Primers for PCR Amplification of Asteraceae Chloroplast DNA

Authors: Panero, José L., and Crozier, Bonnie S.

PRIMERS FOR PCR AMPLIFICATION OF ASTERACEAE CHLOROPLAST DNA

José L. Panero and Bonnie S. Crozier

Section of Integrative Biology, School of Biological Sciences, The University of Texas, Austin, Texas 78712

Abstract: Primers for the amplification and sequencing of DNA fragments from chloroplast genes and non-coding regions are provided to facilitate molecular phylogenetic studies aimed at building a tree-of-life for the Asteraceae. The primers reported here have been extensively tested and some empirical guidelines are included to facilitate their use.

Keywords: chloroplast DNA primers, *acc*K, *mat*K, *ndh*C, *ndh*D, *ndh*F, *ndh*G, *ndh*I, *ndh*J, *ndh*K, *pet*B, *pet*D, *rbc*L, *rpo*B, *rpo*C, *trn*T, *trn*L, 23S, genetics/plant, biochemical methods, nucleic acids.

This paper provides sequencing primer information useful for molecular phylogenetic studies of the family Asteraceae. The initial impetus to develop primers for new markers of the chloroplast was to elucidate phylogenetic relationships of the Heliantheae sensu lato (Panero et al., in prep.). These primers are the products of extensive testing and were used to produce DNA sequence data by the first author in recent comprehensive phylogenetic studies of the family (Panero & Funk, 2002; Panero & Funk, in prep.). We provide these primers to expand the suite of chloroplast markers now in use for phylogenetic studies, and expect that the Asteraceae research community will use these markers and expedite, through collaboration, the creation of a tree-of-life of sunflowers. Below we provide some empirical suggestions that can facilitate the use of these markers for phylogenetic purposes.

Molecular markers were chosen from several areas of the chloroplast genome (Fig. 1) and for the most part amplify DNA of all lineages of the family and some related families in the Asterales (sensu the Angiosperm Phylogeny Group, see http://www.mobot.org/MOBOT/Research/APweb/welcome.html; Stevens, 2003). Exceptions exist however, and are noted below. Markers were originally chosen to elucidate phylogenetic relationships in the Heliantheae.

Subsequently, the study was expanded to include representatives of most lineages of the Asteraceae including many members of subfamilies Mutisioideae, Gochnatioideae, Hecastocleioideae, Carduoideae, and Pertyoideae (see Panero & Funk, 2002 for latest tribal classification). For the most part the Heliantheae primers worked efficiently for most lineages of the Asteraceae but in several cases additional primers were developed to obtain adequate amplifications useful in sequencing reactions.

We developed primers for markers not previously used in phylogenetic studies such as *acc*D, *ndh*C, *ndh*D, *ndh*I, *ndh*J, *ndh*K, *rpo*B, the *ndh*I-*ndh*G and 23S-*trn*A intergenic spacers, and the *trn*A, *pet*B and *pet*D introns. We provide new primers for *mat*K, *rbc*L, and *ndh*F to accomplish the longer runs (600 bp on average) possible using current sequencing technologies. All the primers that we report here were developed by using the *Nicotiana tabacum* chloroplast map of Shinozaki et al. (1986) as an initial source of sequence information and complemented with the *Nicotiana* genome sequence available in Genbank (http://www.ncbi.nlm.nih.gov:80/entrez/viewer.fcgi?val=NC_001879) to provide exact coordinates for most primers (indicated at the end of the primer information in Figs. 2–5). Several primers are the result of extensive cloning of Asteraceae products from
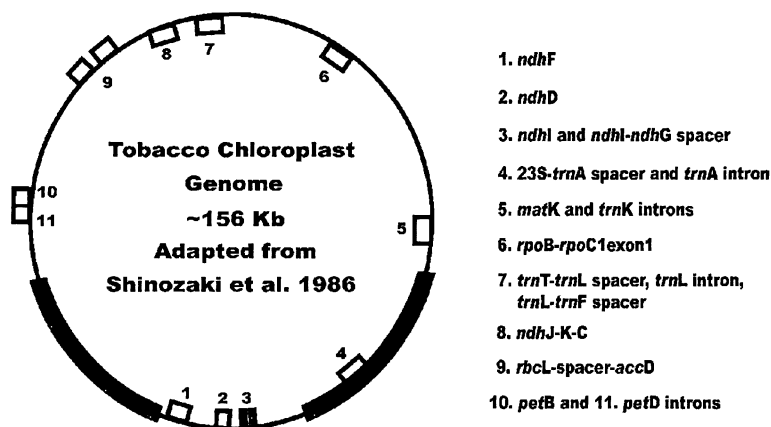
FIG. 1. Locations of markers relative to the *Nicotiana tabacum* chloroplast genome map. Numbers indicate their order of discussion in the text. Dark bands represent the inverted repeat regions of the chloroplast.

poor amplifications based on *Nicotiana* primer sequences. This is especially true for the *pet*B and *pet*D introns, and the *rpo*B and *acc*D genes.

In general we found that data from coding regions is easier to align than noncoding regions and therefore have lower levels of homoplasy caused by ambiguous alignment of the data. However, we have refrained, for the most part, from ranking specific regions of the chloroplast on their perceived usefulness in uncovering phylogenetic signal or ease of alignment or a combination of both. We do not endorse the use of any single marker to obtain quick, credible phylogenetic hypotheses but rather we advocate the use of multiple markers to produce well-supported tree topologies to address outstanding questions in Asteraceae systematics. Because they are physically linked and recombination is minimal, chloroplast genes likely share the same branching histories, so multiple gene studies can result in more robust estimates of phylogeny when several markers are combined (Weins, 1998; de Queiroz, 2002). Indeed, we found that data partitions for large numbers of taxa sampled across the Asteraceae and across the tribe Heliantheae differed only in minor parts of their histories (Panero and Funk, in prep.; Panero et al.,

in prep). However, sparse taxon sampling may increase phylogenetic estimation error (see Pollack et al., 2002; Hillis et al., 2003; Graybeal, 1998). A large number of sequences for representatives of many lineages of the Asteraceae using the primers reported herein are already at the research community's disposal (i.e., Genbank). By publishing these primers we hope to accelerate the production of a robust molecular phylogeny for the entire Asteraceae through the collaborative efforts of sunflower researchers across the world.

AMPLIFICATION AND PRIMER CONSTRUCTION

DNA for all markers was amplified using the following PCR protocol: 5 min denaturation at 95°C, 45 sec annealing at 48°C, and 1 min primer extension at 72°C, followed by 35 cycles with 1 min denaturation at 95°C, 45 sec (2·sec were added in each subsequent cycle) annealing at 48°C, and 1 min primer extension at 72°C. The reaction was terminated with a 10 min final extension at 72°C and then held at 10°C. In some cases, non-specific amplification (multiple band pattern) was eliminated by increasing the annealing temperature 2–5°C. PCR products were purified using filter

centrifugation (Ultrafree®-MC centrifugal filter units, 30 000 nominal molecular weight limit—NMWL, Millipore Corporation, Bedford, Massachusetts, USA, or PCR cleaning columns from Qiagen). Cloning of PCR products was performed using the TOPO TA Cloning® kit from Invitrogen. Ten transformed colonies were used as templates in PCR reactions using the same specifications as outlined above. Colonies having the insert of approximately the same size as the original cloning template were identified by gel electrophoresis of PCR products. One of the colonies containing the right insert was picked at random for plasmid harvest. A sample of the chosen colony was used to inoculate 5 ml of LB growing medium. The liquid medium was incubated at 37°C and shaken at 250 rpm. After 6 hours the cells were lysed and plasmids removed using a Promega Wizard® Plus Miniprep kit. The purified plasmids were used as template in a new PCR reaction whose product was prepared for sequencing.

The program MacVector 5.0 (originally from the Eastman Kodak Company, newer versions sold by Accelrys, Inc.) was used to construct primers with melting temperatures ranging between 50–60°C and G-C content above 45%. When it was not possible to meet these requirements for a specific DNA region, primers with lower G-C contents and melting temperatures were accepted. The program excludes primer designs with hairpins and self-duplexes. Primers were obtained from Integrated DNA Technologies of Coralville, IA 52241, USA (http://www.idtdna.com).

DISCUSSION

As a general rule, we have designed a primer overlap of approximately 30–100 bp for the longer regions that required amplification and sequencing in multiple segments (see maps in Figs. 2–5 for *ndh*F, *ndh*C, *mat*K, *rpo*B, *rpo*C, *ndh*K and *rbc*L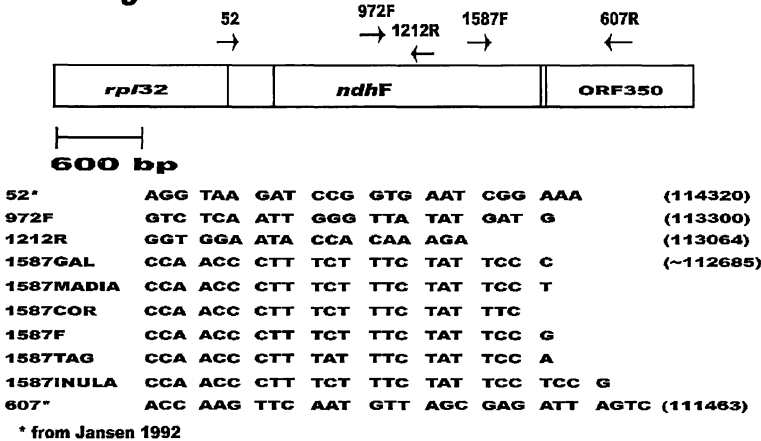). This also provides some sequence editing flexibility when determining where to trim overlapping contig sequence files. We never amplify an entire gene of more than 2000 bp in length using outside primers, instead we develop internal primers to divide the sequencing products into smaller regions. This strategy provides cleaner sequencing results. Below we provide a discussion of our amplification protocols and guidelines for each marker.

*ndh*F. We amplified the *ndh*F gene in two segments using 4 primers, primer pairs 52-1212R and 972F-607 (primers 52 and 607 reported in Jansen, 1992). The internal primer 1587F and derivatives are used to read the final 3′ end of the gene. The 1587F region of *ndh*F is very variable and placing primers that would produce high quality sequences is challenging. The 1587 family of primers was initially developed for the Heliantheae sensu lato and allies and works well for any other sunflower except *Inula* that has an insertion of 3 bp at the end of the priming region. Therefore, it is advisable to sequence the 972F-607 amplification product with primer 972F first, read the 1587F primer region, and then ascertain which primer to use. A map is presented in Fig. 2.

*ndh*D. This gene is difficult to amplify especially when DNA templates are derived from herbarium specimens. We initially amplified the gene using primers derived from the *Nicotiana* sequence and then used the cloned product to design primer pairs ndhDF-672R and 732F-ndhDR (Fig. 2). As we expanded the survey of sunflowers to cover most lineages of the family, we realized that amplification was no longer efficient at the basal lineages of the family. We then created two outside primers ycs5F and psaCR to replace ndhDF and ndhDR respectively. These have the added advantage of producing an amplification product containing the entire gene.

*ndh*I. This is one of the most reliable markers to use for its ease of amplification. When amplifying DNA of genera from different lineages, PCR products using primers

1. *ndh*F gene

52*	AGG TAA GAT CCG GTG AAT CGG AAA	(114320)
972F	GTC TCA ATT GGG TTA TAT GAT G	(113300)
1212R	GGT GGA ATA CCA CAA AGA	(113064)
1587GAL	CCA ACC CTT TCT TTC TAT TCC C	(~112685)
1587MADIA	CCA ACC CTT TCT TTC TAT TCC T	
1587COR	CCA ACC CTT TCT TTC TAT TTC	
1587F	CCA ACC CTT TCT TTC TAT TCC G	
1587TAG	CCA ACC CTT TAT TTC TAT TCC A	
1587INULA	CCA ACC CTT TCT TTC TAT TCC TCC G	
607*	ACC AAG TTC AAT GTT AGC GAG ATT AGTC (111463)	

* from Jansen 1992

2. *ndh*D gene

NDHDF	TTC GAC CTT GTC AAC TGC	(117523)
NDHDR	GAA CTC CTT CTA ACG ACT TAT GC	(118877)
672F	CGG CTA GAA GCA TAC AAG	(118275)
732R	TTG CCG ATT CTA CCC CTA C	(118355)
psaCR	TGG CAT GAA ACA ACT CGC	(119165)
ycs5F	GGG GTC AAT CTT TTA GGA ATA GGG	(117223)

3. *ndh*I and *ndh*I-*ndh*G spacer

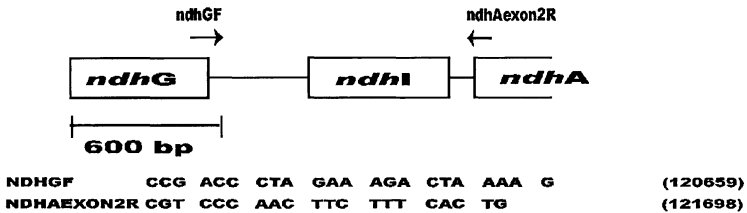| NDHGF	CCG ACC CTA GAA AGA CTA AAA G	(120659)
NDHAEXON2R	CGT CCC AAC TTC TTT CAC TG	(121698)

FIG. 2. Primers developed for the small single copy region. Relative positions of the *ndh*F gene, *ndh*D gene and the *ndh*I and *ndh*I–*ndh*G intergenic spacer regions are shown and the direction of primers indicated by arrows. Numbers in parentheses represent primer starting position based on the *Nicotiana tabacum* chloroplast genome sequence.

ndhGF and ndhAexon2R can be variable in size due to variation in the *ndh*G-*ndh*I spacer. This region has a higher number of base pair changes and can be useful in studies aimed at elucidating phylogenetic relationships among closely related species (Fig. 2). The DNA of some taxa, (e.g., *Doronicum* and *Dimorphotheca*) did not produce amplification product for this marker using the primer pair reported above. We solved this primer-template mismatch by replacing the *ndh*GF primer with *ndh*D's 732R and sequencing with primer ndhAexon2R. This region may also be useful at the order and

4. 23S-*trn*A spacer and *trn*A intron

| 23SF | ATC CAC CGT AAG CCT TTC | (106339) |
| TRNIR | ATT GGT TGG ACC GTA GGT GC | (105260) |

5. *mat*K gene and *trn*K introns

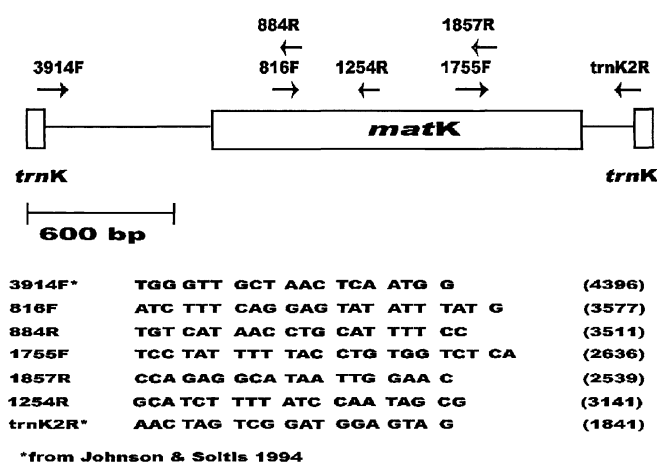| 3914F* | TGG GTT GCT AAC TCA ATG G | (4396) |
816F	ATC TTT CAG GAG TAT ATT TAT G	(3577)
884R	TGT CAT AAC CTG CAT TTT CC	(3511)
1755F	TCC TAT TTT TAC CTG TGG TCT CA	(2636)
1857R	CCA GAG GCA TAA TTG GAA C	(2539)
1254R	GCA TCT TTT ATC CAA TAG CG	(3141)
trnK2R*	AAC TAG TCG GAT GGA GTA G	(1841)

*from Johnson & Soltis 1994
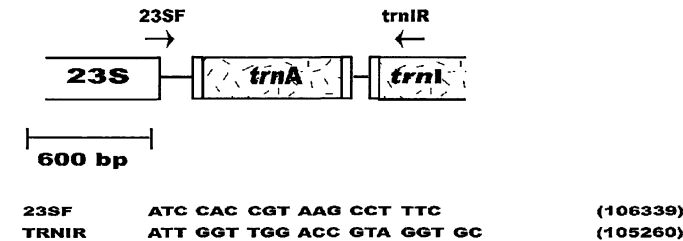
FIG. 3. Primers and their relative positions for the 23S-*trn*A intergenic spacer and *trn*A intron gene, and *mat*K and the *trn*K introns. Numbers in parentheses represent primer starting position based on the *Nicotiana tabacum* chloroplast genome sequence.

subclass levels as we recently amplified DNA of the monocot genus *Nolina* using the primers reported here.

The 23S-trnA spacer, trnA intron, and trnA-trnI spacer. This marker is the only one we developed from the inverted repeat region (Figs. 1 and 3). The primer pair 23SF-trnIR produces a strong amplification product of equivalent size for most genera of the Asteraceae we sampled. The region is not as generally useful for phylogenetic purposes because mutations are mostly autopomorphic and synapomorphies are minimal. However, this region contains a synapomorphy for the Senecioneae and an important synapomorphy shared by the Polymnieae and lineages above it (Panero et al., in prep).

The matK gene and trnK introns. This region is easy to amplify and provides many mutations useful in phylogenetic studies of Asteraceae (Bayer et al., 2002). In some species of tribe Heliantheae and allies, the 3' end of the *trn*K split intron contains a polyT region followed by a polyA region that makes its sequencing challenging, so this region was omitted from our phylogenetic studies. We amplify the *mat*K region using the primer pairs 3914F-884R, 816F-1857R, and 1755F-trnK2R (Fig. 3). For certain members of the basal lineages of the Asteraceae, we developed primer 1254R because 884R was providing bad sequences, good amplification product notwithstanding.

rpoB-rpoC1 genes. These genes provide

a wealth of base substitution mutations that can be useful in phylogenetic studies. The primers reported here have been used to amplify all lineages of the Asteraceae as well as representatives of the Goodeniaceae and Calyceraceae. This region is not easy to amplify in the Asteraceae and it is even more difficult when the DNA is obtained from herbarium specimens. Furthermore, many of the sequences have ambiguous outputs for the first 30–50 bases. The primers reported here are products of extensive cloning of the PCR products for several sunflower genera, and still they are not the best or final primers for this region. The 3′ end of *rpo*B is difficult to amplify for most members of Asteroideae probably because the gene is longer than in members of basal lineages. We amplify the *rpo*B gene in three segments using the following primer pairs: rpoCF-1394R, 1270F-2503R, and rpoB1R or rpoB2R-2426F. To amplify the 3′ end of rpoC1exon2 through rpoC1exon1, we use rpoC904F, 917F, or 952F with rpoBS1R or rpoBS2R. If the aim is to sequence only exon1 of rpoC1 the primers rpoC1462CF or 1485CF with rpoBS1R or rpoBS2R can be used (Fig. 4). For those interested in the *rpo*C1 intron region, our preliminary data (i.e., PCR band size) show that the intron is fairly similar in size across the family.

*trn*L *intron.* The *trn*L intron and the *trn*L-*trn*F intergenic spacer were amplified using the primers C and F of Taberlet et al. (1991). We cloned the PCR product of primers C and F and found the primer C sequence to have changed by one base in the Asteraceae (primer C-ASTER in Fig. 4). This region is quite popular judged by the number of molecular phylogenetic studies based on it; it amplifies readily and sequences are long and of good quality. As mentioned by Taberlet et al. (1991), primer F is not a sequencing primer. Contrary to his advice, we have sequenced PCR product using the F primer and have obtained mixed results. The best sequencing reactions were obtained using primer C-ASTER.

*ndh*J-*ndh*K-*ndh*C *region.* The *ndh*J-K-C
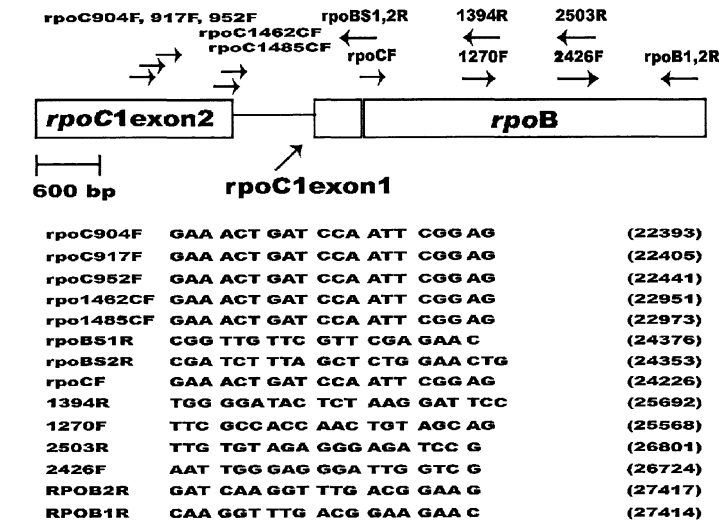
region amplifies reliably but we have not yet used these primers to amplify DNA of members of the basal lineages of the family. Primer pairs ndhJF-ndhK2R and ndhK1F-ndhCR amplify in two segments the entire *ndh*K gene and partial amplifications for genes *ndh*J and *ndh*C (Fig. 4). Primers for this marker need further refinement to obtain amplification products that contain the three genes in their entirety.

*rbc*L *gene.* We amplified the *rbc*L gene using primers widely used in previous studies including studies of members of the Asteraceae here labeled rbcL1 and rbcL2 (Hillis et al., 1996, p. 240; Olmstead, 1992; see Fig. 5). However, sequencing of amplification products using these primers produces mediocre results in the Asteraceae, especially for the initial 30–40 bp of each sequence. Reverse primer rbcL2 is located 103 downstream of the coding region in tobacco but contains a stem-loop structure that may interfere with good sequencing of all taxa. Forward primer rbcL1 is located exactly at the first 26 bases at the start of the gene so that part of the 5′ end of the gene cannot be sequenced using that primer (Fig. 5). We developed internal primers 876F and 911R to be used with rbcL2 and rbcL1, respectively. Amplification of the gene in two segments using these internal primers yields products of high quality for sequencing and the ease in obtaining PCR product is second only to *ndh*I. Our sequencing primers are primers 876F and 911R and read through the priming sites of rbcL1 and 2 when sequencing runs of 750 or more can be obtained. In the future we will develop primers in *atp*B and *acc*D to amplify the complete *rbc*L gene.

*rbc*L-*acc*D *spacer.* The intergenic spacer between *rbc*L and *acc*D is highly variable in size and may be of use to those researchers interested in DNA variation among closely related taxa. We amplified this region along with the 5′ end of the *acc*D gene and sequence the spacer portion using our primer rbcL 1581F anchored in the *rbc*L gene.

*acc*D. The *acc*D gene was sequenced for

6. *rpoC1-rpoB genes*

rpoC904F	GAA ACT GAT CCA ATT CGG AG	(22393)
rpoC917F	GAA ACT GAT CCA ATT CGG AG	(22405)
rpoC952F	GAA ACT GAT CCA ATT CGG AG	(22441)
rpo1462CF	GAA ACT GAT CCA ATT CGG AG	(22951)
rpo1485CF	GAA ACT GAT CCA ATT CGG AG	(22973)
rpoBS1R	CGG TTG TTC GTT CGA GAA C	(24376)
rpoBS2R	CGA TCT TTA GCT CTG GAA CTG	(24353)
rpoCF	GAA ACT GAT CCA ATT CGG AG	(24226)
1394R	TGG GGA TAC TCT AAG GAT CC	(25692)
1270F	TTC GCC ACC AAC TGT AGC AG	(25568)
2503R	TTG TGT AGA GGG AGA TCC G	(26801)
2426F	AAT TGG GAG GGA TTG GTC G	(26724)
RPOB2R	GAT CAA GGT TTG ACG GAA G	(27417)
RPOB1R	CAA GGT TTG ACG GAA GAA C	(27414)

7. *trn*T-*trn*L spacer-*trn*L intron, *trn*L-*trn*F spacer

C-ASTER*	CGA AAT TGG TAG ACG CTA CG

*modified from Taberlet 1991

8. *ndh*J-C-K genes

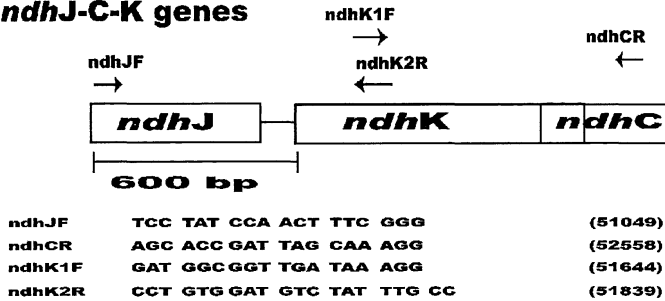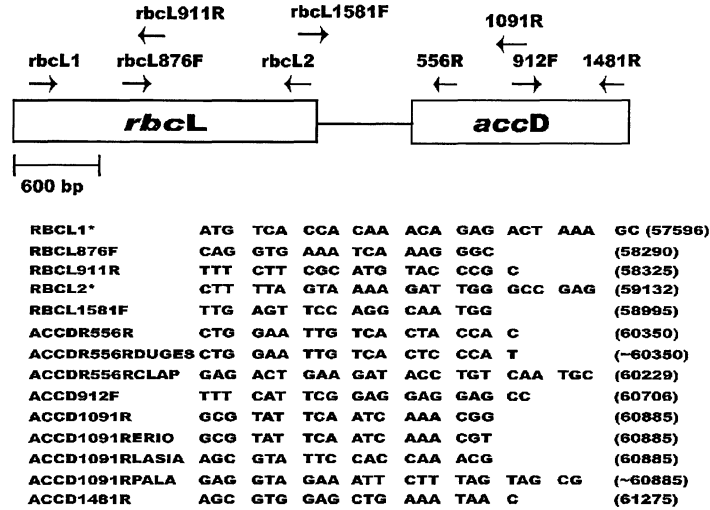| ndhJF	TCC TAT CCA ACT TTC GGG	(51049)
ndhCR	AGC ACC GAT TAG CAA AGG	(52558)
ndhK1F	GAT GGC GGT TGA TAA AGG	(51644)
ndhK2R	CCT GTG GAT GTC TAT TTG CC	(51839)

FIG. 4. Primers and their relative positions for the *rpo*C1-*rpo*B genes, the *trn*L intron-*trn*L-*trn*F intergenic spacer region, and the *ndh*J-C-K genes region. Numbers in parentheses represent primer starting position based on the *Nicotiana tabacum* chloroplast genome sequence.

the Heliantheae and related tribes only. Amplification of basal lineages of the Asteraceae has provided mixed success to date. We amplified this gene in two segments using primer pairs rbcL1581F-accD1091R and rbcL912F-accD1481R. Internal primer accDR556 was used to sequence the 5'end of the gene. However, we normally sequenced some taxa using accD1091R first because the 5′ end of the gene has multiple insertions and deletions that make sequencing this area challenging when they coincide with the 556R primer site. The 3′ end of the gene is amplified using the primer pair accD912F-accD1481R. This area is easy to align because there is little variation in sequences; primer accD912F is used to sequence the 3′ end. Primers for *acc*D and the *rbc*L-*acc*D spacer were obtained after extensive cloning of several taxa of the Heliantheae and thus the primers are efficient at amplifying DNA only of taxa in this and

9. *rbc*L-spacer-*acc*D

RBCL1*	ATG	TCA	CCA	CAA	ACA	GAG	ACT	AAA	GC (57596)
RBCL876F	CAG	GTG	AAA	TCA	AAG	GGC			(58290)
RBCL911R	TTT	CTT	CGC	ATG	TAC	CCG	C		(58325)
RBCL2*	CTT	TTA	GTA	AAA	GAT	TGG	GCC	GAG	(59132)
RBCL1581F	TTG	AGT	TCC	AGG	CAA	TGG			(58995)
ACCDR556R	CTG	GAA	TTG	TCA	CTA	CCA	C		(60350)
ACCDR556RDUGES	CTG	GAA	TTG	TCA	CTC	CCA	T		(~60350)
ACCDR556RCLAP	GAG	ACT	GAA	GAT	ACC	TGT	CAA	TGC	(60229)
ACCD912F	TTT	CAT	TCG	GAG	GAG	GAG	CC		(60706)
ACCD1091R	GCG	TAT	TCA	ATC	AAA	CGG			(60885)
ACCD1091RERIO	GCG	TAT	TCA	ATC	AAA	CGT			(60885)
ACCD1091RLASIA	AGC	GTA	TTC	CAC	CAA	ACG			(60885)
ACCD1091RPALA	GAG	GTA	GAA	ATT	CTT	TAG	TAG	CG	(~60885)
ACCD1481R	AGC	GTG	GAG	CTG	AAA	TAA	C		(61275)

*see Hillis et al., 1996; Olmstead et al., 1992

10. *pet*B and 11. *pet*D introns

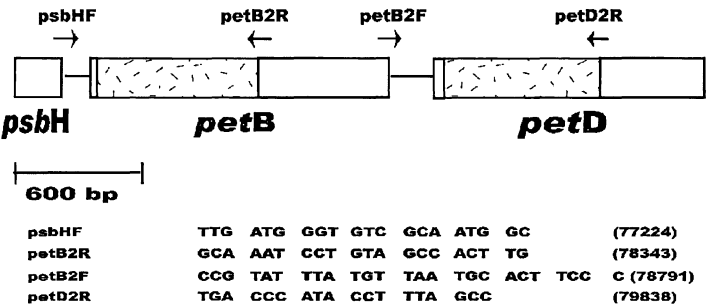| psbHF	TTG	ATG	GGT	GTC	GCA	ATG	GC		(77224)
petB2R	GCA	AAT	CCT	GTA	GCC	ACT	TG		(78343)
petB2F	CCG	TAT	TTA	TGT	TAA	TGC	ACT	TCC	C (78791)
petD2R	TGA	CCC	ATA	CCT	TTA	GCC			(79838)

FIG. 5. Primers and their relative positions for the *rbcL* gene, *rbcL-accD* intergenic spacer, *accD* gene, *petB* and *petD* introns. Numbers in parentheses represent primer starting position based on the *Nicotiana tabacum* chloroplast genome sequence.

related tribes. Further experimentation including cloning may be required to produce primers that can be viewed as truly universal for the family.

petB and petD introns. Primers for the *petB* and *petD* introns amplify very efficiently DNA of members of subfamily Asteroideae (Fig. 5). The primers were developed after extensive cloning of members of tribe Heliantheae. Both introns show little length variation across the taxa sampled and are relatively easy to align when compared to other introns commonly used in phylogenetic studies such as the *rpl*16 intron. We amplified these regions by using the primer pairs psbHF-petB2R and petB2F-petD2R.

CONCLUSIONS

The pace and ease in obtaining molecular data for systematic studies continue to

accelerate and it is plausible that in the near future such studies will rely more heavily on sequence comparisons of whole chloroplast genomes, or of a specific nuclear chromosome rather than on studies limited to one or a few genes. At present however, the cost of doing partial or complete chloroplast genome studies for most genera of Asteraceae or flowering plants is still not economically feasible and most systematists continue to sequence region by region. As mentioned above, sequences using these primers have already been performed for many genera of the family and are available through Genbank. By publishing a detailed description of the primers and protocols used to generate those sequences, we hope to facilitate the Asteraceae research community's efforts to produce a tree-of-life of sunflowers, a comprehensive phylogenetic hypothesis that will stimulate posing more interesting questions about the evolution and radiation of sunflowers across most of our planet.

ACKNOWLEDGEMENTS

LITERATURE CITED

Bayer, R. J., D. G. Greber, and N. H. Bagnall. 2002. Phylogeny of Australian Gnaphalieae (Asteraceae) based on chloroplast and nuclear sequences, the *trn*L intron, the *trn*L/*trn*F intergenic spacer, *mat*K, and ETS. Syst. Bot. 27: 801–814.

de Queiroz, A., R. Lawson, and J. A. Lemos-Espinal. 2002. Phylogenetic relationships of North American garter snakes (*Thamnophis*) based on four mitochondrial genes: How much DNA sequence is enough? Mol. Phyl. Evol. 22: 315–329.

Graybeal, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. 47: 9–17.

Hillis, D. M., D. D. Pollack, J. A. McGuire, and D. J. Zwickl. 2003. Is sparse taxon sampling a problem for phylogenetic inference? Syst. Biol. 52: 124–126.

Hillis D. M., C. Moritz, and B. K. Mable. 1996. Molecular Systematics, 2nd Edition. p. 240. Sinauer Associates, Inc. Sunderland, MA.

Jansen, R. K. 1992. Current research. Plant Molecular Evolution Newsletter 2: 13–14.

Johnson L. A. and D. E. Soltis. 1994. *Mat*K DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. Syst. Bot. 19: 143–156.

Olmstead, R. G., H. J. Michaels, K. M. Scott, and J. D. Palmer. 1992. Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbc*L. Ann. Missouri Bot. Gard. 79: 249–265.

Panero, J. L. and V. A. Funk. 2002. Toward a phylogenetic subfamilial classification for the Compositae. Proc. Biol. Soc. Wash. 115: 909–922.

Pollack, D. D., D. J. Zwickl, J. A. McGuire, and D. M. Hillis. 2002. Incomplete taxon sampling is not a problem for phylogenetic inference. Proc. Natl. Acad. Sci. USA. 98: 10751–10756.

Shinozaki, K., M. Ohme, M. Tanaka, T. Wakasugi, N. Hayshida, T. Matsugayasha, N. Zaita, J. Chunwongse, J. Obokata, K. Yamaguchi-Shinozaki, C. Ohto, K. Torazawa, B. Y. Meng, M. Sugita, H. Deno, T. Kamogashira, K. Yamada, J. Kusuda, F. Takaiwa, A. Kata, N. Tohdoh, H. Shimada, and M. Sugiura. 1986. *Nicotiana tabacum* chloroplast DNA 155,844 bp, pp 1–147. Center for Gene Research, Nagoya University, Chikusa, Nagoya 464, Japan.

Stevens, P. F. (2003). Angiosperm Phylogeny Website. Version 4, May 2003. http://www.mobot.org/MOBOT/research/APweb/.

Taberlet, P., L. Gielly, G. Patou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant. Mol. Biol. 17: 1105–1109.

Weins, J. J. 1998. Combining data sets with different phylogenetic histories. Syst. Biol. 47: 568–581.