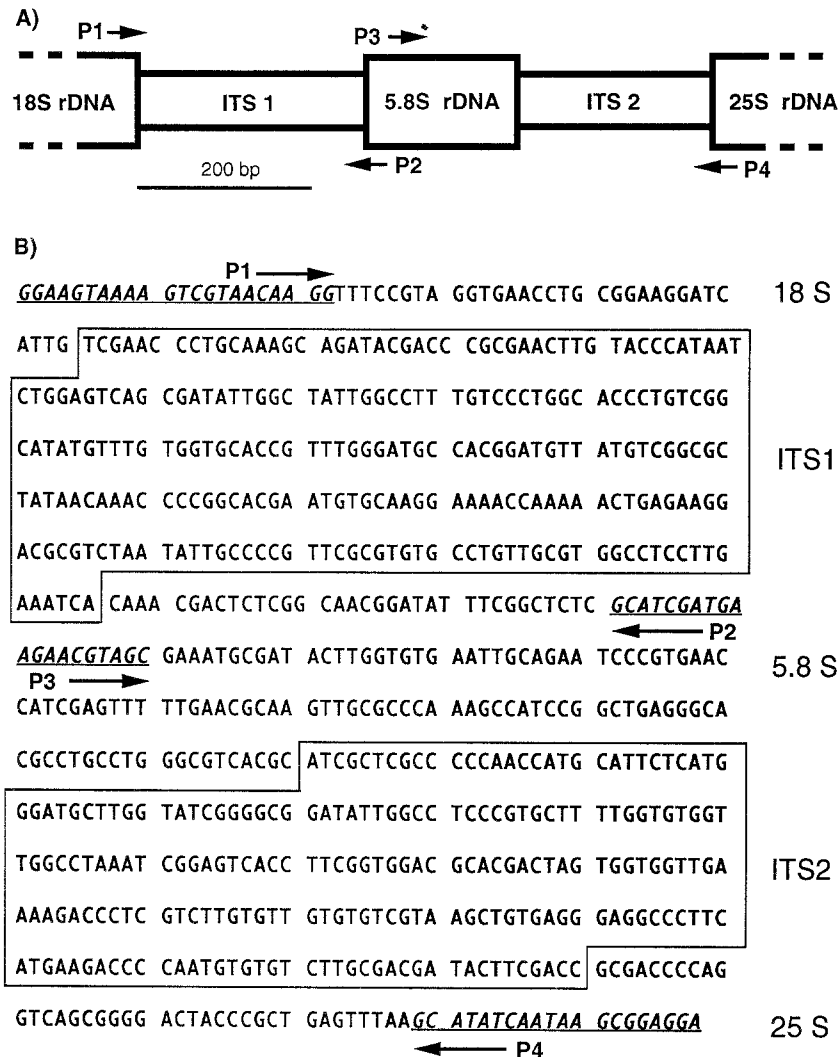


Fig. 1. Internal transcribed (ITS) regions of 18 S–25 S nuclear ribosomal DNA. *A* Diagram showing the general structure of ITS region and primer positions (P 1–P 4). *B* Sequence of ITS region in *Krigia biflora*. The complete sequence between two amplifying primers, P 1 and P 4, is 706 base pairs (bp) in length including 32 bp of 18 S rDNA, 252 bp of ITS 1, 164 bp of 5.8 S rDNA, 220 bp of ITS 2, and 38 bp of 25 S rDNA. The underlined sequences indicated the four sequencing primers. P 3 is complementary to P 2. The sequence of P 2 (P 3) in *K. biflora* is identical to the primer sequence. Sequences for P 1 and P 4 were not verified from *K. biflora*. P 1, P 2, P 3, and P 4 correspond to the ITS 5, ITS 2, ITS 3, and ITS 4, respectively, of White & al. (1990) (see also Baldwin 1992)

of the 30 cycles consisted of 1 min at 95 °C to denature template DNA, 1 min at 55 °C to anneal primers to denatured template DNA, and 45 sec at 72 °C for primer extension. Primer extension time was gradually increased by 3 sec intervals during each cycle. After 30 cycles the PCR reactions were incubated at 72 °C for 7 min to complete primer extension.

Purification of PCR products. Small amounts of PCR products (5 µl) were visualized with ethidium bromide after agarose gel electrophoresis using a mini-gel system. Only a

single band of 700 basepair (bp) was detected and no detectable length variation was observed among species and genera. Typically 5 µl of PCR product shows a stronger signal than 1 µg of *Hind*III digested lambda DNA marker under the UV light.

PCR-generated DNAs were purified by electrophoresis through low-melting agarose gel matrix using 1 × TAE buffer. After staining the gel with ethidium bromide, the agarose block containing DNAs was excised from the gel using a razor blade under low wavelength (370 nm) UV light. Concentrated DNA (30 µl, approximately 1/3 of original PCR reaction volume) was recovered using the GeneClean II system according to manufacturer's protocol (Bio 101 Co.).

Sequencing of ITS region. Sequences were generated by standard dideoxy chain termination reaction using Sequenase (version 1.0, U.S. Biochemical). We modified the Sequenase protocol (version 1.0) using the snap-cooling method of WINSHIP (1989). First, microcentrifuge tubes containing annealing mixtures (7 µl of purified DNA template DNA – approximately 2 µg, 1 µl of 25 µM primer – 0.25 µM, 2 µl of 5 × Sequenase reaction buffer) were placed in boiling water for 3 min to denature the template DNAs and the tube was transferred immediately into an ice bath for 5–15 min to anneal the primers to the template DNAs. Second, the labeling reactions were done in the ice bath rather than at room temperature for 5 min (the labeling reaction components are the same as in the Sequenase protocol). Third, the termination reactions were carried out at 40 °C.

Both strands were sequenced using four ITS primers (Fig. 1). Reactions with two reverse primers, P 2 and P 4, worked well with standard reaction conditions mentioned above. However, reactions with the two forwarding primers, P 1 and P 3, showed weaker sequencing signals. This is probably due to the secondary structures of nrDNA sequences near the priming sites toward to primer extension orientation (see VENKATESWARLU & NAZAR 1991). The inclusion of dimethyl sulfoxide (DMSO) in the sequencing reactions greatly enhanced the sequencing signals for these two primers (see WINSHIP 1989). DMSO concentrations of 10% were added to the annealing, labeling, and termination mixes. DMSO also improved DNA yield of PCR amplification in some samples. Similar effects may also be achieved with the addition of formamide (ZHANG & al. 1992) or the anionic detergent tergitol (type NP-40) (BECHMANN & al. 1990, WANG & al. 1992) both in amplification and in sequencing. The addition of 1 µl of Mn buffer (0.1 M $MnCl_2$ and 0.15 M Sodium isocitrate) to each labeling reaction greatly enhanced the sequencing signal close to the primers.

DNA sequences were separated on a 6% acrylamide gel using wedge-shaped spacers. In order to read the entire ITS region, long and short gels were run for each sample. The short gel was run until the bromphenol blue dye migrated to the bottom of the gel. The long gel was electrophoresed for 12 h at 1400 volts. Sequencing reactions with Mn buffer were electrophoresed less than 2.5 h at the same voltage. Gels were fixed in 10% glacial acetic acid for 15 min, transferred to 3 MM Whatman paper, vacuum dried at 80 °C for 8 h and exposed to Kodak XAR X-ray film for 12–48 h.

Sequence alignment of ITS region. The sequence boundaries among two ITS regions and three coding regions (18 S, 5.8 S, and 25 S) of nuclear rDNA were determined by comparison with the published sequences from *Oryza sativa* (TAKAIWA & al. 1985) and *Vicia faba* (YOKOTA & al. 1989). Sequences were edited using the EyeBall Sequence Editor (ESEE, provided by E. CABOT) and aligned using Clustal V (provided by D. HIGGINS) program with gap adjustments. The 19 new sequences from this study were deposited in GenBank (see Table 1 for the accession numbers). Both unaligned and aligned sequences are also available from the senior author.

Outgroup selection. Five representative genera of the tribe *Lactuceae* were selected as outgroups, including *Lactuca, Stephanomeria, Agoseris, Microseris*, and *Pyrrhopappus. Lactuca* and *Stephanomeria* are traditionally classified in two different subtribes, *Crepidinae* and *Stephanomeriinae*, respectively (STEBBINS 1953). All other genera are included in the

subtribe *Microseridinae* with *Krigia*. The *Stephanomeriinae* is the closest subtribe to the *Microseridinae* (Stebbins 1953, Jansen & al. 1991).

Phylogenetic analysis of ITS sequences. Parsimony analyses were performed with unordered parsimony using PAUP (version 3.0 r, Swofford 1991 b) on a Macintosh Quadra 700. The branch-and-bound algorithm (Hendy & Penny 1982) was used to find the shortest trees in all analyses except bootstrapping. Bootstrap analyses (Felsenstein 1985) were performed using 1000 replicates and the heuristic search algorithm with the mulpars options. Decay analyses (Hillis & Dixon 1989) were performed until all branches collapsed. The evolutionary direction of sequence changes was inferred by outgroup comparison using multiple outgroups. After initial parsimony analyses, multiple populations from the same species were removed. As a result, the data matrix was reduced to 12 taxa, including seven ingroup species and five outgroup genera, for further phylogenetic analyses.

Sequence divergence values between species were calculated by the two-parameter method (Kimura 1980) using the DNADIST program of PHYLIP (version 3.4, Felsenstein 1991). This method allows for the correction of multiple substitutions and differential transition/transversion probability based on empirical observation from the data. The ratio was set at 1.0 based on the actually observed frequencies (Kimura 1981) in the maximum parsimony tree. As a result, it was assumed that there is an equal probability of independent change at all sites (Jukes & Cantor 1969). The composition of the four nucleotide bases (A : C : G : T) was also set to the actual ratios in the data pool. The neighbor joining tree was constructed from the corrected sequence divergence using the NJ program (version 2.0, Saitou & Nei 1987).

Cladistic analyses of combined data. Three other types of data were published previously from the same taxa of *Krigia* and the outgroup genera. The earlier data sets include 35 morphological and chromosomal characters (Kim & Turner 1992), 252 cpDNA restriction site changes (Kim & al. 1992 a) and 50 nrDNA restriction site characters (Kim & Mabry 1991). Two rDNA restriction site changes (characters no. 34 and 37) were removed before combining the data sets because they occur within ITS regions. All other nrDNA restriction site variants occur in the IGS region. Various combinations of the four data sets were analysed using PAUP. In order to identify the phylogenetic signal in the reduced data sets (Hillis 1991), the skewness of tree-lengths distribution was determined using the exhaustive search option of PAUP in all individual data sets and in a combined data set. The four data sets were also combined directly as a single matrix and then parsimony and bootstrap analyses were performed with the same search options employed for the ITS data alone. The degree of character congruence among data sets was evaluated by the Mickevich & Farris (1981) index and the Miyamoto index (see Swofford 1991 a). Individual tree topologies from the four different data sets were also combined in two different ways: 1) the character re-coding method (Doyle 1992), and 2) the combinable consensus method (Bremer 1990). For character re-coding, all trees were weighted equally rather than treating the molecular data set as a single character.

Results

Length variation and base composition of ITS region. Complete sequences of the ITS region were generated for 19 taxa, including all species of *Krigia* and five outgroup genera. The ITS regions vary in length from 632 bp (*Agoseris*) to 636 bp (*Pyrrhopappus*), including 164 bp of the 5.8 S rRNA gene (Table 2). The length of ITS 1 ranges from 246 bp (*Agoseris*) to 253 bp (*Pyrrhopappus*). The length of ITS 2 ranges from 218 bp (*Microseris*) to 222 bp (*Agoseris*).

Sequence alignments required 20 and 5 independent insertion/deletion mutations in ITS 1 and ITS 2, respectively (Fig. 2). Eighteen of the 20 length mutations in ITS 1 and four of the mutations in ITS 2 are one or two bp insertions/deletions.

Table 2. Length and G + C content variation of ITS region from various higher plant groups. References: 1. VENKATESWARLU & NAZAR (1991), 2. KISS & al. (1988), 3. D'OVIDIO (1992), 4. RATHGEBER & CAPESIUS (1989), 5. KAVANAGH & TIMMIS (1988), 6. CHATTERTON & al. (1992 a), 7. CHATTERTON & al. (1992 b), 8. TAKAIWA & al. (1985), 9. YOKOTA & al. (1989), 10. SCHIEBEL & HEMLEBEN (1989), 11. BALDWIN (1992). All others are from the present study. *Members of *Asteraceae*

Taxa, reference	ITS 1		ITS 2		5.8 S	
	Length	G + C %	Length	G + C %	Length	G + C %
Nicotiana rustica[1]	216	69.4	217	65.4	163	55.4
Lycopersicon esculentum[2]	217	67.7	217	70.5	163	53.4
Populus deltoides[3]	220	66.8	207	69.5	164	55.4
Sinapis alba[4]	265	50.6	188	54.3	163	52.8
Cucumis melo[5]	216	55.6	237	60.2	163	58.9
Avena sativa[6]	219	56.2	213	62.4	163	58.9
Hordeum vulgare[7]	218	56.9	217	62.2	163	60.7
Oryza sativa[8]	194 (5)	72.7	233	77.3	163 (4)	59.5
Daucus carota[9]	215	49.4	224	52.3	164	54.2
Vigna radiata[10]	205	60.0	220	59.1	163	52.8
Vicia faba[9]	235	51.9	208	49.6	164	50.6
**Madiinae* (12 Taxa)[11]	255–261	47.7–51.4	216–223	49.5–53.0	164	51.2–53.7
**Krigia* spp. (7 species)	251–254	52.0–53.2	220	55.0–57.7	164	54.9–56.1
**Lactuca sativa*	251	53.8	221	55.1	164	54.9
**Agoseris heterophylla*	246	52.0	222	53.2	164	56.1
**Microseris laciniata*	253	53.4	218	54.8	164	55.5
**Pyrrhopappus grandiflorus*	253	52.6	219	55.7	164	54.3
**Stephanomeria pauciflora*	252	54.0	220	56.8	164	55.5

The longest length mutation was an eight-bp deletion in the ITS 1 of *Agoseris* (coordinates 87–94 in Fig. 2). In addition, two three-bp deletions were detected in the ITS 1 of *Lactuca* (223–225) and the ITS 2 of *Microseris* (640–642). Ten length mutations were hypothesized within *Krigia* and all of these are one or two bp insertions/deletions. The G + C content in the ITS region of the *Lactuceae* is slightly greater than 50% with a narrow range of variation in ITS 1 (52–54%), ITS 2 (55–58%), and 5.8 S gene (54–56%).

Polymorphic nuclear ribosomal DNA. DNA of the hexaploid *K. montana* was isolated from a single individual, nevertheless, the hexaploid showed polymorphic sequences at 20 sites (coordinates 82, 91, 102, 104, 105, 119, 147, 153, 156, 218, 268, 319, 430, 431, 500, 512, 528, 603, 613, and 672 in Fig. 2). All of these sites differ between the two parental species, *K. montana* and *K. biflora*, and the hexaploid shows a combined sequence. The two parental species also show two single-base pair differences in length at positions 106 and 556 (Fig. 2). The hybrid shows the double sequence bands from these positions because of the different length of nrDNA repeating units in the parents.

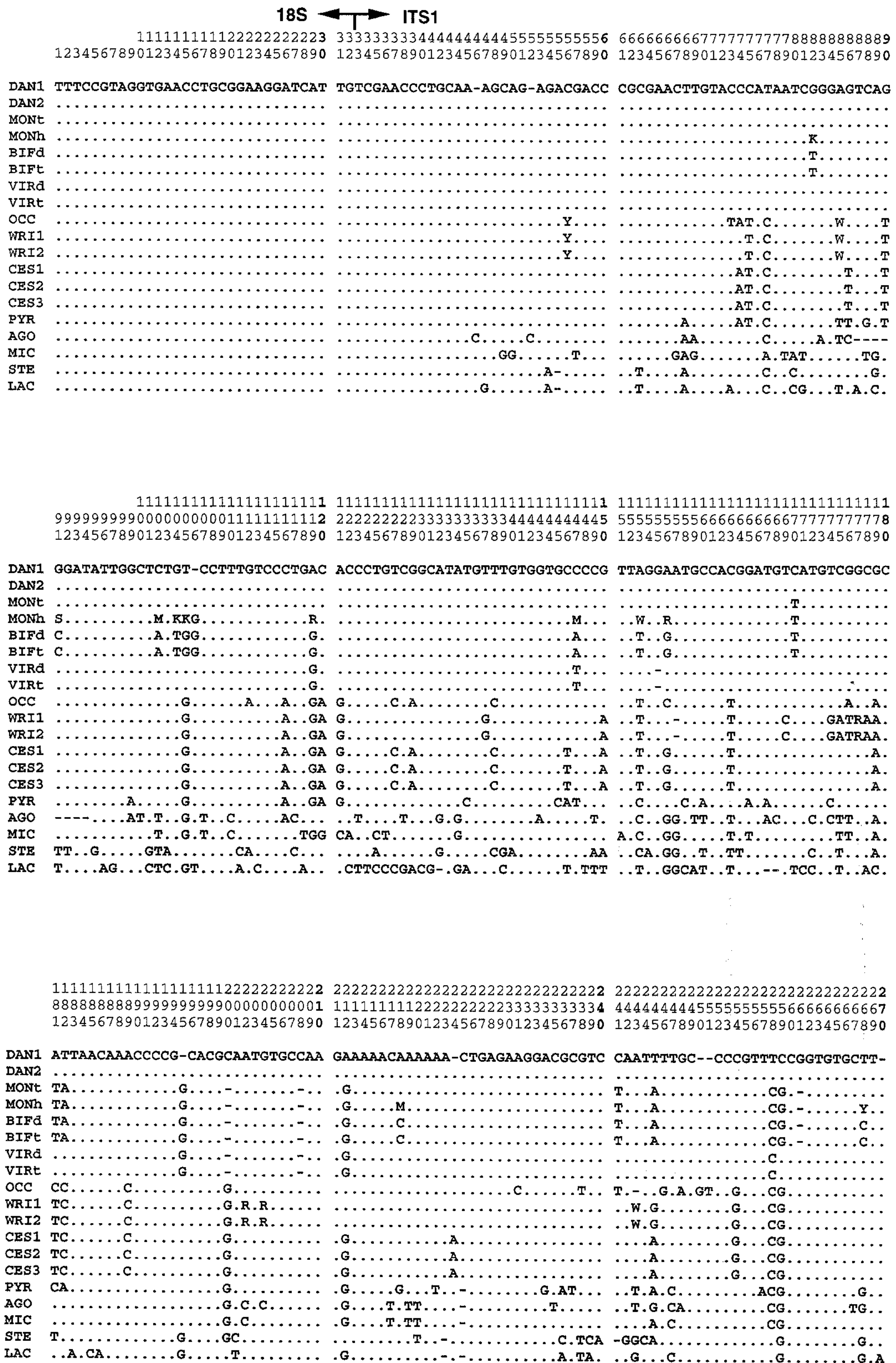


Fig. 2

```
              ITS1 <--|--> 5.8S
     222222222222222222222222222223 333333333333333333333333333333 333333333333333333333333333333
     777777777888888888899999999990 000000000111111111122222222223 333333333444444444455555555556
     123456789012345678901234567890 123456789012345678901234567890 123456789012345678901234567890

DAN1 GTTG-CGTGGCCTCCTTGAAATCACAAACG ACTCTCGGCAACGGATATCTCGGCTCTCGC ATCGATGAAGAACGTAGCGAAATGCGATAC
DAN2 .............................. .............................. ..............................
MONt .............................. .............................. ..............................
MONh .............................. ..................Y........... ..............................
BIFd .............................. ..................T........... ..............................
BIFt .............................. ..................T........... ..............................
VIRd .............................. .............................. ..............................
VIRt .............................. .............................. ..............................
OCC  .............................. .............................. ..............................
WRI1 CG............................ .............................. ..............................
WRI2 CG............................ .............................. ..............................
CES1 .............................. .............................. ..............................
CES2 .............................. .............................. ..............................
CES3 .............................. .............................. ..............................
PYR  ....GT...A.................... .............................. ..............................
AGO  C............................. .............................. ..............................
MIC  ..C.C....T.................... .............................. ..............................
STE  ....G....T......A............. .............................. ..............................
LAC  CAG.T.............G........... ..................T.......A... ..................A...........
```

```
     333333333333333333333333333333 333333333344444444444444444444 444444444444444444444444444444
     666666666777777777788888888889 999999999000000000011111111112 222222222333333333344444444445
     123456789012345678901234567890 123456789012345678901234567890 123456789012345678901234567890

DAN1 TTGGTGTGAATTGCAGAATCCCGTGAACCA TCGAGTTTTTGAACGCAAGTTGCGCCCAAA GCCATCCGGYYGAGGGCACGCCTGCCTGGG
DAN2 .............................. .............................. ..............................
MONt .............................. .............................. ..............................
MONh .............................. .............................. ..............................
BIFd .............................. .............................. .........CT...................
BIFt .............................. .............................. .........CT...................
VIRd .............................. .............................. .........CC...................
VIRt .............................. .............................. .........CC...................
OCC  .............................. .............................. ..............................
WRI1 .............................. ...........................R.. ..............................
WRI2 .............................. ...........................R.. ..............................
CES1 .............................. .............................. .........CC...................
CES2 .............................. .............................. .........CC...................
CES3 .............................. .............................. .........CC...................
PYR  .............................. .............................. .........TT.........T.........
AGO  .............................. ...........................G.. .........CT...................
MIC  .............................. .............................. .........CT...................
STE  .............................. .............................. .....T...AA..............G....
LAC  .............................. ...........................G.. .........CT...................
```

```
     5.8s <--|--> ITS2
     444444444444444444444444444444 444444444444444444455555555555 555555555555555555555555555555
     555555555566666666667777777778 888888888999999999900000000001 111111111222222222233333333334
     123456789012345678901234567890 123456789012345678901234567890 123456789012345678901234567890

DNA1 CGTCACGCATCGCTCGCCCCCAACCATGCA TCC-TTATGGGATGCTTGGCATCGGGGCGG AGATTGGCCTCCCGTGCCTTTGGTGTGGTT
DAN2 .............................. .............................. ..............................
MONt .............................. .T...C........................ ..............................
MONh .............................. .T...C.............Y.......... .K...............Y............
BIFd .............................. .T...C.............T.......... .T...............T............
BIFt .............................. .T...C.............T.......... .T...............T............
VIRd .............................. .............................. ..............................
VIRt .............................. .............................. ..............................
OCC  ...........R.................. .T.............C.............. .T..............MT.....CA.....
WRI1 ...........R.................. .T....................T....... .T...............T.....CA.....
WRI2 ...........R.................. .T....................T....... .T...............T.....CA.....
CES1 .............................. .T............................ .T...............T.....CA.....
CES2 .............................. .T............................ .T...............T.....CA.....
CES3 .............................. .T............................ .T...............T.....CA.....
PYR  ...........R..............A... .......C..............T....... .................T.......C....
AGO  ...........R.........C........ .T.C....A.....AGC...T.T....... .......T.........T..GT...C....
MIC  ...........A..T......C........ ...TA.T.......A............... .T.....A......................
STE  .............................. ...CC..AA..........TG......... .........................C....
LAC  .................T...C.....A.C ...CC..C..TTG..A...TG.T....... .T.A.............T.G..T.TC....
```

Fig. 2

```
     555555555555555555555555555555 555555555555555555555555555556 666666666666666666666666666666
     444444444555555555566666666667 777777777888888888899999999990 000000000111111111122222222223
     123456789012345678901234567890 123456789012345678901234567890 123456789012345678901234567890

DAN1 GGCCTAAATCGGAGT-ACCTTCGGTGGACG CACGACTAGTGGTGGTTGAAAAGACCCTCG TCCTGTGTTGTGCGTCGTAAGCTGTGAGGG
DNA2 .............................. .............................. ..............................
MONt .............................. .............................. ..............................
MONh ...............C.............. .............................. ..Y.........Y.................
BIFd ...............C.............. .............................. ..T.........T.................
BIFt ...............C.............. .............................. ..T.........T.................
VIRd .............................. .............................. ..............................
VIRt .............................. .............................. ..............................
OCC  ...............C...........G.. ....G......................... ..T................T..........
WRI1 ...............C............A. .T..G......................... ..T................T..C.......
WRI2 ...............C............A. .T..G......................... ..T................T..C.......
CES1 ..........A....C...........G.. ....G......................... ..T...........C....T..........
CES2 ..........A....C...........G.. ....G......................... ..T...........C....T..........
CES3 ..........A....C...........G.. ....G......................... ..T...........C....T..........
PYR  ........CT.....C--............ ....................T......... ..T.A..........T.............T
AGO  .T......CATA...C.....T........ ...A................T......... ..T.T.................C.......
MIC  .T......C.T....C.....G...T.... ....................T......... ....AA.G..............C..A....
STE  ........CTA................... ...A................T......... ..T...........................
LAC  .........AA.....T.......C..... ....................T......... ..T.T.....C.T.....G......A....

                                              ITS2 <-|-> 25S
     666666666666666666666666666666 666666666666666666666666666666 6666666667777777777777777777
     333333333444444444455555555556 666666666777777777788888888889 9999999990000000000111111111
     123456789012345678901234567890 123456789012345678901234567890 1234567890123456789012345678

DNA1 AGGCCCTTCATGAAGACCCCAATGTGGTGT CTTGCGACGATGCTTCGACCGCGACCCCAG GTCAGCGGGGACTACCCGCTGAGTTTAA
DAN2 .............................. .............................. ............................
MONt .............................. .............................. ............................
MONh .............................. ...........R.................. ............................
BIFd .........................-.... ...........A.................. ............................
BIFt .........................-.... ...........A.................. ............................
VIRd .............................. .............................. ............................
VIRt .............................. .............................. ............................
OCC  ........G.W..............-.... .............................. ............................
WRI1 ........G.W..............-.... .......YT.........Y........... ............................
WRI2 ........G.W..............-.... .......YT.........Y........... ............................
CES1 ........G................-.... .............................. ............................
CES2 ........G................-.... .............................. ............................
CES3 ........G................-.... .............................. ............................
PYR  G.......G...........T....ATC.. .A..T......................... ............................
AGO  .T.A.T..G..........T..C...TC.. .C.TG......................... ............................
MIC  .....T..G---..........C..-.A.. .............................. ............................
STE  ........GT...............-.... .C.A...........G.............. ............................
LAC  TA.....CATCA........AT...ATC.. ...CG..T...................... ............................
```

Fig. 2. Aligned sequences of the internal transcribed spacer region from 19 taxa of *Krigia* and outgroups. Complete sequences between two amplifying primers (P 1 and P 4 in Fig. 1) are aligned to maximum identity with minimal gap adjustments. Nucleotide positions start at the end of amplifying primer P 1. Dots (.) indicated matched sequences to the first taxon and bars (−) indicated the gaps. Species abbreviations are given in Table 1. Polymorphic sites are represented by following symbols: R = A or G, Y = C or T, M = A or C, K = G or T, W = A or T, S = C or G

A total of 13 ambiguous polymorphic sites was detected from all other species. The majority (9/13) of these are ambiguous nucleotides within the purine (A or G) or the pyrimidine (C or T) bases. The other four are ambiguous between purine and pyrimidine bases.

Sequence divergence of the ITS region. Aligned sequences between the two amplifying primers (P 1 and P 4 in Fig. 1) included 718 positions. No sequence variation was observed for the first 32 bp (the 3′ end of 18 S rRNA coding region) and the last 38 bp (the 5′ end of 25 S rRNA coding region). These regions and all ambiguous gap sites were removed prior to the calculation of sequence divergence and phylogeny reconstruction. A total of 238 variable sites was detected from the ITS

regions. More than half (55.9%) of these variable sites (142/238) were in ITS 1 compared to 36.6% (87/238) in ITS 2 and only 3.8% (9/238) in 5.8 S rDNA.

Sequence divergences among 12 taxa in ITS 1, ITS 2, and whole ITS regions are given in Table 3.

Phylogenetic analyses of ITS sequence. A total of 262 variable sites was detected, 130 of which were phylogenetically informative. The distribution of the number of informative positions by region was 60.8% (79/130) for ITS 1, 36.2% (47/130) for ITS 2, and 3.0% (4/130) for 5.8 S.

A parsimony analysis of the ITS sequences of 18 taxa (without the hybrid *K. montana*) generated a single most parsimonious tree with a length of 405 steps (including autapomorphies), a consistency index of 0.62 (excluding autapomorphies), and a retention index of 0.71. The hybrid taxon, hexaploid *K. montana*, was added by hand to this tree (Fig. 3). Inclusion of a hybrid taxon to the data matrix resulted in six equally most parsimonious trees with the same branch lengths (trees

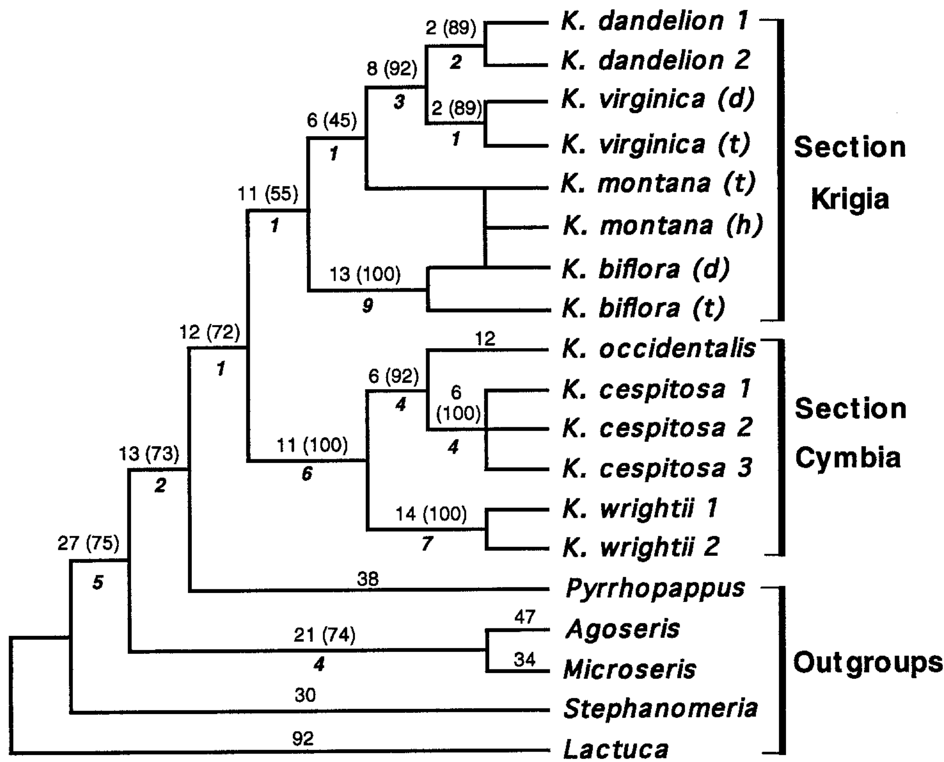
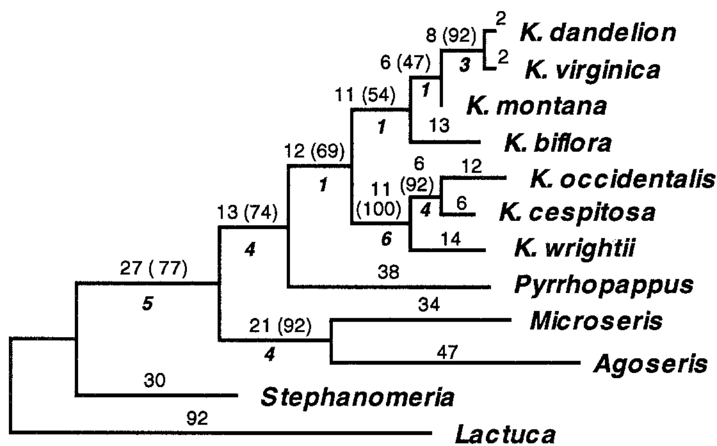
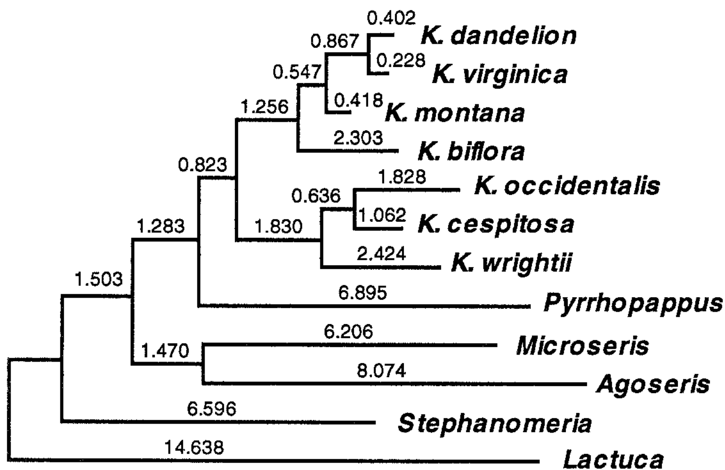


Fig. 3. A single most parsimonious tree of *Krigia* and outgroups. All gap sites were treat as missing characters. The numbers without parentheses above (arabic) and below (italic) the node indicate the number of mutations and the decay index, respectively. The numbers in parentheses represent the percentage of times that a monophyletic group occurred in 1,000 bootstrap replicates. The hybrid taxon (hexaploid *K. montana*) between tetraploid *K. montana* and diploid *K. biflora* was added to the tree by hand after the PAUP analysis. The symbol in parentheses after species name indicate diploid (*d*), tetraploid (*t*), and hexaploid (*h*)

not shown). The strict consensus tree had one unresolved branch at the basal node of *K. biflora* and *K. montana* (cf. McDade 1992).

Decay analyses were continued until all nodes collapsed to assess the robustness of each clade. At tree lengths of 406, 407, 408, 409, 410, 411, and 412 steps, there were 7, 15, 15, 15, 52, 147, and 316 equally parsimonious trees, respectively. All internal nodes collapsed at nine steps longer than the most parsimonious tree. Bootstrap values of internal nodes ranged from 45% to 100%. Exclusion of multiple populations from the data matrix did not change either the tree topology or the number of steps supporting internal nodes, however, it generated slightly different decay indices and bootstrap values (Fig. 4). Similar analyses using only the ITS 1 and ITS 2 (without 5.8 S) sequences also generated the same tree topology as that of full ITS data set.





Figs. 4–5. Phylogenetic trees of seven species of *Krigia* and five outgroup genera using ITS sequence data. – Fig. 4. A single most parsimonious tree from reduced data set. Exclusion of multiple populations from data matrix (Fig. 2) resulted the same tree topology compared to full data set (see Fig. 3 for more detailed tree description). – Fig. 5. Neighbor-joining tree showing the relative distances among branches. Relative distances of the nodes and branches are given as 100 × value

Outgroup selection also did not change the tree topology for ingroup taxa. *Pyrrhopappus* was identified as a sister genus to *Krigia* using *Lactuca* as an outgroup. Inclusion or exclusion of any of the five outgroup genera did not affect the tree topology in *Krigia* if *Pyrrhopappus* was included. *Agoseris* and *Microseris* form a monophyletic clade among outgroup genera. *Stephanomeria* of the subtribe *Stephanomeriinae* was positioned as a sister genus to the subtribe *Microseridinae*.

The two sections of genus *Krigia*, *Krigia* and *Cymbia*, form monophyletic lineages. Within sect. *Cymbia*, *K. cespitosa* and *K. occidentalis* form a monophyletic clade and their sister species was *K. wrightii*. Within sect. *Krigia*, the perennial species *K. dandelion* and the annual species *K. virginica* form a terminal monophyletic clade. A closely related species pair, *K. biflora* and *K. montana*, remained as independent lineages rather than forming a clade.

Neighbor joining analyses also generated a tree topology identical to the parsimony tree (Fig. 5). Evolutionary distances of internal nodes and terminal branches were largely heterogeneous.

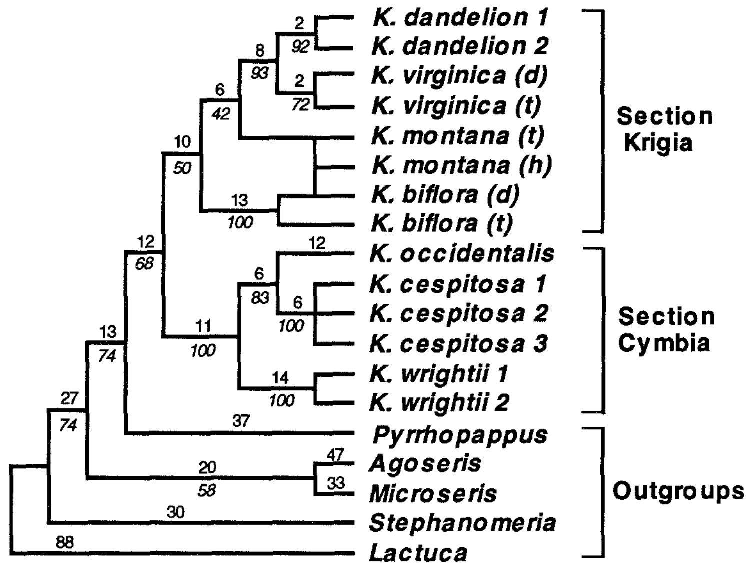
Gap treatments and tree topology. A total of 25 independent insertion/deletion mutations was observed from the aligned ITS data set (Fig. 2). Thirteen of these are potentially phylogenetically informative (coordinates 55, 106, 195, 200, 208, 223, 224, 225, 261, 275, 484, 556, and 656 in Fig. 2).

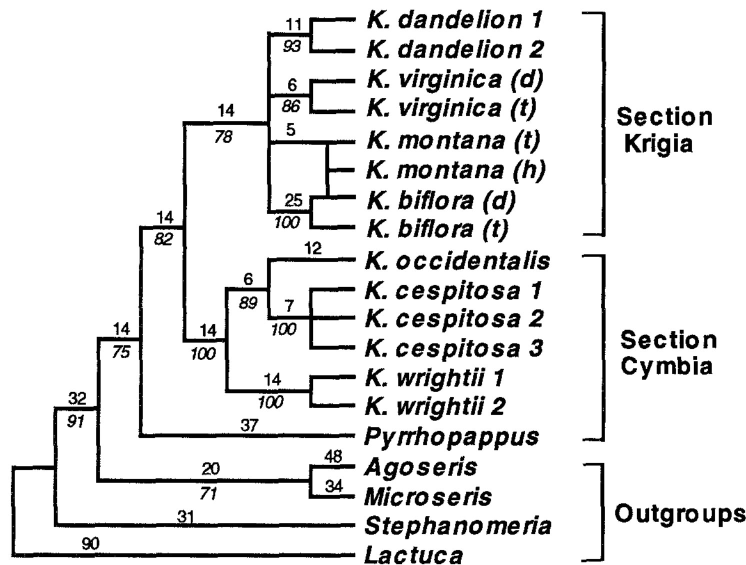
To investigate the possible impact of gaps on phylogenetic tree construction, we treated the gaps in three different ways: 1) gaps were simply coded as missing characters; 2) gap regions were excluded; and 3) gap areas were excluded and were coded as unweighted binary characters (present or absent).

The previous section described the results of phylogenetic analyses using the first option (all gaps as missing characters). Removing gap areas from the data matrix (the second option) generated a single most parsimonious tree with a length of 397 (including autapomorphies), a consistency index of 0.62 (excluding autapomorphies), and a retention index of 0.71 (Fig. 6). The tree topology, CI, and RI (Fig. 6) were identical to that of the first option (Fig. 3).

In the third option, all 13 gap areas were removed and 13 binary characters were added to the data matrix. The homology of three (coordinates 223, 224, and 225) of 13 characters is somewhat ambiguous. This data matrix contains 705 aligned nucleotide sites and 13 binary characters (data not shown). Phylogenetic analyses generated two most parsimonious trees with a length of 424 steps (including autapomorphies), a consistency index of 0.61 (excluding autapomorphies), and a retention index of 0.71 (Fig. 7). The two trees show different topologies for members of the sect. *Krigia*. One of the trees has the identical topology to that of the first and second options (Figs. 3 and 6). However, the second tree has a novel ITS topology showing the [(*K. dandelion*, *K. virginica*) (*K. montana*, *K. biflora*)] relationships, which was not observed in trees from any other type of data, including cpDNA (KIM & al. 1992 a) and morphology (KIM & TURNER 1992). The strict consensus tree of the two trees shows totally unresolved relationships among the members of sect. *Krigia* (Fig. 7).

The ambiguity in gap coding in *Krigia* argued for deleting all potentially synapomorphic gap areas (the second option) from the data matrix in subsequent phylogenetic comparisons.





Figs. 6–7. Phylogenetic trees of ITS data based on the different gap treatments. The arabic number above and the italic number below the node indicate the number of mutations and the bootstrap percentage from 1,000 replicates, respectively. The symbol in parentheses after species name indicate diploid (*d*), tetraploid (*t*), and hexaploid (*h*). – Fig. 6. A single most parsimonious tree of *Krigia* and related genera. All synapomorphic gap areas were excluded from data matrix prior to tree construction. The tree has identical topology, consistency index and retention index compare to that of the Fig. 3. – Fig. 7. The strict consensus tree from two equally parsimonious trees. All synapomorphic gap areas were removed and added as a binary characters (present or absent) to the data matrix

Phylogenetic analyses of combined data. To simplify the individual data sets, all multiple populations of the same species were removed and *Pyrrhopappus* was selected as a single outgroup. The reduced data sets have a substantially smaller

number of synapomorphies than the full data sets (Table 4); nevertheless, parsimony analyses of each reduced data set showed tree topologies identical to the full data set (Fig. 8 A–D). There are 10,395 possible trees in the reduced unrooted eight taxon analyses using exhaustive search options. All data sets showed strong skewness of tree-length distribution, indicating that there is strong phylogenetic signal in all reduced data sets (Table 4). The skewness of the tree-length distribution varies from $g_1 = -1.112$ in cpDNA data to $g_1 = -0.639$ in ITS data.

The combined morphology, cpDNA, nrDNA restriction site, and ITS sequence data set contained 205 synapomorphic characters, including 18 morphological characters, 47 ITS nucleotide positions, 12 nrDNA restriction sites, and 128 cpDNA restriction sites. Phylogenetic analyses of the combined data resulted in a single most parsimonious tree with 250 steps and a consistency index of 0.83 (Fig. 8 E). The combined data show a strong skewness of the tree-lengths distribution ($g_1 = -0.979$). The tree from the combined data had the same two monophyletic sections, *Krigia* and *Cymbia*, that were present in trees from all individual data sets (Fig. 8). Within sect. *Cymbia*, there was good congruence between the combined tree and each of the individual trees. Within sect. *Krigia*, two monophyletic species pairs were identified. The *K. dandelion-K. virginica* clade in the combined data was present in the ITS sequence and nrDNA restriction site trees. However, this group was not supported by the morphological or the cpDNA data. In contrast, the *K. biflora-K. montana* clade in the combined data occurred in both the cpDNA and morphological trees. This group was not supported in either the ITS or the nrDNA trees. Thus, both conflict and congruence were observed between the combined tree and each of the individual trees.

The combinable consensus tree topology among four individual trees collapsed at the base of sect. *Krigia* (Fig. 8 F). The combinbed tree topology using the character re-coding method (DOYLE 1992) generated two equally parsimonious trees. Their strict consensus tree was identical to the combinable consensus tree (Fig. 8 F).

Character incongruence within/among data sets. Different degrees of character incongruence were observed in each of four data sets (Table 4). ITS and rDNA data show more character conflict than morphological and cpDNA data. When all data sets are combined there are 45 extra steps on the best fitting tree ($R = 205$, $L = 250$, Table 4, Fig. 8 E), which is defined as total character incongruence. Nearly half (21/45) of the character incongruence in the combined data set originated from ITS data set (Table 4). Another 18 extra steps were required for the rDNA, morphological, and cpDNA data. These 39 extra steps explain the character incongruence within the individual data sets, which leaves six additional steps originating from incongruence among the four data sets (Table 4). As a result, 13.3% (6/45) of the total character incongruence is due to the disparity among four data sets (I_{MF} value of MICKEVICH & FARRIS 1981, Table 5). The I_{MF} index may underestimate the character incongruence among data sets (SWOFFORD 1991 a). In particular, the value will generate an inaccurate estimate of character incongruence when the number of characters in each data set is unbalanced, which is the case in four *Krigia* data sets (Table 4).

As an alternative method MIYAMOTO suggested a procedure for partitioning between from within data set incongruence (see SWOFFORD 1991 a, I_M index). In order to calculate the I_M index, each data set was mapped onto the trees computed

Table 3. Sequence divergences of ITS region from *Krigia* species and outgroup genera. Observed numbers of nucleotide differences (pairwise comparisons) are shown above the diagonal and sequence divergences corrected for multiple substitutions using two-parameter method (KIMURA 1980) are given below the diagonal. Multiple populations for the same species are not included in this table. Species abbreviation are given in Table 1

A. ITS 1 distance matrix

	DAN	MON	BIF	VIR	OCC	WRI	CES	PYR	AGO	MIC	STE	LAC
DAN	–	8	19	4	31	28	28	42	50	45	48	66
MON	0.0327	–	11	8	30	27	24	38	50	44	48	68
BIF	0.0806	0.0454	–	16	34	32	27	41	53	47	53	71
VIR	0.0161	0.0327	0.0674	–	30	27	25	37	49	42	51	68
OCC	0.1350	0.1319	0.1519	0.1319	–	23	15	43	57	56	58	76
WRI	0.1227	0.1190	0.1437	0.1194	0.0991	–	17	41	49	50	58	74
CES	0.1198	0.1026	0.1169	0.1074	0.0621	0.0716	–	36	54	49	55	72
PYR	0.1897	0.1708	0.1869	0.1670	0.1946	0.1871	0.1578	–	55	62	71	83
AGO	0.2383	0.2413	0.2594	0.2359	0.2804	0.2355	0.2592	0.2642	–	46	58	67
MIC	0.2054	0.2028	0.2198	0.1926	0.2677	0.2363	0.2264	0.2999	0.2182	–	63	79
STE	0.2217	0.2225	0.2507	0.2391	0.2799	0.2838	0.2600	0.3548	0.2900	0.3096	–	67
LAC	0.3308	0.3477	0.3666	0.3463	0.4012	0.3910	0.3725	0.4458	0.3525	0.4226	0.3411	–

B. ITS 2 distance matrix

	DAN	MON	BIF	VIR	OCC	WRI	CES	PYR	AGO	MIC	STE	LAC
DAN	–	2	8	0	11	14	12	20	35	23	17	45
MON	0.0090	–	6	2	11	14	12	22	35	25	19	47
BIF	0.0368	0.0274	–	8	11	14	12	23	37	29	21	42
VIR	0.0000	0.0090	0.0368	–	11	14	12	20	35	23	17	45
OCC	0.0513	0.0513	0.0510	0.0513	–	7	3	24	37	29	24	46
WRI	0.0664	0.0664	0.0660	0.0664	0.0324	–	8	25	37	30	27	47
CES	0.0559	0.0559	0.0555	0.0559	0.0136	0.0370	–	25	38	30	23	46
PYR	0.0963	0.1065	0.1118	0.0963	0.1173	0.1236	0.1221	–	33	32	22	21
AGO	0.1768	0.1768	0.1883	0.1768	0.1889	0.1908	0.1940	0.1670	–	34	32	52
MIC	0.1141	0.1247	0.1452	0.1141	0.1457	0.1526	0.1506	0.1648	0.1741	–	32	52
STE	0.0804	0.0903	0.1005	0.0804	0.1163	0.1330	0.1106	0.1072	0.1607	0.1618	–	41
LAC	0.2355	0.2474	0.2187	0.2355	0.2435	0.2521	0.2423	0.2150	0.2765	0.2861	0.2117	–

C. Combined ITS 1 and ITS 2 distance matrix

	DAN	MON	BIF	VIR	OCC	WRI	CES	PYR	AGO	MIC	STE	LAC
DAN	–	10	27	4	42	42	40	62	85	68	56	111
MON	0.0214	–	17	10	41	41	36	60	85	69	67	115
BIF	0.0595	0.0369	–	24	45	46	39	64	90	76	74	113
VIR	0.0085	0.0214	0.0527	–	41	41	37	57	84	65	68	113
OCC	0.0946	0.0928	0.1024	0.0928	–	30	18	67	94	85	82	122
WRI	0.0956	0.0936	0.1057	0.0938	0.0669	–	25	66	86	80	85	121
CES	0.0892	0.0802	0.0872	0.0827	0.0390	0.0551	–	61	92	79	78	118
PYR	0.1444	0.1398	0.1505	0.1328	0.1574	0.1566	0.1409	–	88	94	93	124
AGO	0.2085	0.2098	0.2245	0.2071	0.2355	0.2139	0.2275	0.2168	–	80	90	119
MIC	0.1616	0.1653	0.1837	0.1549	0.2082	0.1960	0.1900	0.2344	0.1970	–	95	131
STE	0.1518	0.1571	0.1759	0.1600	0.1981	0.2086	0.1858	0.2291	0.2254	0.2366	–	108
LAC	0.2841	0.2982	0.2928	0.2916	0.3223	0.3218	0.3078	0.3285	0.3147	0.3551	0.2767	–

D. ITS region (ITS 1, ITS 2, and 5.8 S rDNA) distance matrix

	DAN	MON	BIF	VIR	OCC	WRI	CES	PYR	AGO	MIC	STE	LAC
DAN	–	10	28	4	42	42	40	63	86	68	69	115
MON	0.0160	–	18	10	41	41	36	61	86	69	71	119
BIF	0.0457	0.0290	–	26	46	47	41	67	92	77	79	116
VIR	0.0063	0.0160	0.0423	–	41	41	37	60	86	66	72	118
OCC	0.0695	0.0681	0.0768	0.0681	–	30	18	68	95	85	86	126
WRI	0.0701	0.0686	0.0791	0.0687	0.0493	–	25	67	85	80	89	124
CES	0.0657	0.0591	0.0675	0.0608	0.0289	0.0408	–	64	94	80	82	123
PYR	0.1070	0.1036	0.1146	0.1019	0.1161	0.1154	0.1080	–	91	96	98	130
AGO	0.1515	0.1522	0.1639	0.1519	0.1698	0.1530	0.1664	0.1608	–	81	95	122
MIC	0.1170	0.1194	0.1340	0.1137	0.1494	0.1408	0.1387	0.1709	0.1428	–	99	135
STE	0.1177	0.1215	0.1366	0.1235	0.1503	0.1574	0.1419	0.1738	0.1705	0.1768	–	116
LAC	0.2094	0.2187	0.2124	0.2159	0.2345	0.2319	0.2267	0.2425	0.2266	0.2552	0.2123	–

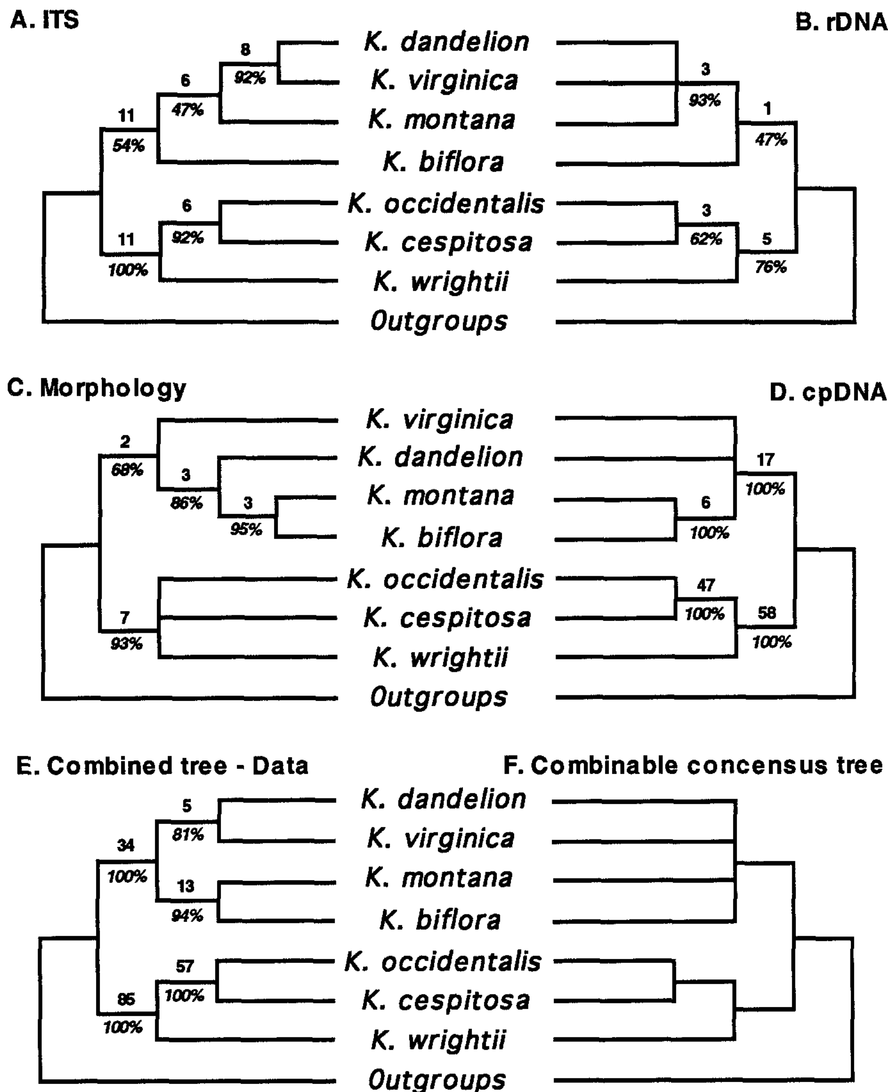


Fig. 8. Maximum parsimony trees showing the relationships among seven species of *Krigia* from various data sets. Numbers above the node indicate the numbers of changes supporting the nodes and the percentages below the node represent the bootstrap values. *A* ITS sequence data (this paper). *B* Nuclear ribosomal DNA restriction site data (Kim & Mabry 1991). *C* Morphological and cytological data (Kim & Turner 1992). *D* Chloroplast DNA restriction site data (Kim & al. 1992 a). *E* Tree derived from combined data set of A, B, C, and D. *F* Combinable consensus tree topology of four trees (A–D). The strict consensus tree from character recoding analyses (Doyle 1992) generated the identical topology to the tree F

for the other data sets and the number of extra steps than minimum numbers of synapomorphies was calculated from all possible combinations of data sets (t value in Table 5). The within-data-set incongruence was also calculated from all possible combinations of two data sets (s value in Table 5). Finally, the total character incongruence was calculated by adding all possible combinations among four data

Table 4. Summary of parsimony analyses on *Krigia* data sets. All information is derived from reduced data sets unless otherwise indicated. Only synapomorphies were used in parsimony analyses

	ITS	rDNA (IGS)	Morphology	cpDNA	Combined
Minimum number of synapomorphies (full data sets)	130	28	24	159	341
Minimum number of synapomorphies (R) (reduced data sets)	47	12	18	128	205
Length of most parsimonious trees (L)	68	17	23	136	250
Number of extra steps (s)	21	5	5	8	45
Number of equally parsimonious trees	1	1	3	1	1
Consistency indexes (CI)	0.735	0.707	0.783	0.941	0.832
Retention indexes (RI)	0.766	0.737	0.844	0.963	0.878
Skewness of tree-length distribution (g_1 value)	−0.639	−0.804	−0.805	−1.112	−0.979
Figure illustrating consensus tree	8 A	8 B	8 C	8 D	8 E

Table 5. Character incongruence among *Krigia* data sets. Incongruence between data sets: *e* number of extra steps than most parsimonious tree required to change the tree topology from row to column, incongruence within data set: *s* number of extra steps than minimum numbers of synapomorphies (see Table 4) required to change the tree topology from row to column, t = e + s, *M* MIYAMOTO incongruence index (I_M value, SWOFFORD 1991 a), *F* MICKEVICH & FARRIS (1981) incongruence index (I_{MF} value)

From To	ITS	rDNA (IGS)	Morphology	cpDNA	Combined
ITS	– s = 21 –	e = 0 t = 5 M = 0.000	e = 5 t = 10 M = 0.277	e = 6 t = 14 M = 0.194	e = 5 t = 50 M = 0.083
rDNA	e = 0 t = 21 F = 0.000	– s = 5 –	e = 5 t = 10 M = 0.444	e = 6 t = 14 M = 0.409	e = 5 t = 50 M = 0.138
Morphology	e = 5 t = 26 F = 0.103	e = 3 t = 8 F = 0.300	– s = 5 –	e = 0 t = 8 M = 0.000	e = 2 t = 47 M = 0.074
cpDNA	e = 1 t = 22 F = 0.033	e = 3 t = 8 F = 0.188	e = 0 t = 5 F = 0.000	– s = 8 –	e = 0 t = 45 M = 0.000
Combined	e = 1 t = 22 F = 0.015	e = 3 t = 8 F = 0.057	e = 2 t = 7 F = 0.038	e = 0 t = 8 F = 0.000	– s = 45 –

sets using the following formula: $I_M = [(26 + 22 + 10 + 14 + 8 + 8 + 10 + 14 + 21 + 5 + 5 + 8) - (21 + 5 + 8 + 21 + 5 + 5 + 8 + 5 + 21 + 5 + 5 + 8)]/(26 + 22 + 10 + 14 + 8 + 8 + 10 + 14 + 21 + 5 + 5 + 8) = (151-117)/151 = 0.225$.

The average I_M among all possible combinations of two data sets was 0.221. Thus, the average I_M may also represent the I_M index of the combined data (see M value in Table 5). Almost twice as much character incongruence was identified using the I_M index (22.5%) than the I_{MF} index (13.3%). Up to 44% character incongruence was observed between rDNA restriction site data and morphological or cpDNA data. ITS data also showed substantial character incongruence with morphological and cpDNA data. However, no character incongruence was observed between either IGS and ITS data, or cpDNA and morphological data. The combined data also showed less character incongruence with cpDNA and morphological data than with ITS and rDNA data (Table 5).

Topological constraint analyses from one data set to another tree also show the degree of conflict between the two data sets (Table 5). However, the number of extra steps required to change the tree topology from one data set to the other was often asymmetric; for instance, six steps were required from cpDNA data on to ITS data, while only one step was needed from ITS data on to cpDNA tree. No extra steps were required between either cpDNA and morphological data, or ITS and IGS data. In contrast, several steps were required between the two nrDNA data sets and morphological/cpDNA data sets.

Discussion

Evolution of ITS region. The length of ITS 1 in the *Lactuceae* (*Asteraceae* subfam. *Cichorioideae*) is longer than that of ITS 2; the differences ranged from 24 bp to 35 bp. Similar length variation was also reported from the subtribe *Madiinae* (subfam. *Asteroideae*) of *Asteraceae* (BALDWIN 1992) and 20 other genera of tribes *Heliantheae* and *Eupatorieae*, including *Helianthus, Polymnia, Eupatorium, Helenium, Encelia, Gaillardia, Psilostrophe, Coreopsis, Hymenoxys*, etc. (K.-J. KIM & R. JANSEN, unpubl.).

The results from two subfamilies of *Asteraceae* indicated that ITS 1 is probably longer than ITS 2 in all members of the family. This is not always the case in other flowering plant families. For example, ITS 1 (235 bp) is 27 bp longer than ITS 2 (208) in *Vicia faba* (*Fabaceae*) while ITS 2 (220) is 15 bp longer than ITS 1 (205) in *Vigna radiata* of the same family (Table 2). Both ITSs show similar size in two genera (*Nicotiana* and *Lycopersicon*) of *Solanaceae* and two genera (*Avena* and *Hordeum*) of *Poaceae*. However, ITS 1 (194) is 39 bp shorter than ITS 2 (233 bp) in a third genus (*Oryza*) of the *Poaceae*. The most extreme length difference (77 bp) is reported from *Sinapsis alba* of the *Brassicaceae* (Table 2).

In the *Lactuceae*, length variation is more common in ITS 1 (20 variants) than in ITS 2 (5 variants) (Fig. 2). Most of the length variants are one or two base-pair insertions/deletions and occur at single direct repeats such as AA(A)A (positions 46 and 224 in Fig. 2), CC(CC)C (position 250 in Fig. 2) and G(G) (positions 195, 275, and 656). In addition, length mutations are also located in short direct and inverted repeats. The majority of these length variants are probably due to slippage during DNA replication (LEVINSON & GUTMAN 1987).

Short insertion/deletion mutations are the primary source of length variation in *Asteraceae* (this study and BALDWIN 1992). However, this is not in the case in other families. Length variation up to 200 bp was reported from nine genera of *Fabaceae* (JORGENSEN & CLUSTER 1988). Differences of up to 100 bp were also reported among individuals within a population of *Lisianthus skinneri* (*Gentianaceae*) (SYTSMA & SCHAAL 1990). These large insertions/deletions may be due to unequal crossing over of rDNA repeat units (SMITH 1976). Sequence information from both groups may help to understand the origin of large length variants. Length mutations also occur frequently at different positions even though they have the same length.

ITS 1 changes faster than ITS 2 not only in length but also in sequence. ITS 1 sequences evolve two times faster than ITS 2 (average sequence divergence 9.6% vs. 4.3%, Table 3) in *Krigia*. Up to 44.6% of the sequence difference between *Pyrrhopappus* and *Lactuca* was in ITS 1, whereas, the difference was only 21.5% in ITS 2. Differential evolutionary rates between ITS 1 and ITS 2 were also reported from the *Madiinae* group (BALDWIN 1992), even though the author did not mention this fact. In this case, ITS 1 changes approximately 1.5 times faster than ITS 2. The reason for different levels of sequence conservation between the two ITS is unclear. We can only speculate that the two ITS sequences are under different evolutionary constraints. Comparative sequencing from other closely related plant groups may help to resolve this question. The high sequence divergence, along with high levels of homoplasy in ITS 1, among three subtribes of the *Lactuceae* raises serious concerns about the utility of the ITS 1 sequences in phylogeny reconstruction at higher taxonomic levels (above the subtribe level) in *Asteraceae*.

The combination of ITS 1, ITS 2, and 5.8 S sequences reduced significantly the overall divergence values (Table 3 D). However, combining the sequences does not diminish the problems associated with ITS 1. High sequence variability in the ITS regions may relate to their structural variability. Recent computer-simulated secondary structures of plant ITS sequences (VENKATESWARLU & NAZAR 1991) suggest that the detailed structure is not strongly conserved except for the core-like features, which may act to bring the mature termini into close proximity for splicing.

Two different ITS types in a hybrid plant. Restriction site analysis of entire nrDNA from *Krigia* demonstrated the existence of two different sizes of the repeating units from the hexaploid *K. montana* (KIM & MABRY 1991). Uniparentally inherited cpDNA data (KIM & al. 1992 b) together with biparentally inherited nrDNA data identified the tetraploid *K. montana* as the paternal genome donor and the diploid *K. biflora* as the maternal one. The presence of the 20 polymorphic nucleotide sites and the two length variants that distinguish *K. montana* and *K. biflora* in the hexaploid *K. montana* indicate that the two rDNA types occur in a single individual. The ITS results are concordant with earlier studies using nrDNA, cpDNA, cytology, and morphology (CHAMBERS 1965, KIM & TURNER 1992). We have no information about a possible genetic recombination between the two types of nrDNA repeating units in the hybrid genome because our ITS data are based on the direct sequencing of amplified DNA. Sequencing of several individual molecules through cloning of PCR product may provide an understanding of the concerted evolution of the nrDNA repeat in this hybrid polyploid.

The lack of length and sequence variation in the ITS region of the tetraploids of *K. biflora* and *K. virginica* compared with their diploid progenitors support previous hypotheses of their autpolyploid origin (CHINNAPPA 1981, KIM & TURNER 1992). In contrast to nrDNA data, large numbers of cpDNA mutations were accumulated in the tetraploid populations (KIM & al. 1992 b). The contrasting results between cytoplasmic and nuclear genomes may indicate either cytoplasmic gene flow without detectable nuclear gene flow (DORADO & al. 1992) or the biased gene conversion of nrDNA toward to the diploid rDNA repeat type (HILLIS & al. 1991).

Phylogenetic comparisons among data sets. The phylogenetic tree obtained from the ITS sequences (Figs. 3–5) has complete resolution of species relationships (Fig. 8 A) and is congruent with the tree based on nrDNA restriction site data (Fig. 8 B). The ITS tree also agrees well with morphological, cytological, and cpDNA data with respect to three clades: 1) *Pyrrhopappus* is a sister genus of *Krigia*; 2) both sections *Krigia* and *Cymbia* are monophyletic; and 3) *Krigia wrightii* is a sister species to the *K. cespitosa-K. occidentalis* group.

The two most recent classifications (STEBBINS 1953, JEFFREY 1966) of tribe *Lactuceae* recognized two generic groups within the subtribe *Microseridinae*. Both systems group together the two North American genera, *Krigia* and *Pyrrhopappus*, and a South American genus, *Picrosia*, based on several morphological and cytological features (STEBBINS & al. 1953). Pollen morphology provided the most conclusive evidence for the close relationships between *Krigia* and *Pyrrhopappus* because these genera share three unique features, including absence of polar thickening, symmetrical polar lacunae and narrow parapolar ektexine (FEUER & TOMB 1977, NORTHINGTON 1974, KIM & TURNER 1992). Although these shared pollen characters are unusual in *Microseridinae*, the sister group relationship of these two genera was not accepted widely for two reasons. First, similar pollen types also occur sporadically in other subtribes of *Lactuceae* (BLACKMORE 1981). Second, some workers emphasized the distinctive position of *Pyrrhopappus* from all other genera of *Microseridinae* because of the occurrence of chalcone compounds (NORTHINGTON 1974, HARBORNE 1977). However, two chalcone compounds were identified recently from *K. dandelion* (K.-J. KIM & T. MABRY, unpubl.). Chloroplast DNA restriction site data also supported the monophyly of the two genera (JANSEN & al. 1991). Our ITS tree identified *Pyrrhopappus* as a sister genus of *Krigia* using *Lactuca* as an outgroup. Thus, ITS data are in agreement with morphological, cytological, chemical, and cpDNA data for the sister group relationship of *Krigia* and *Pyrrhopappus*.

The two sections of *Krigia* (*Cymbia* and *Krigia*) are readily distinguished from each other by several morphological and cytological features (SHINNERS 1947, KIM & TURNER 1992). Some earlier workers emphasized the distinctive morphological features of the two groups and treated them as separate genera (GRAY 1886, STANDLEY 1911). Section *Cymbia* is easily recognized by its reflexed phyllaries, a self-compatible breeding system and reflexed style branches. The section includes only three annual species: *K. cespitosa, K. wrightii*, and *K. occidentalis* (SHINNERS 1947, KIM & TURNER 1992). Section *Krigia*, which is characterized by erect phyllaries and a base chromosome number of 5, includes an annual species, *K. virginica*, and three perennial species, *K. montana, K. biflora*, and *K. dandelion*. Chloroplast

DNA data strongly support the monophyly of the two sections (KIM & al. 1992a). The monophyly of sect. *Cymbia* was always stronger than the sect. *Krigia* in both cpDNA and morphological trees. The distinctness of the two sections was also indicated by the high level of cpDNA sequence divergence. Similar results are also observed with the ITS sequence data. The average sequence divergence between the two sections was 11.5% in ITS 1 and 5.8% in ITS 2. This value is comparable to some generic differences in the subtribe *Madiinae* (BALDWIN 1992). More ITS sequence data from diverse plant groups may provide a more accurate estimate of the range of sequence divergence within and between genera. The ITS sequence data also show strong agreement with morphological and cpDNA data with respect to the monophyly of the two sections and the level of divergence between them.

Relationships among the three species of sect. *Cymbia* were unresolved in the morphological tree (Fig. 8 C) even though they are readily distinguished by achene morphology and chromosome number. All three species (*K. wrightii, K. occidentalis*, and *K. cespitosa*) have unique chromosome numbers in the genus (n = 9, 6, and 4, respectively). Both cpDNA and nrDNA restriction site data provided support for the sister species relationship of *K. occidentalis* and *K. cespitosa* (Fig. 8 B, D). ITS sequence data also supported the monophyly of the group comprising these two species. As expected, the combination of morphological data with other data set improved the resolution in sect. *Cymbia*.

Complete congruence of the ITS tree with other lines of evidence with respect to the three major clades within *Krigia* and to generic relationships demonstrate the utility of ITS sequence data for examining phylogenetic relationships at the intergeneric, intersectional, and interspecific levels.

Limits of phylogenetic utility of ITS sequence data. Despite significant congruence of the ITS sequence data with other data sets, the ITS tree is substantially different from morphological and cpDNA trees for the species relationships of sect. *Krigia*. The three perennial species in the section (*K. dandelion, K. montana*, and *K. biflora*) do not form a monophyletic lineage in the ITS phylogeny because the annual species, *K. virginica*, is nested within the perennial group. In contrast, the three perennial species form a strong monophyletic group in the morphological phylogeny. The perennial species share several derived characters, including many bristles and scales (more than 20), whereas the annual species (*K. virginica*) has only five scales and bristles (KIM & TURNER 1992). All three morphological characters which defined the perennial node are not related functionally with each other. Monophyly of the perennial species require three additional steps than the most parsimonious tree in ITS data.

Within the perennial clade, a sister species relationship between *K. biflora* and *K. montana* was firmly established by both morphological and cpDNA evidence. Cytological and breeding experiments further support the close relationships between these two species (CHAMBERS 1965, CHINNAPA 1981). However, ITS data do not indicate a sister group relationship between *K. biflora* and *K. montana* because the *K. virginica-dandelion* clade was nested within these two species. Monophyly of the *K. montana-biflora* clade requires only one additional step than the most parsimonious ITS tree. No conflict was observed between ITS data and IGS data even though they show slightly different degree of resolution. Conflict among data sets may due to homoplasy or different evolutionary constraints among data sets.

A higher level of homoplasy was observed for ITS (CI = 0.74) and nrDNA data (CI = 0.70), compared to morphological (CI = 0.78) and cpDNA data (CI = 0.94). As is often the case, the number of characters did not affect to levels of homoplasy in the four *Krigia* data sets, all of which have the same number of taxa (SANDERSON & DONOGHUE 1989). Most of the homoplasy in the two nrDNA-based data sets is concentrated in the two monophyletic nodes in sect. *Krigia*. Thus, the distribution of homoplasy may be one of causes of the different tree topologies in the two nrDNA phylogenies. The congruent nature of homoplasy between the two nrDNA-based data sets may reflect similar constraints on the evolution of the nrDNA repeat. This is not a surprise because concerted evolution is well known in nrDNA repeating units (DOVER 1982, HILLIS & al. 1991). In this situation, all nucleotide positions in the nrDNA repeating units are highly dependent on each other. The results from *Krigia* suggest that the nrDNA-based data may have an intrinsic weakness for the phylogenetic reconstruction in some plant groups. Our results also indicate that conflict between nrDNA and cpDNA data reported from several plant groups may not always be caused by the uniparentally inherited chloroplast genome.

Combining the data in *Krigia*. There has been substantial controversy about how different data sets should be combined in phylogenetic analyses (BULL & al. 1993). One possible method is the consensus approach, but this has not been widely accepted because it usually results in less resolved trees than the individual data sets. Of the several available consensus methods, combinable component analysis usually generates more resolved trees (BREMER 1990). Our combinable component consensus tree, however, showed unresolved relationships among four species of the sect. *Krigia* because of the conflict between nrDNA tree(s) and cpDNA/morphological tree(s).

Alternatively, tree topologies from each data set can be coded as an ordered multistate character or multiple binary characters (DOYLE 1992) and then the combined recoded data matrix can be re-analyzed using tree construction programs. This approach permits flexibility for weighting different trees. The combined tree from four *Krigia* trees using this approach resulted in a topology identical with that of the combinable component consensus (Fig. 8 F).

The most frequently used combining method is the direct combination of different data sets so that all of the characters are considered in tree construction (e.g., total evidence, MIYAMOTO 1985, HILLIS 1987, KLUGE 1989, DONOGHUE & SANDERSON 1992). Total evidence provides the best estimate of the phylogeny if different data sets are not in conflict with each other. If different data sets show significant conflict, however, total evidence may result in an unacceptable tree topology (SWOFFORD 1991 a, SHAFFER & al. 1991, BULL & al. 1993).

The phylogenetic tree of *Krigia* derived from combining the four data sets shows both congruence and incongruence with the trees generated from the individual data sets (Fig. 8). Unresolved relationships among the three species of sect. *Cymbia* in morphological data (Fig. 8 C) were fully resolved in all three molecular data sets. Thus, the direct combination of morphological and molecular data provided an improved resolution for the morphological phylogeny. Complementary resolution is best achieved in this case by combining the cpDNA and morphological data. Unresolved relationships among three branches of sect. *Krigia* in the cpDNA tree

and unresolved relationships among the three species of sect. *Cymbia* in the morphological tree were completely resolved by combining these two data sets.

Direct combination of the ITS data with morphological or cpDNA data shows a mosaic tree topology from individual trees for the species relationships of sect. *Krigia* (Fig. 8 E). The monophyly of the *K. montana-K. biflora* branch in the combined data set(s) is due to the morphological and/or cpDNA data. In contrast, the monophyly of the *K. dandelion-K. virginica* clade in the combined data is influenced primarily by the ITS data. The high degree (up to 44%) of character conflict between the nrDNA-based data and the cpDNA/morphological data indicate that the direct combination of the four data sets may result in an incorrect tree topology even though a combined data set shows the least conflict to each of individual data set (Table 5). We believe that the monophyly of *K. dandelion* and *K. virginica* in the combined data is unlikely and it may due to concerted evolution of the nrDNA repeat. Our data suggest that careful character analyses of each data set is necessary prior to combining the data sets.

Note added in proof:

The degree of (in)congruence between data sets can be calculated by the indices of I_{MF} (Mickevich & Farris 1981) or I_M (Swofford 1991 a). However, examples in both papers are only for calculations using two data sets. In the current paper, we applied the two methods to four data sets. Because of requests from other laboratories, below we provide calculations for the four data sets used in this paper.

I_{MF} = Between data set incongruence/Total incongruence
= [Total incongruence − Within data set incongruence]/Total incongruence
= [Number of extra steps in combined data (St) − Sum of number of extra steps from each data set]/St ------- (see Table 4)
$= [45 - (21 + 5 + 5 + 8)]/45$
$= [45 - 39]/45$
$= 0.133$

I_M $= [\sum t - \sum S]/\sum t$ ------- (see Table 5 for the definitions of t, s, and e. There are 12 possible combinations for four data sets.)
$= \sum e/\sum t$ (because $t = s + e$)
$= 34/151$
$= 0.225$

We thank J. Leebens-Mack, L. Brouillet, J. Averett, and J. P. Huelsenbeck for their helpful comments on various versions of this paper. We also thank L. H. Rieseberg and K. J. Sytsma for their critical reviews of the manuscript. This work was supported by a NSF grant DEB-9200707 to RKJ.

References

Appels, R., Dvorak, K., 1982: Relative rates of divergence of species and gene sequences within the rDNA region of species in the *Triticeae*: implications for the maintenance of homogeneity of a repeated gene family. − Theor. Appl. Genet. **63**: 361–365.

Baldwin, B. G., 1992: Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the *Compositae*. − Mol. Phyl. Evol. **1**: 3–16.

– 1993: Molecular phylogenetics of *Calycadenia* (*Compositae*) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evolution re-examined. – Amer. J. Bot. **80**: 222–238.

BARKER, R. F., HARBERD, N. P., JARVIS, M. G., FLAVELL, R. B., 1988: Structure and evolution of the intergenic region in a ribosomal DNA repeat unit of wheat. – J. Mol. Biol. **201**: 1–17.

BECHMANN, B., LÜKE, W., HUNSMANN, G., 1990: Improvement of PCR amplified DNA sequencing with the aid of detergents. – Nucl. Acids Res. **18**: 1309.

BLACKMORE, S., 1981: Palynology and intergeneric relationships in subtribe *Hyoseridinae* (*Compositae: Lactuceae*). – Bot. J. Linn. Soc. **82**: 1–13.

BREMER, K., 1990: Combinable component consensus. – Cladistics **6**: 369–372.

BULL, J. J., HUELSENBECK, J. P., CUNNINGHAM, C. W., SWOFFORD, D. L., WADDELL, P. J., 1993: Partitioning and combining data in phylogenetic analysis. – Syst. Biol. **42**: 384–397.

CHAMBERS, C. L., 1965: An allopolyploid *Krigia* from the southern Appalachians. – Amer. J. Bot. (Abstr.) **52**: 654.

CHATTERTON, N. J., HSIAO, C., ASAY, K. H., JENSEN, K. B., WANG, R. R.-C., 1992a: Nucleotide sequence of the internal transcribed spacer region of rDNA in the primitive oat species, *Avena longiglumis* DURIEU (*Gramineae*). – Pl. Mol. Biol. **20**: 163–164.

– – – – – 1992b: Nucleotide sequence of the internal transcribed spacer region of rDNA in barley, *Hordium vulgare* L. (*Gramineae*). – Pl. Mol. Biol. **20**: 165–166.

CHINNAPPA, C. C., 1981: Cytological studies in *Krigia* (*Asteraceae*). – Canad. J. Genet. Cytol. **23**: 671–678.

DONOGHUE, M. J., SANDERSON, M. J., 1992: The suitability of molecular and morphological evidence in reconstructing plant phylogeny. – In SOLTIS, P. S., SOLTIS, D. E., DOYLE, J. J., (Eds): Molecular systematics of plants, pp. 340–368. – New York: Chapman and Hall.

DORADO, O., RIESEBERG, L. H., ARIAS, D. M., 1992: Chloroplast introgression in southern California sunflowers. – Evolution **46**: 566–572.

DOVER, G. A., 1982: Molecular drive: a cohesive mode of species evolution. – Nature **299**: 159–165.

D'OVIDIO, R., 1992: Nucleotide sequence of a 5.8 S rDNA gene and of the internal transcribed spacers from *Populus deltoides*. – Pl. Mol. Biol. **19**: 1069–1072.

DOYLE, J. J., 1992: Gene trees and species trees: molecular systematics as one character taxonomy. – Syst. Bot. **17**: 144–163.

– DOYLE, J. A., 1987: A rapid DNA isolation procedure for small quantities of fresh leaf tissue. – Phytochem. Bull. **19**: 11–15.

FELSENSTEIN, J., 1985: Confidence limits on phylogenies: an approach using the bootstrap. – Evolution **39**: 783–791.

– 1991: PHYLIP (Phylogeny inference package), version 3.4. Computer program distributed by the author. – St Louis: The University of Washington.

FEUER, S., TOMB, A. S., 1977: Pollen morphology and detailed structure of family *Compositae*, tribe *Cichorieae* II. subtribe *Microseridinae*. – Amer. J. Bot. **64**: 230–245.

FIELD, G. F., OLSEN, G. L., LANE, D. J., GIOVANNONI, S. J., GHISELIN, M. T., RAFF, E. C., PACE, N. R., RAFF, R. A., 1988: Molecular phylogeny of the animal kingdom. – Science **239**: 748–753.

GONZALEZ, I. L., SYLVESTER, J. E., SMITH, T. F., STAMBOLIAN, D., SCHMICKEL, R. D., 1990: Ribosomal RNA gene sequences and hominoid phylogeny. – Mol. Biol. Evol. **7**: 203–219.

GRAY, A., 1886: *Apogon* and *Krigia*. – In: Synoptical flora of North America, pp. 410–412. – New York: Ivison, Baikeman, Taylor.

GYLLENSTEN, U. B., ERLICH, H. A., 1988: Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. – Proc. Natl. Acad. Sci. USA **85**: 7652–7656.

HAMBY, R. K., ZIMMER, E. A., 1992: Ribosomal RNA as a phylogenetic tool in plant systematics. – In SOLTIS, P. S., SOLTIS, D. E., DOYLE, J. J., (Eds): Molecular systematics of plants, pp. 50–91. – New York: Chapman and Hall.

HARBORNE, J. B., 1977: Variations in pigment patterns in *Pyrrhopappus* and related taxa of the *Cichorieae*. – Phytochemistry **16**: 927–928.

HARRISON, R. G., 1991: Molecular changes at speciation. – Annu. Rev. Ecol. Syst. **22**: 281–308.

HENDY, M. D., PENNY, D., 1982: Branch and bound algorithms to determine minimal evolutionary trees. – Math. Biosci. **59**: 227–290.

HILLIS, D. M., 1987: Molecular versus morphological approaches to systematics. – Annu. Rev. Ecol. Syst. **18**: 23–42.

– 1991: Discriminating between phylogenetic signal and random noise in DNA sequences. – In MIYAMOTO, M. M., CRACRAFT, J., (Eds): Phylogenetic analysis of DNA sequences, pp. 278–294. – New York: Oxford University Press.

– DIXON, M. T., 1989: Vertebrate phylogeny: evidence from 28 S ribosomal DNA sequences. – In FERNHOLM, B., BREMER, K., JORNVALL, H., (Eds): The hierarchy of life, pp. 355–367. Nobel Symposium **70**. – Amsterdam: Elsevier.

– MORITZ, C., PORTER, C. A., BAKER, R. J., 1991: Evidence for biased gene conversion in concerted evolution of ribosomal DNA. – Science **251**: 308–310.

JANSEN, R. K., WALLACE, R. S., KIM, K.-J., CHAMBERS, K. L., 1991: Systematic implications of chloroplast DNA variation in the subtribe *Microseridinae* (*Asteraceae: Lactuceae*). – Amer. J. Bot. **78**: 1015–1027.

JEFFREY, C., 1966: Notes on *Compositae:* I. The *Cichorieae* in east tropical Africa. – Kew Bull. **18**: 427–482.

JORGENSEN, R. A., CLUSTER, P. D., 1988: Modes and tempos in the evolution of nuclear ribosomal DNA: new characters for evolutionary studies and new markers for genetic and population studies. – Ann. Missouri Bot. Gard. **75**: 1238–1247.

JUKES, T. H., CANTOR, C. R., 1969: Evolution of protein molecules. – In MUNRO, H. N., (Ed.): Mammalian protein metabolism, pp. 21–123. – New York: Academic Press.

KAVANAGH, T. A., TIMMIS, J. N., 1988: Structure of melon rDNA and nucleotide sequence of the 17–25 S spacer region. – Theor. Appl. Genet. **76**: 673–680.

KIM, K.-J., MABRY, T. J., 1991: Phylogenetic and evolutionary implications of nuclear ribosomal DNA variation in dwarf dandelions (*Krigia, Lactuceae, Asteraceae*). – Pl. Syst. Evol. **177**: 53–69.

– TURNER, B. L., 1992: Systematic overview of *Krigia* (*Asteraceae: Lactuceae*). – Brittonia **44**: 173–198.

– – JANSEN, R. K., 1992 a: Phylogenetic and evolutionary implications of interspecific chloroplast DNA variation in *Krigia* (*Lactuceae-Asteraceae*). – Syst. Bot. **17**: 449–469.

– JANSEN, R. K., TURNER, B. L., 1992 b: Evolutionary implications of intraspecific chloroplast DNA variation in dwarf dandelions (*Krigia-Asteraceae*). – Amer. J. Bot. **79**: 708–715.

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.

– 1981: Estimation of evolutionary distances between homologous nucleotide sequences. – Proc. Natl. Acad. Sci. USA **78**: 454–458.

KISS, T., MIHÁLY, K., STEFFEN, A., SOLYMOSY, F., 1988: Nucleotide sequence of the 17 S–25 S spacer region from tomato rDNA. – Nucl. Acids Res. **16**: 7179.

KLUGE, A. G., 1989: A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (*Boidae, Serpentes*). – Syst. Zool. **38**: 7–25.

LAKE, A. J., 1985: Evolving ribosome structure: domains in archaebacteria, eubacteria, eocytes, and eukaryotes. – Annu. Rev. Biochem. **54**: 507–530.

LEE, S. B., TAYLOR, J. W., 1992: Phylogeny of five fungus-like protoctistan *Phytophthora* species, inferred from the internal transcribed spacers of ribosomal DNA. – Mol. Biol. Evol. **9**: 636–653.

LEVINSON, G., GUTMAN, G. A., 1987: Slipped-strand mispairing: a major mechanism for DNA sequence evolution. – Mol. Biol. Evol. **4**: 203–221.

MCDADE, L. A., 1992: Hybrids and phylogenetic systematics II. The impact of hybrid on cladistic analysis. – Evolution **46**: 1329–1346.

MICKEVICH, M. F., FARRIS, J. S., 1981: The implications of congruence in *Menidia*. – Syst. Zool. **30**: 351–370.

MIYAMOTO, M. M., 1985: Consensus cladograms and general classifications. – Cladistics **1**: 186–189.

NORTHINGTON, D. K., 1974: Systematic studies of the genus *Pyrrhopappus* (*Compositae, Cichorieae*). – Spec. Publ. Mus. Texas Tech. Univ. **6**: 1–38.

O'DONNELL, K., 1992: Ribosomal DNA internal transcribed spacers are highly divergent in the phytopathogenic ascomycete *Fusarium sambucinum* (*Gibberella pulicaris*). – Curr. Genet. **22**: 213–220.

OLSEN, G. J., 1987: Earliest phylogenetic branchings: comparing rRNA-based evolutionary trees inferred with various techniques. – Cold Spring Harbor Symp. Quant. Biol. **52**: 825–837.

RATHGEBER, J., CAPESIUS, I., 1989: Nucleotide sequence of the 18 S–25 S spacer region from mustard DNA. – Nucl. Acids Res. **17**: 7522.

RIESEBERG, L. H., SOLTIS, D. E., PALMER, J. D., 1988: A molecular re-examination of introgression between *Helianthus annuus* and *H. bolanderi* (*Compositae*). – Evolution **42**: 227–238.

ROGERS, S. O., BENDICH, A. J., 1987: Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. – Pl. Mol. Biol. **9**: 509–520.

SAGHAI-MAROOF, M. A., SOLIMAN, K. M., JORGENSEN, R. A., ALLARD, R. W., 1984: Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. – Proc. Natl. Acad. Sci. USA **81**: 8014–8018.

SANDERSON, M. J., DONOGHUE, M. J., 1989: Patterns of variation in levels of homoplasy. – Evolution **43**: 1781–1795.

SAITOU, N., NEI, M., 1987: The neighbor-joining method: a new method for reconstructing phylogenetic trees. – Mol. Biol. Evol. **4**: 406–425.

SCHIEBEL, K., HEMLEBEN, V., 1989: Nucleotide sequence of the 18 S–25 S spacer region from rDNA of mung bean. – Nucl. Acids Res. **17**: 2852.

SHAFFER, H. B., CLARK, J. M., KRAUS, F., 1991: When molecules and morphology clash: a phylogenetic analysis of North American ambystomatid salamanders (*Caudata: Ambystomatidae*). – Syst. Zool. **40**: 284–303.

SHINNERS, L. H., 1947: Revision of the genus *Krigia* SCHREBER. – Wrightii **1**: 187–206.

SMITH, G. P., 1976: Evolution of repeated DNA sequences by unequal crossover. – Science **191**: 528–535.

STANDLEY, P. C., 1911: A revision of the cichoriaceous genera *Krigia, Cynthia*, and *Cymbia*. – Contr. U. S. Natl. Herb. **13**: 351–357.

STEBBINS, G. L., 1953: A new classification of the tribe *Cichorieae*, family *Compositae*. – Madroño **12**: 33–64.

– JENKINS, J. A., WALTER, M. S., 1953: Chromosomes and phylogeny in the *Compositae*, tribe *Cichorieae*. – Univ. Calif. Publ. Bot. **26**: 401–429.

SWOFFORD, D. L., 1991 a: When are phylogeny estimates from molecular and morphological data incongruent. – In MIYAMOTO, M. M., CRACRAFT, J., (Eds): Phylogenetic analysis of DNA sequences, pp. 295–333. – New York: Oxford University Press.
– 1991 b: PAUP: phylogenetic analysis using parsimony, version 3.0 r. Computer program distributed by the author. – Urbana-Champain: Illinois Natural History Survey.
SYTSMA, K. J., SCHAAL, B. A., 1990: Ribosomal DNA variation within and among individuals of *Lisianthus* (*Gentianaceae*) populations. – Pl. Syst. Evol. **170**: 97–106.
TAKAIWA, F., OONO, K., SUGIURA, M., 1985: Nucleotide sequence of the 17 S–25 S spacer region from mice rDNA. – Pl. Mol. Biol. **4**: 355–364.
TURNER, B. L., KIM, K.-J., 1990: Overview of the genus *Pyrrhopappus* (*Asteraceae: Lactuceae*) with emphasis upon chloroplast DNA restriction site data. – Amer. J. Bot. **77**: 845–850.
UEKI, M., UCHIZAWA, E., YAKURA, K., 1992: The nucleotide sequence of the rRNA intergenic spacer region in a wild species of the genus *Vicia, V. angustifolia.* – Pl. Mol. Biol. **18**: 175–178.
VENKATESWARLU, K., NAZAR, R., 1991: A conserved core structure in the 18 S–25 S rRNA intergenic region from tobacco, *Nicotiana rustica.* – Pl. Mol. Biol. **17**: 189–194.
WANG, B., FANG, Q., WILLIAMS, W. V., WEINER, D. B., 1992: Double-stranded DNA sequencing by linear amplification with *Taq* DNA polymerase. – Biotechniques **13**: 527–530.
WHITE, T. J., BRUNS, T., LEE, S., TAYLOR, J., 1990: Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. – In INNIS, M. A., GELFAND, D. H., SNINSKY, J. J., WHITE, T. J., (Eds): PCR protocols: a guide to methods and applications, pp. 315–322. – San Diego: Academic Press.
WINSHIP, P. R., 1989: An improved method for direct sequencing PCR amplified material using dimethyl sulphoxide. – Nucl. Acids Res. **17**: 1266.
YAKURA, K., KATO, A., TANIFUJI, S., 1984: Length heterogeneity in the large spacer of *Vicia fava* rDNA is due to differing number of 325 bp repetitive elements. – Mol. Gen. Genet. **139**: 400–405.
– NISHIKAWA, K., 1992: The nucleotide sequences of the rDNA spacer region between the 25 S and 18 S rRNA genes in a species of the genus *Vicia, V. hirsuta.* – Pl. Mol. Biol. **19**: 537–539.
YOKOTA, Y., KAWATA, T., IIDA, Y., KATA, A., TANIFUJI, S., 1989: Nucleotide sequences of the 5.8 S rRNA gene and internal transcribed spacer regions in carrot and broad bean ribosomal DNA. – J. Mol. Evol. **29**: 294–301.
ZHANG, W., READING, C., DEISSEROTH, A. B., 1992: Improved PCR sequencing with formamide. – Trends Genet. **8**: 332.

Address of the authors: KI-JOONG KIM and ROBERT K. JANSEN, Department of Botany, University of Texas, Austin, TX 78713-7640, USA.

Accepted September 29, 1993 by B. L. TURNER