# PHYLOGENETIC POSITION AND GENERIC LIMITS OF ARABIDOPSIS <br> (BRASSICACEAE) BASED ON <br> SEQUENCES OF NUCLEAR <br> RIBOSOMAL DNA ${ }^{1}$ 

Steve L. O'Kane, Jr. ${ }^{2}$ and Ihsan A. Al-Shehbaz ${ }^{3}$


#### Abstract

The primary goals of this study were to assess the generic limits and monophyly of Arabidopsis and to investigate its relationships to related taxa in the family Brassicaceae. Sequences of the internal transcribed spacer region (ITS-1 and ITS-2) of nuclear ribosomal DNA, including 5.8 S rDNA, were used in maximum parsimony analyses to construct phylogenetic trees. An attempt was made to include all species currently or recently included in Arabidopsis, as well as species suggested to be close relatives. Our findings show that Arabidopsis, as traditionally recognized, is polyphyletic. The genus, as recircumscribed based on our results, (1) now includes species previously placed in Cardaminopsis and Hylandra as well as three species of Arabis and (2) excludes species now placed in Crucihimalaya, Beringia, Olimarabidopsis, Pseudoarabidopsis, and Ianhedgea. Key words: Arabidopsis, Arabis, Beringia, Brassicaceae, Crucihimalaya, ITS phylogeny, Olimarabidopsis, Pseudoarabidopsis.


Arabidopsis thaliana (L.) Heynh. was first recommended as a model plant for experimental genetics over a half century ago (Laibach, 1943). In recent years, many biologists worldwide have focused their research on this plant. As indicated by Patrusky (1991), the widespread acceptance of $A$. thaliana as a model organism is attributed to the discovery that it has one of the smallest genomes of any flowering plant, a low chromosome number ( $n=5$ ), and that its genome contains few repetitive sequences and little intergenic spacer DNA. A surprising recent finding by Blanc et al. (2000), however, showed that although $A$. thaliana has a remarkably small genome, much of the DNA is present in more than one copy. In addition to these important attributes, A. thaliana has a short generation time (four to six weeks), a small size (dozens can be grown in a small pot), and can easily be grown on synthetic media (Meyerowitz, 1989; Meyerowitz \& Pruitt, 1985). The species has been used extensively in developmental, evolutionary, and ge-
netic studies and has played a major role in understanding the various biological processes in higher plants (see references in Somerville \& Meyerowitz, 2002). The intraspecific phylogeny of $A$. thaliana has been examined by Vander Zwan et al. (2000). Despite the acceptance of A. thaliana as a model organism and the sequencing and mapping of its nuclear genome (The Arabidopsis Genome Initiative, 2000; Cooke et al., 1996), little is known about the other species of Arabidopsis sensu lato, and their closest relatives.

A small number of molecular phylogenetic studies have included a few members of Arabidopsis Heynh. sensu lato (Price et al., 1994; O`Kane et al., 1996; Galloway et al., 1998; Koch et al., 1999, 2000, 2001; Yang et al., 1999). However, none of these studies attempted to examine all of the taxa either currently or previously included in the genus, and they included only a small number of other, sometimes distantly related, genera. The last comprehensive taxonomic account (Schulz, 1924),

[^0]which recognized 11 species, is unsatisfactory. As many as 50 species have been placed in the genus, and, although many of these are now placed in other genera (Al-Shehbaz et al., 1999), their phylogenetic relationships remain unresolved. Monophyly of the genus has not yet been critically determined, and even basic biological information, such as chromosome numbers, generation time, and breeding system of the members of the genus, is lacking.

Generic delimitation is perhaps one of the most difficult and frequently encountered problems in the systematics of the Brassicaceae (Al-Shehbaz, 1973; Rollins, 1993), and Arabidopsis clearly demonstrates this problem. There has been a lack of agreement among taxonomists on the number of species that belong to Arabidopsis and on the characters that indicate its generic boundaries (e.g., Ball, 1993; Löve, 1961; Rollins, 1993). The generic limits of Arabidopsis have been highly unnatural, and there were no well-defined characters separating it from several closely associated genera (but see our taxonomic revision based on the results of this current work, O'Kane \& Al-Shehbaz, 1997; AlShehbaz \& O'Kane, 2002a). Some individual Arabidopsis species have been transferred among several other genera. An example is A. thaliana, which on the basis of Schulz's (1924) synonymy was previously placed in at least nine other genera, including Arabis L., Conringia Adans., Crucifera E. H. L. Krause, Erysimum L., Hesperis L., Nasturtium R. Br., Pilosella Kostel., Sisymbrium L., and Stenophragma Čelak.

Arabidopsis has been closely associated with three different genera, Cardaminopsis (C. A. Mey.) Hayek, Arabis L., and Halimolobos Tausch. Schulz $(1924,1936)$ considered its nearest relative to be Halimolobos, and separated the latter as being coarser herbs with the styles much narrower than the fruit, as opposed to Arabidopsis, which were seen as slender herbs with the styles slightly narrower than the fruit. These alleged differences are not mutually exclusive, and species recognized by him in one genus can easily be accommodated in the other. Löve (1961) and Hylander (1957) indicated a relationship with Cardaminopsis based on natural interspecific hybridization. Hedge (1968) suggested a closer relationship between Arabidopsis and Arabis and indicated that the two differ only in the cotyledonary position. He further suggested that Arabidopsis wallichii (Hook. f. \& Thoms.) Busch probably represents the link between the two genera. An (1987) and Jafri's (1973) transfer of several species from Arabis to Arabidopsis was probably influenced by Hedge's view.

Molecular-based results (O`Kane et al., 1996; Kamm et al., 1995; Mummenhoff \& Hurka, 1994) agree with Löve's (1961) and Hylander's (1957) hypothesis in showing $A$. thaliana to be most closely related to species placed in Cardaminopsis. In anticipation of results published here and to make the names available for floristic works in progress, we previously published the needed nomenclatural innovations for the genus Arabidopsis (O'Kane \& AlShehbaz, 1997) and have established several new genera to accommodate excluded species (Al-Shehbaz et al., 1999). In brief, Arabidopsis includes only A. thaliana and species previously included, or suggested to be, in Cardaminopsis (Jones \& Akeroyd, 1993a, 1993b). Species now excluded from Arabidopsis are placed in Thellungiella O. E. Schulz (Al-Shehbaz \& O’Kane, 1995), Neotorularia Hedge \& J. Léonard (Al-Shehbaz \& O’Kane, 1997), Ianhedgea Al-Shehbaz \& O’Kane (Al-Shehbaz \& O’Kane, 1999), Crucihimalaya Al-Shehbaz et al., Olimarabidopsis Al-Shehbaz et al., and Pseudoarabidopsis Al-Shehbaz et al. (Al-Shehbaz et al., 1999), and Beringia Price et al. (Price et al., 2001).

Our primary objectives are to determine the generic limits of a morphologically coherent, monophyletic Arabidopsis and to reconstruct a robust interpretation of its phylogenetic neighborhood. A well-corroborated phylogeny of the group will allow better evolutionary interpretations to be made of the massive amounts of data now accumulating for $A$. thaliana. Workers will know which species to compare to $A$. thaliana when making interpretations of evolutionary processes. Furthermore, these initial steps will provide a better understanding of morphological character evolution in the Brassicaceae, a family of great economic importance fraught with taxonomic problems related to an under-developed understanding of character evolution and generic delimitation.

## Materials and Methods

## TAXON SAMPLING,

We included representatives of all taxa (at least at the generic level) that are now or have been included in Arabidopsis (e.g., Schulz, 1924; Hedge, 1965; Jafri, 1973; Al-Shehbaz, 1988; Ball, 1993). Taxa shown to lie near Arabidopsis in other molecular studies have also been included (Price et al., 1994; O’Kane et al., 1996; Galloway et al., 1998), as have a sampling of taxa from elsewhere in the Brassicaceae. Phylogenetic trees were initially rooted by Cleome lutea Hook. of the Cleomaceae, a family basal to the Brassicaceae (Rodman et al., 1993; Judd et al., 1994; Hall et al., 2002). Included
taxa, as well as voucher information and some nomenclatural comments, are given in Table 1. Where possible, plant materials were collected in the field and dried in powdered silica gel. In some cases tissue was obtained from plants grown from seeds. Where fresh or dried material was not available, we used tissue from herbarium specimens; the sequence for Arabis scabra All. was obtained from GenBank.

## DNA EXTRACTION, PCR AMPLIFICATION, AND

 SEQUENCINGTotal DNA was extracted from dried tissue ground in a pinch of sterile sand by a modified CTAB procedure as previously described (O'Kane et al., 1996). Amplification of the internal transcribed spacer region (including ITS-1, 5.8S, and ITS-2) was done using the conditions given in O'Kane et al. (1996) except that some ITS regions were amplified as a single unit using primer ITS1$18 S$ ( $5^{\prime}$ CGTAACAAGGTTTCCGTAGG $3^{\prime}$ ) and ITS-4 (White et al., 1990) rather than as two overlapping pieces. PCR products were purified from $0.8 \%$ agarose gels containing 1X TAE using Wizard PCR Preps (Promega). Sequences were obtained either manually using the fmol ${ }^{\text {©® }}$ DNA Sequencing System (Promega) or from the automated sequencer at the University of Iowa using the same primers as were used to amplify the product. GenBank accession numbers are given in Table 1. Sequences of the allotetraploid Arabidopsis suecica (Fries) Norrl. were obtained from cloned PCR products as previously reported (O'Kane et al., 1996).

SEQUENCE ALIGNMENT AND PHYLOGENETIC ANALYSES

Sequences were aligned with the computer program MALIGN 2.7 (Wheeler \& Gladstein; available at $<\mathrm{ftp}$ ://ftp.amnh.org/pub/people/wheeler/malign/>) using the following empirically determined parameters: internal 7 , extragap 5 , leading 3 , trailing 3 , matrix 0 323303223033230 , aspr, spr, quick, keepa2, keept3, and score 2 . The matrix parameters weight transversions as 3 , transitions as 2 , intitial gaps as 7 , gap extensions as 5 , and initial and ending gaps as 3 (not a factor in our sequences). Most-parsimonious trees were found using PAUP* 4.0b4 (Swofford, 2000). In our analyses all characters were considered to be of equal weight and gaps were coded as missing data. Two hundred and fifty replicates of random addition using Fitch parsimony were performed using Tree Bisection Reconnection (TBR), Mulpars, multistate $=$ polymorphism, gaps coded as missing, and Collapse branches if maximum length is zero. Boot-
strap support was obtained from 500 replicates using a single round of simple taxon addition. Decay values (Bremer, 1988; Donoghue et al., 1992) were found using the AutoDecay program 4.01 of Eriksson (1998). Clade Significance (Lee, 2000) was implemented in PAUP by the AutoCladeS program (T. Eriksson; available at <http://www.bergianska.se/index_forskning_soft.html $>$ ). We found that this new measure of support indicates those clades that have the highest support based on the other two measures. The information content of the data was assessed by the gl statistic (Hillis \& Huelsenbeck, 1992) based on 100,000 random trees and by the Permutation Tail Probability (PTP) (Faith \& Cranston, 1992) based on 200 heuristic searches of randomized data (PAUP* parameters as above except simple addition was used rather than random addition).

## Results

In nearly all samples there was no evidence of heterogeneity among individual ITS copies. Rarely, two different bases were present at a given position as indicated by two bands on an autoradiograph or as two clear peaks on a chromatogram. In these cases the base position was coded using the appropriate ambiguity code. We interpret this rare "heterozygosity" as incomplete homogenization of the ITS copies and not as evidence of hybridization; an individual sequence would show much more variation if hybridization were involved. Although initial analyses used Cleome lutea as the outgroup, inclusion of this taxon added to the complexity of sequence gaps and to areas with ambiguous alignments. Of the taxa included in this study, Berteroella maximowiczii (Palib.) O. E. Schulz was found to be strongly supported as the basal-most taxon. Further analyses, then, used B. maximowiczii as the outgroup. The resulting multiple alignment of the internal transcribed spacer region (ITS) was 716 base-pairs in length. Sequences are deposited in GenBank (accession numbers in Table 1), and the full alignment is available from the first author. Two regions of the alignment that were extremely sensitive to alignment parameters and could not be improved by eye (114-139 and 461-507) were not used in the phylogenetic analyses. In all, 384 bases were invariant, 79 were parsimony uninformative, and 253 were parsimony informative.

Parsimony searches yielded 24 distinct mostparsimonious trees of length 951, consistency index (CI) 0.48 , consistency index excluding uninformative characters (CIU) 0.43 , retention index (RI) 0.73 , and rescaled consistency index (RC) of 0.35 . The gl statistic of the data was -0.5982 , which
Table 1. Voucher information and GenBank accession numbers for samples included in this study. Voucher information is given as collector, collector number, and herbarium of deposit; identical superscripts indicate samples with identical rDNA sequences.

| Taxon | Voucher information | Source locality | Number |
| :---: | :---: | :---: | :---: |
| Arabidopsis arenosa (L.) Lawalrée | O'Kane \& Mirek 3650 (MO) ${ }^{1}$ | Poland | U52188 |
|  | Tzvelev et al. 79 (LE) ${ }^{1}$ | Russia | U52187 |
|  | O'Kane \& Kuciel 3664 (MO) ${ }^{1}$ | Poland | U43233 |
| A. cebennensis (DC.) O'Kane \& Al-Shehbaz | Coste s.n. (1890) )BUCS) | France | AF137545 |
| A. croatica (Schott) O'Kane \& Al-Shehbaz | Kummert s.n. (1977) (W) | Yugoslavia | AF137546 |
| A. halleri (L.) O'Kane \& Al-Shehbaz subsp. gemmifera (Matsum.) O'Kane \& Al-Shehbaz | O'Kane 3693 (MO) ${ }^{\text {2 }}$ | Japan | AF137544 |
|  | O'Kane \& Berkutenko 3678 (MO) ${ }^{2}$ | Russian Far East | AF137542 |
|  | O'Kane \& Probatova 3683 (MO) ${ }^{2}$ | Russian Far East | AF137543 |
| A. halleri (L.) O'Kane \& Al-Shehbaz subsp. halleri | Koteja 123 (MO) | Poland | AF137541 |
| A. halleri (L.) O'Kane \& Al-Shehbaz subsp. ovirensis (Wulfen) O’Kane \& Al-Shehbaz | O'Kane \& Dihoru 3611 (MO) | Romania | AF137540 |
| A. lyrata (L.) O'Kane \& Al-Shehbaz subsp. kamchatica (Fisch. ex DC.) O'Kane \& Al-Shehbaz | O'Kane \& Berkutenko 3679 (MO) ${ }^{3}$ | Russian Far East |  |
|  | $\text { O'Kane } 3684(\mathrm{MO})^{3}$ | Japan | U96268 |
|  | O'Kane \& Berkutenko 3681 (MO) ${ }^{3}$ | Russian Far East | U96267 |
| A. lyrata (L.) O'Kane \& Al-Shehbaz subsp. petraea (L.) O'Kane \& Al-Shehbaz | Kharkevich et al. 22-7-83 (VLA) ${ }^{4}$ | Russian Far East | AF137539 |
|  | Barkalov \& Bezdeleva 6-8-89 (VLA) ${ }^{4}$ | Russian Far East | U96270 |
|  | Probatova \& Seledets 7-8-80 (VLA) ${ }^{4}$ | Russian Far East | U96269 |
| A. neglecta (Schult.) O'Kane \& Al-Shehbaz | Hornickova s.n. (1977) (BRA) ${ }^{1}$ | Slovakia | U52186 |
| A. suecica (Fr.) Norrl. | Lampinen 2450 (H) | Finland | U52185 |
| A. thaliana (L.) Heynh. (sample 1) | Podlech 17544 (MO) | Afghanistan | U43225 |
| (sample 2) | O'Kane \& Krahulec 3638 (MO) | Czech Republic | U43224 |
| Arabis alpina L. | O'Kane \& Dihoru 3618 (MO) | Romania | AF137559 |
| Arabis drummondii S. Watson | O'Kane 3676 (MO) | Montana, U.S.A. | AF137575 |
| Arabis flagellosa Miq. | O'Kane \& Kato 3689 (MO) | Japan | AF137560 |
| Arabis lyallii S. Watson | O'Kane 3673 (MO) | Montana, U.S.A. | AF137561 |
| Arabis nuttallii B. L. Rob. | O'Kane 3672 (MO) | Montana, U.S.A. | AF137562 |
| Arabis pendula L. <br> Arabis scabra All. | O'Kane \& Berkutenko 3682 (MO) | Russian Far East | AF137572 |
| Arabis scabra All. | From GenBank |  | X98630 |
| Beringia bursifolia (DC.) R. A. Price, Al-Shehbaz \& O'Kane subsp. bursifolia | Abolin 53 (LE) | Russia | AF137557 |
| Berteroella maximowiczii (Palib.) O. E. Schulz | Cheo \& Yen 208 (GH) | China | AF137573 |
| Braya glabella (Richardson) S. Watson | Neely 3174 (CS) | Colorado, U.S.A. | AF137578 |
| Camelina microcarpa Andrz. ex DC. | O'Kane \& Dihoru 3596 (MO) | Romania | AF137574 |
| Capsella bursa-pastoris (L.) Medik. | Al-Shehbaz \& O'Kane 9401 (MO) | Missouri, U.S.A. | AF137570 |
| Cleome lutea Hook. <br> Crucihimalaya himalaica (Edgew) Al-Shehbaz, O'Kane \& R. A. Price | O'Kane 3773 (ISTC) | Utah, U.S.A. | AF137588 |
| Crucihimalaya himalaica (Edgew.) Al-Shehbaz, O'Kane \& R. A. Price | Al-Shehbaz 9354 (MO) | China | AF137553 |

Table 1. Continued.

| Taxon | Voucher information | Source locality | Number |
| :---: | :---: | :---: | :---: |
| Crucihimalaya kneuckeri (Bornm.) Al-Shehbaz, O'Kane \& R. A. Price | Kaiser 632 (S) | Egypt (Sinai) | AF137550 |
| Crucihimalaya lasiocarpa (Hook. f. \& Thoms.) Al-Shehbaz, O'Kane \& R. A. Price | Al-Shehbaz 9333 (MO) | China | AF137556 |
| Crucihimalaya mollissima (C. A. Mey.) Al-Shehbaz, O'Kane \& R. A. Price | Wendelbo \& Ekberg 9642 (E) | Afghanistan | AF137552 |
| Crucihimalaya ovczinnikovii (Botsch.) Al-Shehbaz, O'Kane \& R. A. Price | Kamelin s.n. (LE) | Tajikistan | AF137551 |
| Crucihimalaya stricta (Cambess.) Al-Shehbaz, O'Kane \& R. A. Price | Rechinger 30655 (W) | Pakistan | AF137554 |
| Crucihimalaya wallichii (Hook. f. \& Thoms.) Al-Shehbaz, O'Kane \& R. A. Price | Palmer 66 (K) | Afghanistan | AF137555 |
| Dichasianthus subtilissimus (Popov) Ovez. \& Junussov | Chukavinah \& Kenzikaeva 5473 (LE) | Tajikistan | AF137594 |
| Dimorphocarpa wislizenii (Engelm.) Rollins | Porter 4467 (SJC) | New Mexico, U.S.A. | AF137593 |
| Dithyrea californica Harv. | Heil 6401 (SJC) | Mexico | AF137592 |
| Drabopsis nuda (Belang. ex Boiss.) Stapf | Rechinger 54514 (MO) | Iran | AF137577 |
| Eutrema penlandii Rollins | O'Kane 2157 (CS) | Colorado, U.S.A. | AF137580 |
| Halimolobos palmeri (Hemsl.) O. E. Schulz var. acutiloba Rollins | Magaña 5961 (MO) | Mexico | AF137569 |
| Halimolobus diffusa A. Gray var. jaegeri (Munz) Rollins | Morefeld et al. 4549 (MO) | Nevada, U.S.A. | AF137567 |
| Ianhedgea minutiflora (Hook. f. \& Thoms.) Al-Shehbaz \& O'Kane | Rasoul 3541 (W) | Afghanistan | AF137568 |
| Lyrocarpa coulteri Hook. \& Harv. | Heil 6399 (SJC) | Mexico | AF137591 |
| Neotorularia gamosepala (Hedge) Al-Shehbaz \& O'Kane | Podlech 12379 (E) | Afghanistan | AF137565 |
| Neotorularia humilis (C. A. Mey.) Hedge \& J. Léonard | Shao-Xing s.n. (PE) | China | AF137566 |
| Neotorularia torulosa (Desf.) Hedge \& J. Léonard | Rechinger 40534 (W) | Iran | AF137571 |
| Nerisyrenia linearifolia (S. Watson) Greene | Sawyer 77 (SJC) | New Mexico, U.S.A. | AF137587 |
| Neslia paniculata (L.) Desv. subsp. paniculata | Bzowska \& Rabczak 329 (MO) | Poland | AF137576 |
| Olimarabidopsis cabulica (Hook. f. \& Thoms.) Al-Shehbaz, O'Kane \& R. A. Price | Grubov et al. s.n. (LE) | Kyrgyzstan | AF137548 |
| Olimarabidopsis pumila (Stephan) Al-Shehbaz, O'Kane \& R. A. Price | Hedge \& Wendelbo W8650 (E) | Afghanistan | AF137549 |
| Olimarabidopsis umbrosa (Botsch. \& Vved.) Al-Shehbaz, O'Kane \& R. A. Price | Rechinger 18359 (US) | Afghanistan | AF137547 |
| Paysonia densipila (Rollins) O'Kane \& Al-Shehbaz | Rebman \& Dierig 2894 (ISTC) | Tennessee, U.S.A. | AF137586 |
| Paysonia stonensis (Rollins) O'Kane \& Al-Shehbaz | Rebman 2892 (ISTC) | Tennessee, U.S.A. | AF137585 |
| Physaria acutifolia Rydb. | O'Kane 3706 (ISTC) | Utah, U.S.A. | AF137582 |
| Physaria didymocarpa (Hook.) A. Gray var. didymocarpa | O'Kane 3794 (ISTC) | Montana, U.S.A. | AF137583 |
| Physaria pruinosa (Greene) O'Kane \& Al-Shehbaz | O'Kane \& Anderson 3739 (ISTC) | Colorado, U.S.A. | AF137584 |
| Pseudoarabidopsis toxophylla (M. Bieb.) Al-Shehbaz, O'Kane \& R. A. Price | Skvortsov s.n. (1987) (MO) | Kazakstan | AF137558 |
| Smelowskia calycina (Stephan ex Willd.) C. A. Mey. | O'Kane 3675 (MO) | Montana, U.S.A. | AF137581 |
| Sphaerocardamum macropetalum (Rollins) Rollins | Rollins \& Roby 7489 (MO) | Coahuila, Mexico | AF137589 |
| Synthlipsis greggii A. Gray | Rollins \& Tryon 5893 (MO) | Nuevo Leon, Mexico | AF137590 |
| Thellungiella halophila (C. A. Mey.) O. E. Schulz | Pavlov et al. 1192 (LE) | Kazakstan | AF137563 |
| Thellungiella parvula (Schrenk) Al-Shehbaz \& O'Kane | Davis \& Dodds 18677 (BM) | Turkey | AF137579 |
| Thellungiella salsuginea (Pall.) O. E. Schulz | Ledingham 7937 (MO) | Canada | AF137564 |

indicates strong phylogenetic signal in the data ( $P$ $<0.01$ ). The Permutation Tail Probability (PTP) also indicated strong signal $(P=0.005)$. Figure 1 shows the strict consensus tree of the 24 most parsimonious trees.

## Discussion

## RELATIONSHIPS AND CIRCUMSCRIPTION OF ARABIDOPSIS

The relationships among the species included in this study are almost entirely consistent with results previously published for smaller taxon samples in the Brassicaceae focusing on Arabidopsis (e.g., Price et al., 1994; Galloway et al., 1998; Koch et al., 1999, 2000, 2001; Yang et al., 1999). Like those studies, our research indicates that Arabidopsis as traditionally circumscribed is a highly artificial group. In fact, even the tribe Sisymbrieae, the traditional placement for Arabidopsis (Schulz, 1924, 1936; Al-Shehbaz, 1984, 1988), is itself artificial. The confused circumscription of Arabidopsis, as based on morphological grounds, was noted in previous taxonomic treatments. Hylander (1957: 602), for example, recognized that if Cardaminopsis and Arabidopsis are combined, as seemed likely, the limits of Arabidopsis "would thereby be considerably widened-or, perhaps more correctly, drawn in quite another way." Jones (1964) also indicated that two species of Arabis, A. pedemontana Boiss. and $A$. cebennensis DC., might best be included in Cardaminopsis. Thus, at least as early as 1964, taxonomic problems were anticipated in Arabidopsis, Cardaminopsis, and Arabis.

Results from our study are sufficient to allow a revision of the taxonomy of the genus Arabidopsis. A strongly supported clade (bootstrap support $97 \%$, decay index 6, clade support 0.014 ; see Fig. 1) containing $A$. thaliana, the type species of the genus, defines the limits of a recircumscribed genus Arabidopsis. As stated above, we have anticipated the publication of these results by redefining the circumscription of Arabidopsis (O'Kane \& Al-Shehbaz, 1997), transferring species to previously recognized genera (Al-Shehbaz \& O’Kane, 1995, 1997), and lastly by erecting several new genera for species previously included in Arabidopsis (AlShehbaz et al., 1999; Al-Shehbaz \& O’Kane, 1999, 2002a; Price et al., 2001). Names in bold type in Figure 1 represent new genera erected for species previously included in Arabidopsis by other authors, and the bold letter "A" indicates species previously included in Arabidopsis. In every case we have made genera monophyletic (sensu Hennig, 1966; holophyletic of Ashlock, 1971). Species previously
included in Arabidopsis are now placed in Beringia, Crucihimalaya, Olimarabidopsis, Pseudoarabidopsis, Ianhedgea, Neotorularia, and Thellungiella. Arabidopsis sensu novo is distinguished from other genera in the Brassicaceae by having short petiolate but not auriculate or amplexicaul cauline leaves, the presence of simple trichomes, these often mixed with few-forked ones but not stellate hairs, well-defined basal rosettes at least in young plants, white to lavender (rarely almost purple) but never yellow flowers, erect to slightly ascending non-saccate or slightly saccate inner sepals, siliques at least slightly torulose, much longer than they are wide and glabrous, compressed (rarely subterete to terete), seeds uniseriate in the silique, and cotyledons accumbent or rarely incumbent (O’Kane \& Al-Shehbaz, 1997). Habit ranges from annual to short- or long-lived perennials. Chromosome numbers vary from $x=n=5$ in A. thaliana to $x=8$ in the remaining species except for $A$. suecica, which is an allotetraploid $(2 n=26)$ derived from $A$. thaliana $(2 n=10)$ and $A$. arenosa (L.) Lawalrée ( $2 n=16$ ) (Mummenhoff \& Hurka, 1994; O'Kane et al., 1996, and references therein). Keys to the species and subspecies are given in O'Kane and Al-Shehbaz (1997). As circumscribed here, the genera Cardaminopsis and Hylandra Á. Löve are united with Arabidopsis. Arabidopsis is a monophyletic genus consisting of $A$. arenosa, $A$. cebennensis (DC.) O'Kane \& Al-Shehbaz, A. croatica (Schott) O'Kane \& Al-Shehbaz, A. halleri (L.) O'Kane \& Al-Shehbaz, A. lyrata (L.) O'Kane \& AlShehbaz, A. neglecta (Schultes) O’Kane \& Al-Shehbaz, A. suecica, and A. thaliana. Although we have no sequences of $A$. pedemontana (Boiss.) O'Kane \& Al-Shehbaz, it, too, clearly belongs in Arabidopsis based on its morphological relationship to $A$. cebennensis.

All other species previously included in Arabidopsis are more distantly related, especially those Himalayan species now included in Crucihimalaya and the Middle Eastern and central Asian Olimarabidopsis. Morphologically, Crucihimalaya differs from Arabidopsis in that it has at least some stellate trichomes, whereas Arabidopsis has forked trichomes. Olimarabidopsis differs from Arabidopsis in its yellow, rather than white or lavender petals, pubescent, rather than glabrous fruits, and auriculate rather than attenuate or petiolate cauline leaves. Al-Shehbaz et al. (1999) presented an analysis of all species formerly placed in Arabidopsis. Based on the results presented here, Arabidopsis is a genus of circumboreal and circum-north-temperate species. Most species, however, are confined to Europe. Workers conducting comparative research us-


Figure 1. Strict consensus tree of 24 most-parsimonious trees. Whole numbers indicate bootstrap support; values in parentheses are the Decay Index; decimal values are Clade Significance. Dashed branches indicate branches with $<50 \%$ bootstrap support. Large bold " $A$ " indicates species variously placed in Arabidopsis in previous taxonomic treatments. Large asterisks indicate species traditionally placed in Arabis. Genera in bold type are segregates from Arabidopsis named elsewhere as a result of these analyses (see Discussion).
ing A. thaliana as a model organism can now confidently use the other species of a bettercircumscribed Arabidopsis, all of which are found in the sister group to $A$. thaliana, as experimental organisms. This sister group relationship implies that all species of Arabidopsis are equally related to $A$. thaliana. Unfortunately, the sister group to the genus Arabidopsis cannot be given with confidence. A trichotomy appears below Arabidopsis (Fig. 1): (Arabidopsis clade)("Capsella-Pseudoarabidopsis" clade)("Crucihimalaya-Olimarabidopsis" clade). Galloway et al. (1998), using sequences of arginine decarboxylase and a much smaller taxon sample ( 28 species from throughout the family), confidently placed Capsella in a sister group relationship to Arabidopsis. Unfortunately, their analysis did not include any members of the "Crucihimalaya-Olimarabidopsis" clade. Koch et al. (1999) found, using ITS sequences (and with low bootstrap support), Capsella to be sister to Arabidopsis and Olimarabidopsis, but their analysis did not include any members of Crucihimalaya. Koch et al. (2001) obtained similar results using plastidic mat K and nuclear Chs sequences. Additional work is needed to resolve this issue, but assuming that the "CapsellaPseudoarabidopsis clade" is sister to Arabidopsis appears to be a valid working hypothesis.

TAXONOMIC IMPLICATIONS ELSEWHERE IN THE BRASSICACEAE

Although our intent was not to study the genus Arabis in any detail, our results mirror those of Koch et al. $(1999,2000,2001)$ in showing Arabis, as traditionally recognized, to be polyphyletic even after A. lyrata L., A. pedemontana, and A. cebennensis are transferred to Arabidopsis (Fig. 1). In our analysis, the genus Boechera Á. Löve \& D. Löve seems to be the proper home for $x=7$ species like A. lyallii A. Gray and A. drummondii S. Watson, though not all of the necessary generic transfers have been made. Arabis in its strictest interpretation will consist only of those species in the clade with A. alpina L., the lectotype of Arabis. Arabis glabra (L.) Bernh. belongs to Turritis L. and A. pauciflora (Grimm) Garcke belongs to Fourraea Greuter \& Burdet (Koch et al., 1999). But to which genus does $A$. pendula L. or $A$. turrita L. belong (Koch et al., 1999)? Including species once thought to be related to Arabidopsis in our study has also raised other taxonomic questions. Neotorularia, Braya Sternb. \& Hoppe, and Dichasianthus Ovcz. \& Junussov form a well-supported clade (Fig. 1) with Neotorularia being paraphyletic. Thellungiella and

Eutrema R. Br. also form a well-supported clade, with Thellungiella being paraphyletic.

A surprising result of our study was the discovery of a clade of species all possessing pollen with more than the usual three colpi (see Polycolpate clade, Fig. 1). Palynological studies (Rollins, 1979; Rollins \& Banerjee, 1979) showed that among some genera thought not to be closely related in the Brassicaceae, colpi range from four to ten. These genera, according to Schulz (1936), are as follows: Physaria (Nutt. ex Torr. \& A. Gray) A. Gray (tribe Lepidieae, subtribe Physariinae), Dithyrea Harv. and its recent segregate Dimorphocarpa Rollins (Lepidieae, Iberidinae), Lyrocarpa Hook. \& Harv. (Lepidieae, Lyrocarpinae), Nerisyrenia Greene (as Greggia A. Gray) and Synthlipsis A. Gray (Lepideae, Capsellinae), and Lesquerella S. Watson (Drabeae). The results presented here also suggest that within this "polycolpate clade" taxonomic revisions are needed in Lesquerella and Physaria. We have recently united these two genera (Al-Shehbaz \& O'Kane, 2002b), except that the auriculate-leaved species formerly placed in Lesquerella are recognized as a distinct genus, Paysonia O'Kane \& AlShehbaz (O'Kane \& Al-Shehbaz, 2002).

Future work in the family will certainly yield further taxonomic alignments since there is rampant morphological convergence (Al-Shehbaz et al., 1999; Koch et al., 1999) and because previous taxonomy in the family has relied heavily on fruit morphology (e.g., Rollins, 1993; Al-Shehbaz, 1984) to the exclusion of floral and vegetative features (AlShehbaz et al., 1999). Molecular techniques in concert with a reevaluation of morphological characters are rapidly reshaping our understanding of the family (e.g., Bailey \& Doyle, 1999; Bailey et al., 2002; Bowman et al., 2000; Koch et al., 1999, 2000, 2001; Mummenhoff \& Koch, 1994; Mummenhoff et al., 1997a, b, 2001a, b; Price \& Palmer, 1996; Rodman et al., 1996; Warwick \& Black, 1993, 1997). Characterizing the membership of Arabidopsis and sketching its relationships to related genera, we believe, contributes to this growing body of knowledge.

Literature Cited
Al-Shehbaz, I. A. 1973. The biosystematics of the genus Thelypodium (Cruciferae). Contr. Gray Herb. 204: 3148.

- 1984. The tribes of Cruciferae (Brassicaceae) in the southeastern United States. J. Arnold Arbor. 65: 343-373.
. 1988. The genera of Arabideae (Cruciferae; Brassicaceae) in the southeastern United States. J. Arnold Arbor. 69: 85-106.
_ \& S. L. O'Kane, Jr. 1995. Placement of Arabidop-
sis parvula in Thellungiella (Brassicaceae). Novon 5: 309-310.
$-\&-$ 1997. Arabidopsis gamosepala and $A$. tuemurnica belong to Neotorularia (Brassicaceae). Novon 7: 93-94.
—_\& —— 1999. Ianhedgea (Brassicaceae), a new generic name replacing the illegitimate Microsisymbrium. Edinburgh J. Bot. 56: 321-327.
$-\&-.2002 \mathrm{a}$. Taxonomy and phylogeny of Arabidopsis (Brassicaceae). (22 August, 2002), C. R. Somerville \& E. M. Meyerowitz (editors), The Arabidopsis Book. American Society of Plant Biologists, Rockville, Maryland. doi/10.1199/tab. 0001, <http:// www.aspb.org/publications/arabidopsis/>.
$-\& \xrightarrow{-}$ 2002b. Lesquerella is united with Physaria (Brassicaceae). Novon 12: 319-329.
$-\_$- \& R. A. Price. 1999. Generic placement of species excluded from Arabidopsis. Novon 9: 296307.

An, Z. X. 1987. Arabidopsis. Pp. 280-288 in T.-Y. Cheo (editor), Flora Reipublicae Popularis Sinicae, Vol. 33. Science Press, Beijing.
The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408: 796-815.
Ashlock, P. D. 1971. Monophyly and associated terms. Syst. Zool. 20: 63-69.
Bailey, C. D. \& J. J. Doyle. 1999. Potential phylogenetic utility of the low-copy nuclear gene pistillata in dicotyledonous plants: Comparison to nrDNA ITS and $\operatorname{trn} \mathrm{L}$ intron in Sphaerocardamum and other Brassicaceae. Molec. Phylogenet. Evol. 13: 20-30.

- , R. A. Price \& J. J. Doyle. 2002. Systematics of the Halimolobine Brassicaceae: Evidence from three loci and morphology. Syst. Bot. 27: 318-332.
Ball, P. W. 1993. Arabidopsis. Pp. 322-323 in T. G. Tutin, N. A. Burges, A. O. Chater, J. R. Edmondson, V. H. Heywood, D. M. Moore, D. H. Valentine, S. M. Walters \& D. A. Webb (editors), Flora Europaea, ed. 2, Vol. 1. Cambridge Univ. Press, Cambridge.
Blanc, G., A. Barakat, R. Guyot, R. Cooke \& M. Delseny. 2000. Extensive duplication and reshuffling in the Ar abidopsis Genome. Pl. Cell 12: 1093-1101.
Bowman, J. L., H. Brüggemann, J.-Y. Lee \& K. Mummenhoff. 1999. Evolutionary changes in floral structure within Lepidium L. (Brassicaceae). Int. J. Pl. Sci. 160: 917-929.
Bremer, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42: 795-803.
Cooke, R., M. Raynal, M. Laudié, F. Grellet, M. Delseny, P.-C. Morris, D. Guerrier, J. Giraudat, F. Quigley, G. Clabault, Y.-F. Li, R. Mache, M. Krivitzky, I. J.-J. Gy, M. Kreis, A. Lecharny, Y. Parmentier, J. Marbach, J. Fleck, B. Clément, G. Philipps, C. Hervé, C. Bardet, D. Tremousaygue, B. Lescure, C. Lacomme, D. Roby, M.F. Jourjon, P. Chabrier, J.-L. Charpenteau, T. Desprez, J. Amselem, H. Chiapello \& H. Höfte. 1996. Further progress towards a catalogue of all Arabidoposis genes: Analysis of a set of 5000 non-redundant ESTs. PI. J. 9: 101-124.
Donoghue, M. J., R. G. Olmstead, J. F. Smith \& J. D. Palmer. 1992. Phylogenetic relationships of Dipsacales based on $r b c \mathrm{~L}$ sequences. Ann. Missouri Bot. Gard. 79: 333-345.
Eriksson, T. 1998. AutoDecay, ver. 4.01 (program distrib-
uted by the author). Department of Botany, Stockholm University, Stockholm.
Faith, D. P. \& P. S. Cranston. 1992. Probability, parsimony, and Popper. Syst. Biol. 41: 252-257.
Galloway, G. L., R. L. Malmberg \& R. A. Price. 1998. Phylogenetic utility of the nuclear gene arginine decarboxylase: An example from Brassicaceae. Molec. Biol. Evol. 15: 1312-1320.
Hall, J. C., K. J. Sytsma \& H. H. Iltis. 2002. Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. Amer. J. Bot. 89: 1826-1842.
Hedge, I. C. 1968. Arabidopsis. Pp. 489-490 in K. H. Rechinger (editor), Flora Iranica, Vol. 57. Akademische Druck- u. Verlagsanstalt, Graz, Austria.
Hennig, W. 1966 (1999 reissue). Phylogenetic Systematics. Univ. Illinois Press, Chicago.
Hillis, D. M. \& J. P. Huelsenbeck. 1992. Signal, noise, and reliability in molecular phylogenetic analysis. Heredity 83: 189-195.
Hylander, N. 1957. Cardaminopsis suecica (Fr.) Hiit., A northern amphidiploid species. Bull. Jard. Bot. Etat 27: 591-604.
Jafri, S. M. H. 1973. Brassicaceae. Pp. 1-308 in E. Nasir \& S. I. Ali (editors), Flora of West Pakistan, Vol. 55. Ferozsons, Karachi.
Jones, B. M. G. 1964. Arabis. Pp. 290-294 in T. G. Tutin, V. H. Heywood, N. A. Burges, D. H. Valentine, S. M. Walters \& D. A. Webb (editors), Flora Europaea, Vol. 1. Cambridge Univ. Press, Cambridge.
- \& J. R. Akeroyd. 1993a. Cardaminopsis. Pp. 351352 in T. G. Tutin, N. A. Burges, A. O. Chater, J. R. Edmondson, V. H. Heywood, D. M. Moore, D. H. Valentine, S. M. Walters \& D. A. Webb (editors), Flora Europaea, ed. 2, Vol. 1. Cambridge Univ. Press, Cambridge.
— \& —. 1993b. Arabis. Pp. 352-356 in T. G. Tutin, N. A. Burges, A. O. Chater, J. R. Edmondson, V. H. Heywood, D. M. Moore, D. H. Valentine, S. M. Walters \& D. A. Webb (editors), Flora Europaea, ed. 2, Vol. 1. Cambridge Univ. Press, Cambridge.

Judd, W. S., R. W. Sanders \& M. J. Donoghue. 1994. Angiosperm family pairs: Preliminary cladistic analyses. Harvard Pap. Bot. 5: 1-51.
Kamm, A., I. Galasso, T. Schmidt \& J. S. Heslop-Harrison. 1995. Analysis of a repetitive DNA family from Arabidopsis arenosa and relationships between Arabidopsis species. Pl. Molec. Biol. 27: 853-862.
Koch, M., J. Bishop \& T. Mitchell-Olds. 1999. Molecular systematics and evolution of Arabidopsis and Arabis. Pl. Biol. 1: 529-537.
——, B. Haubold \& T. Mitchell-Olds. 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in Arabidopsis, Arabis and related genera. Molec. Biol. Evol. 17: 1483-1498.
$-\_$, \& . 2001. Molecular systematics of the Brassicaceae: Evidence from coding plastidic matK and nuclear Chs sequences. Amer. J. Bot. 88: 534-544.
Laibach, F. 1943. Arabidopsis thaliana (L.) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen. Bot. Archiv 44: 439-455.
Lee, M. S. Y. 2000. Tree robustness and clade significance. Syst. Biol. 49: 829-836.
Löve, Á. 1961. Hylandra-A new genus of Cruciferae. Svensk Bot. Tidskr. 55: 211-217.
Meyerowitz, E. M. 1989. Arabidopsis, a useful weed. Cell 56: 263-269.
\& R. E. Pruitt. 1985. Arabidopsis thaliana and plant molecular genetics. Science 22: 1214-1218.
Mummenhoff, K. \& H. Hurka. 1994. Subunit polypeptide composition of Rubisco and the origin of allopolyploid Arabidopsis suecica (Brassicaceae). Biochem. Syst. Ecol. 22: 807-812.
\& M. Koch. 1994. Chloroplast DNA restriction site variation and phylogenetic relationships in the genus Thlaspi sensu lato (Brassicaceae). Syst. Bot. 19: 73-88.
-, A. Franzke \& M. Koch. 1997a. Molecular data reveal convergence in fruit characters used in the classification of Thlaspi s.I. (Brassicaceae). Bot. J. Linn. Soc. 125: 183-199.
$—,-\&-$ 1997b. Molecular phylogenetics of Thlaspi s.l. (Brassicaceae) based on chloroplast DNA restriction site variation and sequences of the internal transcribed spacers of nuclear ribosomal DNA. Canad. J. Bot. 75: 469-482.
_, H. Brüggemann \& J. Bowman. 2001a. Chloroplast DNA phylogeny and biogeography of the genus Lepidium (Brassicaceae). Amer. J. Bot. 88: 2051-2063.
, U. Coja \& H. Brüggemann. 2001b. Pachyphrag$m a$ (Brassicaceae) revisited: Molecular data indicate close relationship to Thlaspi s.str. Folia Geobot. 36: 293-302.
O’Kane, S. L., Jr. \& I. A. Al-Shehbaz. 1997. A synopsis of Arabidopsis (Brassicaceae). Novon 7: 323-327.
$\_\&-$ 2002. Paysonia, a new genus segregated from Lesquerella (Brassicaceae). Novon 12: 379381.
B. A. Schaal \& I. A. Al-Shehbaz. 1996. The origins of Arabidopsis suecica (Brassicaceae) as indicated by nuclear rDNA sequences. Syst. Bot. 21: 559-566.
Patrusky, B. 1991. Drosophila botanica (the fruit fly of plant biology). Mosaic 22: 32-43.
Price, R. A. \& J. D. Palmer. 1996. Chloroplast DNA based phylogenies and fruit shape in the tribe Arabideae (Brassicaceae). Amer. J. Bot. 83: S187-S188.
, J. D. Palmer \& I. A. Al-Shehbaz. 1994. Systematic relationships of Arabidopsis: A molecular and morphological perspective. Pp. 7-19 in E. M. Meyerowitz \& C. R. Somerville (editors), Arabidopsis. Cold Spring Harbor Laboratory Press, New York.
$—$ I. A. Al-Shehbaz \& S. L. O'Kane, Jr. 2001. Beringia (Brassicaceae), a new genus of Arabidopsid affinities from Russia and North America. Novon 11:332336.

Rodman, J. E., R. A. Price, K. Karol, E. Conti, K. J. Sytsma \& J. D. Palmer. 1993. Nucleotide sequences of
the $r b c \mathrm{~L}$ gene indicate monophyly of mustard oil plants. Ann. Missouri Bot. Gard. 80: 686-699.

- K. G. Karol, R. A. Price \& K. J. Sytsma. 1996. Molecules, morphology, and Dahlgren's expanded order Capparales. Syst. Bot. 21: 289-307.
Rollins, R. C. 1979. Dithyrea and a related genus (Cruciferae). Publ. Bussey Inst. Harvard Univ. 1979: 3-32. - 1993. The Cruciferae of Continental North America. Stanford Univ. Press, Stanford.
\& U. C. Banerjee. 1979. Pollens of the Cruciferae. Publ. Bussey Inst. Harvard Univ. 1979: 33-64.
Schulz, O. E. 1924. Cruciferae-Sismbrieae. In: A. Engler (editor), Pflanzenreich IV. 105(Heft 86): 1-388. Verlag von Wilhelm Engelmann, Leipzig.
- 1936. Cruciferae. In A. Engler \& K. Prantl (editors), Natürl. Pflanzenfam., ed. 2, 17B: 227-658. Verlag von Wilhelm Engelmann, Leipzig.
Somerville, C. R. \& E. M. Meyerowitz (editors). 2002. The Arabidopsis Book. <http://www.aspb.org/publications/ arabidopsis/>.
Swofford, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer, Sunderland, Massachusetts.
Vander Zwan, C., S. A. Brodie \& J. J. Campanella. 2000. The intraspecific phylogenetics of Arabidopsis thaliana in worldwide populations. Syst. Bot. 25: 47-59.
Warwick, S. I. \& L. D. Black. 1993. Molecular relationships in subtribe Brassicinae (Cruciferae, tribe Brassiceae). Canad. J. Bot. 71: 906-918.
$-\&-1997$. Phylogenetic implications of chloroplast DNA restriction site variation in subtribes Raphaninae and Cakilinae (Brassicaceae, tribe Brassiceae). Canad. J. Bot. 75: 960-973.
White, T. J., T. Bruns, S. Lee \& J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315-322 in M. A. Innis, D. H. Gelfand, J. J. Sninsky \& T. J. White (editors), PCR Protocols. Academic Press, New York.
Yang, W.-W., K.-N. Lai, P.-Y. Tai, D.-P. Ma \& W.-H. Li. 1999. Molecular phylogenetic studies of Brassica, Rorippa, Arabidopsis and allied genera based on the internal transcribed spacer region of $18 \mathrm{~S}-25 \mathrm{~S}$ rDNA. Molec. Phylogenet. Evol. 13: 455-462.

Note added in proof.
The genus Beringia was renamed in the following article.
Al-Shehbaz, I. A. \& S. L. O'Kane. 2003. Transberingia, a new generic name replacing the illegitimate Beringia (Brassicaceae). Novon 13: 396.


[^0]:    ${ }^{1}$ Research and fieldwork were supported by grants from the National Science Foundation (DEB-9208433), the National Geographic Society (NGS-5068-93), and the University of Northern Iowa. Special thanks to Barbara Schaal, in whose lab this work was initiated, and to the support of the Missouri Botanical Garden. We offer gratitude to the many hosts, field companions, and herbarium curators who aided in this study, especially Abdulla Abbas, Nogman Aralbaev, Isa O. Baitulin, Alexandra Berkutenko, Ram Chaudhary, Gheorghe Dihoru, Vladimir Dorofeyev, Yang Guang, Sun Hang, Josef Holub, H. Kato, the late Sigizmund Kharkevich, Franta Krahulec, Hanna Kuciel, Karol Marhold, Zbigniew Mirek, Klaus Mummenhoff, Noriaki Murakami, Nonna Pavlova, and Boris Syomkin. Reviews by Donovan Bailey and an anonymous reviewer improved this paper substantially.
    ${ }^{2}$ Department of Biology, University of Northern Iowa, Cedar Falls, Iowa 50614-0421, U.S.A. steve.okane@uni.edu.
    ${ }^{3}$ Missouri Botanical Garden, P.O. Box 299, St. Louis, Missouri 63166-0299, U.S.A. ihsan.al-shehbaz@mobot.org.

