American Journal of Botany 88(10): 1881–1887. 2001.

SPECIES RELATIONSHIPS IN *LACTUCA* S.L. (LACTUCEAE, ASTERACEAE) INFERRED FROM AFLP FINGERPRINTS[1]

WIM J. M. KOOPMAN,[2,3] MARTIN J. ZEVENBERGEN,[2,4] AND RONALD G. VAN DEN BERG[2]

[2]Biosystematics Group, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, The Netherlands

An AFLP data set comprising 95 accessions from 20 species of *Lactuca* s.l. (sensu lato) and related genera was generated using the primer combinations E35/M48 and E35/M49. In phenetic analyses of a data subset, clustering with UPGMA based on Jaccard's similarity coefficient resulted in the highest cophenetic correlation, and the results were comparable to those of a principal coordinates analysis. In analyses of the total data set, phenetic and cladistic analyses showed similar tree topologies for the well-supported parts of the trees. The validity of cladistic analysis of AFLP data is discussed. The results do not support a distinction among the *serriola*-like species *L. sativa, L. serriola, L. dregeana,* and *L. altaica,* which is in line with previous results. Therefore, we postulate that these species are conspecific. The *serriola*-like species *L. aculeata* occupies a clearly separate position, making it an ideal outgroup for studies of the closest relatives of *L. sativa.* The subsect. *Lactuca* as a group is well supported by our data, but the positions of *L. saligna* and *L. virosa* relative to the *serriola*-like species remain unclear. The close relationship between the sect. *Mulgedium* species *L. tatarica* and *L. sibirica* is corroborated by the present AFLP results and by additional crossability data.

Key words: AFLPs; Asteraceae; Compositae; *Lactuca*; Lactuceae; lettuce; molecular phylogeny; phenetic relationships.

Cultivated lettuce (*Lactuca sativa* L.) is the world's most important leafy salad vegetable (McGuire et al., 1993). The taxonomic status of this species, the boundaries among *L. sativa* and close relatives, and the boundaries of the genus *Lactuca* L. s.l. (sensu lato) (Lactuceae, Asteraceae) itself have been the subject of controversy among taxonomists for many decades. One of the most widely used classifications today is that of Feráková (1977), comprising the European species of *Lactuca.* She subdivides the genus into four sections: *Lactuca, Mulgedium* (Cass.) C.B. Clarke, *Lactucopsis* (Schultz-Bip. ex Vis. et Panc.) Rouy., and *Phaenixopus* (Cass.) Benth. Section *Lactuca* is subdivided into the subsections *Lactuca* and *Cyanicae* DC. The subsection *Lactuca* comprises *L. sativa, L. serriola* L., *L. altaica* Fisch. et C.A. Mey., *L. saligna* L., *L. virosa* L., and *L. livida* Boiss. et Reut. *Lactuca livida* is closely related to *L. virosa* (Velasco Negueruela, 1981). The species *L. sativa, L. serriola,* and *L. altaica* are closely related and probably conspecific (see Koopman et al., 1998, for a discussion). The lesser known southwest Asian species *L. aculeata* Boiss. & Kotschy ex Boiss., *L. scarioloides* Boiss., *L. azerbaijanica* Rech., *L. georgica* Grossh., and the South-African species *L. dregeana* DC. are also closely related to *L. sativa*/*serriola*/*altaica* (Zohary, 1991). These species could all be classified in Feráková's subsection *Lactuca* if her classification were to be extended to include non-European species. The species of subsect. *Lactuca* comprise the readily accessible part of the lettuce gene pool and form potentially valuable gene sources for lettuce breeding (Zohary, 1991). *Lactuca serriola, L. saligna, L. virosa,* and to a lesser extent *L. altaica* are already commonly used as lettuce genitors. The *Lactuca* species outside subsect. *Lactuca,* as well as species from genera closely related to *Lactuca,* are interesting candidates for broadening the lettuce gene pool (Koopman et al., 1998).

In a previous study, Koopman et al. (1998) used ITS-1 (internal transcribed spacer-1) sequences to examine the relationships of species within or close to the lettuce gene pool. The study enabled straightforward conclusions on the generic and infrageneric boundaries of *Lactuca,* but was inconclusive as to the relationships among closely related species, e.g., within subsect. *Lactuca.* Koopman et al. (1998) concluded that additional information from a more variable marker was needed to resolve these relationships. A study by Hill et al. (1996) demonstrated that AFLPs (Vos et al., 1995) are variable markers useful for studying relationships among closely related species of *Lactuca.* Therefore, in the present study we used AFLP markers to further elucidate the relationships among *Lactuca* species and species from related genera. Our study had four foci: (1) the distinction between *L. sativa* and *L. serriola,* (2) the distinction between *L. serriola* and the *serriola*-like species *L. dregeana, L. altaica,* and *L. aculeata,* (3) the position of *L. saligna* and *L. virosa* relative to these *serriola*-like species, and (4) the detection of clusters/clades of closely related species outside subsect. *Lactuca.*

Data were analyzed both phenetically and cladistically, and the validity of cladistic analysis of AFLP data was discussed. A subset of data was used to compare various combinations of similarity coefficients and clustering methods for phenetic analyses.

MATERIALS AND METHODS

Plant material—We used 95 accessions from a previous ITS-1 sequence study (Koopman et al., 1998), representing 20 species of *Lactuca* and related genera. The species are listed in Table 1 according to the subtribal classification of Bremer (1994) and the generic and specific classification of Feráková (1977) and Iwatsuki et al. (1995). The choice of species, the major generic

[1] Manuscript received 28 November 2000; revision accepted 9 March 2001.

The authors thank Rik van Wijk of Keygene N.V. for suggesting the primer combinations for a pilot study and supplying a reference gel; the researchers of the Laboratory of Plant Breeding, Wageningen University, The Netherlands, for making their laboratory space and equipment available to us; and Prof. Dr. Ir. L.J.G. van der Maesen, Dr. F.T. Bakker, and N. Groendijk-Wilders of the Biosystematics Group, and Dr. J.H. de Jong of the Laboratory of Genetics for useful comments on the manuscript. This work was supported in part by a grant from Enza Zaden B.V., Leen de Mos B.V., Nickerson-Zwaan B.V., Nunhems Zaden B.V., Rijk Zwaan B.V., Novartis Seeds B.V., and Seminis Vegetable Seeds.

[3] Author for reprint requests (e-mail: wim.koopman@biosysnhn.dpw.wag-ur.nl).

[4] Current address: Keygene N.V., P.O. Box 216, 6700 AE Wageningen, The Netherlands.

TABLE 1. Lactuceae species used in this study. The subtribal classification follows Bremer (1994); generic and specific classification of European species follows Feráková (1977). The Asiatic species *L. indica* is classified in the non-European section *Tuberosae*, according to Iwatsuki et al. (1995).

Species	No. of accessions
Subtribe Lactucinae Dumort., genus *Lactuca* L.	
Lactuca sect. *Lactuca* subsect. *Lactuca*	
Lactuca sativa L.	10
Lactuca serriola L.	10
Lactuca dregeana DC.	2
Lactuca altaica Fisch. et C.A. Mey.	2
Lactuca aculeata Boiss. & Kotschy ex Boiss.	2
Lactuca saligna L.	10
Lactuca virosa L.	11
Lactuca sect. *Lactuca* subsect. *Cyanicae* DC.	
Lactuca tenerrima Pourr.	5
Lactuca perennis L.	5
Lactuca sect. *Mulgedium* (Cass.) C.B. Clarke	
Lactuca tatarica (L.) C.A. Mey.	6
Lactuca sibirica (L.) Benth. ex Marim.	5
Lactuca sect. *Lactucopsis* (Schultz-Bip. ex Vis. et Panc.) Rouy.	
Lactuca quercina L.	1
Lactuca sect. *Phaenixopus* (Cass.) Benth.	
Lactuca viminea (L.) J. & C. Presl	5
Lactuca sect. *Tuberosae* Boiss.	
Lactuca indica L.	5
Subtribe Lactucinae, other genera	
Mycelis muralis (L.) Dumort.	4
Steptorhamphus tuberosus (Jacq.) Grossh.	1
Cicerbita plumieri (L.) Kirschl.	3
Cicerbita alpina (L.) Wallr.	3
Prenanthes purpurea L.	2
Unassigned to a subtribe	
Cichorium intybus L.	3

concepts in *Lactuca* and related genera, and the delimitation of *Lactuca* were discussed in Koopman et al. (1998). Details on the accessions were given in Koopman et al. (1998) and on the website of the Centre for Genetic Resources, The Netherlands (CGN) at http://www.plant.wageningen-ur.nl/CGN. Each accession was represented by two plants. Voucher specimens of the plant material in rosette, bolting, and flowering stages were deposited at the Herbarium Vadense (WAG), supplemented with photographs of the plants in all three stages and with pappus preparations and fruit samples. All plants were grown under standard greenhouse conditions.

DNA extraction—Fresh young leaf tissue was collected from each plant, frozen in liquid nitrogen and kept at −70°C. Nuclei were isolated (one plant per accession), and DNA was further purified using phenol/chloroform extraction as described by Vosman et al. (1992).

AFLP analysis—The AFLP procedure was performed according to Van Eck et al. (1995) with minor modifications. In the restriction/ligation reaction ~250 ng of genomic DNA was digested for 1 h at 37°C using 2.5 U (units) EcoRI, 2.5 U MseI, and 8 μL 5× restriction-ligation buffer (5× RL buffer) in a total volume of 40 μL. Restriction/ligation was continued for another 3 h after addition of 10 μL of ligation mixture (containing 5 pmol EcoRI adapter, 50 pmol MseI adapter, 1.0 μL 10 mmol/L ATP, 2.0 μL 5× RL buffer, and 1.0 U T4 DNA ligase). The subsequent selection of biotinylated restriction fragments with streptavidin-coated Dynabeads was replaced by a tenfold dilution of the restriction/ligation mixture with distilled water.

Preselective amplification was performed using the primers E01 (EcoRI + A) and M02 (MseI + C). The resulting product was diluted 50-fold with T0.1E buffer (10 mmol/L Tris pH 8.0, 0.1 mmol/L EDTA). The final restriction fragment amplification was performed using primers with three selective nucleotides. The EcoRI primer in this reaction was labeled with ^{33}P. A pilot study was conducted to test ten primer combinations: E33/M59, E35/M48, E35/M49, E35/M59, E35/M60, E38/M54, E44/M48, E44/M49, E45/M48, and E45/M49. The test data set contained four plants of *L. sativa*, two plants of *L. saligna*, and one plant from each of the other species in Table 1. The AFLP fragments for this experiment were separated on a 0.35-mm sequence system (Gibco BRL/Life Technologies, Rockville, Maryland, USA) and visualized on Kodak X-OMAT LS Scientific Imaging Film (Eastman Kodak, Rochester, New York, USA). Selection of primer combinations was based on the number of bands per lane, the number of bands that were constant among the species, and the absence of very fat bands or smears. Primer combinations E35/M48 (EcoRI + ACA/MseI + CAC) and E35/M49 (EcoRI + ACA/MseI + CAG) were selected to generate the final data set. The AFLP procedure for E35/M48 was performed as above. For E35/M49, final restriction fragment amplification and separation and visualization of the AFLP fragments was performed according to Arens et al. (1998). *Lactuca sativa* 'Norden' served as size standard on each gel. A reference gel with fragment lengths of 'Norden' was kindly provided by Keygene N.V. (Wageningen, The Netherlands).

Data analysis—AFLP fragments were scored as present/absent. Fragment scoring and lane matching were performed automatically on digital images of the autoradiograms, using Phoretix 1D advanced Version 4.00 (Phoretix International, Newcastle upon Tyne, UK). All but the faintest bands were scored, where necessary scores and matches were corrected manually. Fragments scored ranged from 112 to 453 nucleotides for E35/M48 and from 111 to 502 nucleotides for E35/M49. Data from both primer combinations were combined in one data set. The data set was analyzed in two steps.

Firstly, a data subset was constructed comprising *L. sativa* and its closest relatives, *L. serriola*, *L. dregeana*, *L. altaica*, and *L. aculeata*. In the following these will be referred to as the "*serriola*-like species." The subset was used to compare various similarity coefficients and clustering methods and to study the relationships among the *serriola*-like species in detail. Clustering methods and similarity coefficients were tested using the procedures SIMQUAL, SAHN, and TREE from the program NTSYSpc version 2.02k (Applied Biostatistics, Setauket, New York, USA). The "TM" option was set to "FIND" to enable detection of all possible trees. The clustering methods UPGMA, WPGMA, Complete-link, and Single-link were applied in all possible combinations with the similarity coefficients Dice, Jaccard, and Simple matching. Clustering methods and similarity coefficients are described in Rohlf (1993). Cophenetic correlation coefficients (r) were calculated and compared for each of the combinations using the procedures COPH and MXCOMP from NTSYSpc 2.02k. These coefficients indicate the correlation between a similarity matrix and the phenetic tree resulting from it after a cluster analysis, and thus are a measure for the goodness of fit of the cluster analysis to the similarity matrix.

Species relationships among the *serriola*-like species were studied using a principal coordinates analysis (PCO). Jaccard's similarity coefficient and the procedures DCENTER, EIGEN, and MXPLOT from NTSYSpc 2.02k were used to perform the PCO.

Secondly, analyses were performed on the entire data set, containing all accessions from Table 1. This data set was used to compare phenetic and cladistic analysis of the AFLP data and to detect well-supported species clusters/clades within *Lactuca* s.l. The cluster analysis was performed with TREECON 1.2 (Van de Peer and De Wachter, 1994), which enabled bootstrapping of the resulting phenogram. Nei and Li's (1979) dissimilarity coefficient and UPGMA clustering were used; bootstrap values were calculated in 1000 replications. Cladistic analyses and determination of phylogenetic signal in the data set were performed using PAUP version 4.0a (Swofford, 1999). Parsimony settings were: ACCTRAN and "collapse of zero length branches" (max). Phylogenetic signal was determined from the tree-length distribution of 100 000 trees, using the g_1-statistic (Hillis and Huelsenbeck, 1992). The lettuce data set contained >25 taxa and >500 variable characters, and therefore the critical value of −0.08 was used. A g_1-statistic lower than this critical value indicates the presence of significant phylogenetic signal in the corresponding data set (Hillis and Huelsenbeck, 1992).

The cladistic analyses started as a jackknife analysis using 10 000 replicates of a fast heuristic search, nominal deletion of 37% of the characters, and

TABLE 2. Cophenetic correlation coefficients for a data subset containing *Lactuca sativa, L. serriola, L. dregeana, L. altaica,* and *L. aculeata* accessions; total number of trees found in the analysis are in brackets. When multiple trees were found, only the highest cophenetic value is shown. For the Single-link method, all trees are equivalent and thus have the same cophenetic values.

Clustering/similarity	Dice	Jaccard	Simple matching
UPGMA	0.974 (1)	0.979 (1)	0.955 (1)
WPGMA	0.963 (1)	0.968 (1)	0.953 (1)
Complete-link	0.969 (2)	0.973 (2)	0.948 (2)
Single-link	0.951 (1)	0.957 (1)	0.941 (8)

''Jac'' resampling. A 50% majority rule consensus tree was calculated based on the jackknife analysis and used as a constraint tree for a heuristic search. The heuristic search comprised 10 000 random-addition sequences and tree bisection-reconnection (TBR) branch swapping with ''multrees'' switched off. A second search was performed using four cycles of successive weighting. The strict consensus of the heuristic search above was used as a starting point. Characters were reweighted by the maximum value of the rescaled consistency indices, and the searches were conducted with 100 random-addition sequences, tree-bisection-reconnection (TBR) branch swapping, and ''multrees'' on. Jackknife values for the resulting tree were calculated as above.

RESULTS

Total number of bands scored was 544 for E35/M48 and 521 for E35/M49, all of which were polymorphic. Band numbers for the individual accessions ranged from 16 to 109 (average 59.0 bands/lane) for E35/M48 and from 28 to 103 for E35/M49 (average 54.6 bands/lane).

Table 2 shows the cophenetic correlation coefficients from analyses of the data subset containing *L. sativa* and its closest relatives, *L. serriola, L. dregeana, L. altaica,* and *L. aculeata* (the *serriola*-like species). UPGMA clustering yielded the highest cophenetic correlation in all cases, Single-link the lowest. The ranking of WPGMA and Complete-link was less consistent (see Table 2). Among the similarity coefficients, Jaccard consistently yielded the highest cophenetic correlations, followed by Dice and Simple matching. Since a similar ranking of similarity coefficients and clustering methods was found by Mace, Lester, and Gebhardt (1999) and Mace, Gebhardt, and Lester (1999) for *Solanum* and *Datura/Brugmansia,* it possibly applies to all AFLP data sets. The combination of Jaccard similarity with UPGMA clustering yielded the highest cophenetic correlation and is therefore considered most suitable for determining phenetic species relationships in *Lactuca* s.l. The combination of UPGMA with the Dice or Nei and Li (equaling 1-Dice) coefficient is also suitable for our data set, since trees based on these coefficients were identical to those based on Jaccard's coefficient. Identical topologies for the Jaccard and Dice coefficients were also found by Milbourne et al. (1997) for AFLP data from cultivated potato.

Species relationships among the *serriola*-like species were studied in detail with a PCO (Fig. 1). The first principal coordinate describes 18% of the total variation and separates three groups. Group 1 contains the *L. altaica* accessions and some of the *L. serriola* accessions. The *L. serriola* accessions CGN 15684 and CGN 5900 also fall in this group, although they cluster in group 2a in the cluster analysis (see below). The *L. altaica* accessions fall among the *L. serriola* accessions. Group 2a contains *L. sativa, L. serriola,* and *L. dregeana.* The *L. dregeana* accessions fall among the *L. serriola* accessions. Group 2b contains most *L. sativa* accessions and the *L. serriola* oilseed accessions. Note that the *L. sativa* accessions also include an oilseed accession, CGN 9356. The second principal coordinate describes 12% of the total variation and clearly sets

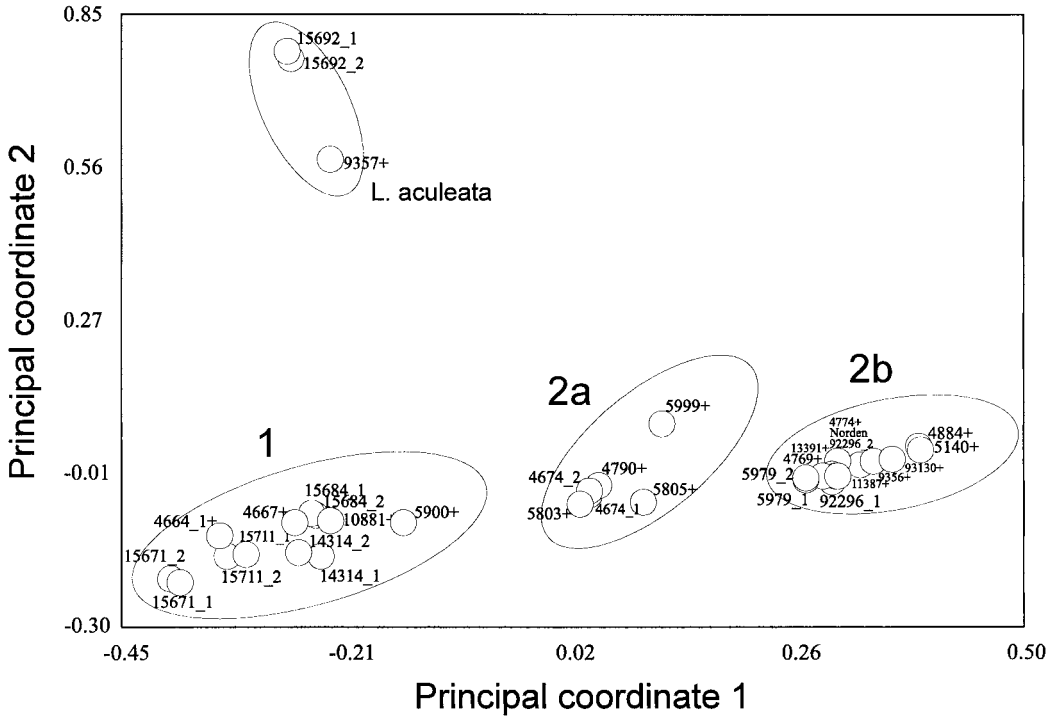

Fig. 1. Principal coordinates analysis of a data subset. Numbers 1, 2a, and 2b indicate the different groups of species referred to in the text. 1 = *Lactuca serriola/L. altaica,* 2a = *L. sativa/L. serriola/L. dregeana,* 2b = *L. sativa/L. serriola.*

apart *L. aculeata* from *L. sativa, L. serriola, L. dregeana,* and *L. altaica.*

In the cluster analysis comprising all accessions, all species except *L. sativa* and *L. serriola* have their own distinct branches (Fig. 2a). However, *L. altaica* (30% support) and *L. dregeana* (98%) cluster within *L. sativa*/*L. serriola.* Subsection *Lactuca* (the *serriola*-like species together with *L. virosa* and *L. saligna*) is well supported (99%). *Lactuca virosa* clusters more closely to the *serriola*-like species than does *L. saligna,* but the branch determining this order is poorly supported (52%). The cluster including only the *serriola*-like species is well supported (100%) and consists of four groups. These groups are identical to those in the PCO, except for the position of CGN 15684 and CGN 5900 (see above). The cluster with *L. aculeata* is strongly supported (100%), but the *L. serriola*/*L. altaica* cluster (group 1) and the *L. sativa/L. serriola/ L. dregeana* cluster (group 2a), are not (26% and 18%, respectively). The cluster with only *L. sativa* accessions (group 2b) is strongly supported (96%), but falls entirely within the *L. sativa/L. serriola/L. dregeana* cluster. The only well-supported species cluster outside subsect. *Lactuca* is that of *L. sibirica, L. tatarica,* and *L. quercina* (99%).

In the cladistic analyses, the g_1-statistic for the combined data set was -0.39, indicating significant phylogenetic signal. The heuristic search with random-addition sequences yielded 40 shortest trees of 4628 steps (retention index [RI] = 0.76, consistency index [CI] = 0.23, rescaled consistency index [RC] = 0.18). The search with successive weighting yielded a single tree of 752 steps (RI = 0.84, CI = 0.45, RC = 0.38). This single tree was compatible to the strict consensus tree of the search with random-addition sequences, but slightly more resolved (Fig. 2b). Topology and bootstrap/jackknife supports for clusters/clades with a support >70% are similar in the cladogram (Fig. 2b) and the phenogram (Fig. 2a).

DISCUSSION

Distinction between L. sativa and L. serriola—*Lactuca sativa* and *L. serriola* group in three clusters/clades, but the distinction between these clusters/clades is weakly supported. Two of the clusters/clades contain both *L. sativa* and *L. serriola* accessions. This is consistent with the AFLP results of Hill et al. (1996), showing a large *L. sativa*/*L. serriola* cluster with *L. serriola* accessions branching off basally to a large subcluster containing all *L. sativa* accessions and one *L. serriola* accession. This *L. serriola* is a ''landrace type,'' intermediate between *L. sativa* and *L. serriola.* The *L. sativa*/*L. serriola* cluster is clearly separated from *L. saligna, L. virosa, L. indica,* and *L. perennis.*

The *L. serriola* and *L. sativa* oilseed accessions in the present study fell within *L. sativa.* This is consistent with the results of Frietema de Vries, Van der Meijden, and Brandenburg (1994) and Frietema de Vries (1996). In their principal components analysis of morphological data, the *L. sativa* and *L. serriola* accessions fall in two partly overlapping groups. According to the text, the oilseed accessions are included in the *L. sativa* cluster, although an accompanying figure depicts them as intermediate between *L. sativa* and *L. serriola* (Frietema de Vries, Van der Meijden, and Brandenburg, 1994; Frietema de Vries, 1996).

Given the lack of distinction between *L. sativa* and *L. serriola* in the present study, the position of the *L. serriola* oilseeds within *L. sativa* in the study of Frietema de Vries, Van der Meijden, and Brandenburg (1994) and Frietema de Vries (1996), the presence of a *L. serriola* ''landrace type'' within *L. sativa* in the study of Hill et al. (1996), and the close similarity of *L. sativa* and *L. serriola* in other characters (discussed in Koopman et al., 1998), we support the conclusion of Frietema de Vries, Van der Meijden, and Brandenburg (1994) and Frietema de Vries (1996) that *L. sativa* and *L. serriola* are conspecific. However, we do not support the distinction of *L. sativa* subsp. *sativa* and *L. sativa* subsp. *serriola,* as proposed by Frietema de Vries (1996). In our opinion, the species are too similar even to maintain them as subspecies. Therefore, we consider the earliest name, *L. sativa,* the correct name for both *L. serriola* and *L. sativa.*

Position of L. dregeana, L. altaica, and L. aculeata relative to L. serriola—*Lactuca dregeana* accessions fell within the mixed *L. sativa*/*L. serriola* cluster/clade in all our analyses. Most accessions within this cluster/clade show a mixture of *L. serriola* and *L. sativa* characteristics. For example, *L. sativa* accession CGN 5999 has an especially rigid, nearly woody stem, and spines on the midribs beneath, characteristics usually associated with *L. serriola.* On the other hand, *L. serriola* accessions CGN 5803 and CGN 4674 show spineless lower midribs, somewhat fleshy leaves, and involucres that are not completely reflexed when the fruits are ripe. These characteristics are usually associated with *L. sativa.* The *L. dregeana* accessions show a similar combination of characteristics. They resemble *L. sativa* in their somewhat fleshy leaves and involucres that are not completely reflexed when the achenes are ripe. On the other hand, they show *L. serriola* characteristics such as a rigid, spiny stem, spiny lower midribs, and dark brown, spotted achenes.

The combination of morphological characteristics and the position of *L. dregeana* in the mixed *sativa*/*serriola* cluster/ clade in our AFLP analyses suggest that *L. dregeana* escaped from cultivation. The fact that *L. dregeana* is endemic to South Africa could mean that it originated from the primitive lettuce cultivars introduced there by European settlers in the 17th century. Lettuce seed production in the Cape was reported as early as 1652–1654 (Karsten, 1951) and could easily have led to escapes to the wild by wind dispersal of achenes from cultivars with loose involucres. After taking into account the morphology of *L. dregeana,* its position in the AFLP analyses, and its possible origin in cultivated lettuce, *L. dregeana* probably does not deserve a species status, but it should be regarded conspecific with *L. sativa*/*L. serriola.*

The *L. altaica* accessions in the present study fell within a group of *L. serriola* accessions, corroborating previous ITS-1 results and the conclusion that *L. altaica* is probably conspecific with *L. serriola* (Koopman et al., 1998). However, this conclusion is based on only two *L. altaica* accessions. Recently, additional wild material of *L. altaica* and its relatives *L. serriola* and *L. saligna* was collected in Uzbekistan (Van Soest, 1997). A study on this material is currently being carried out to further elucidate the relationships and taxonomic status of *L. altaica.*

The accessions of *L. aculeata* form a clearly distinct group among the *serriola*-like species, with a 100% jackknife and bootstrap support. The position of *L. aculeata* separate from, yet closely related to, the other *serriola*-like species is well supported by our earlier ITS-1 study (Koopman et al., 1998). This distinct position of *L. aculeata* within the *serriola*-like

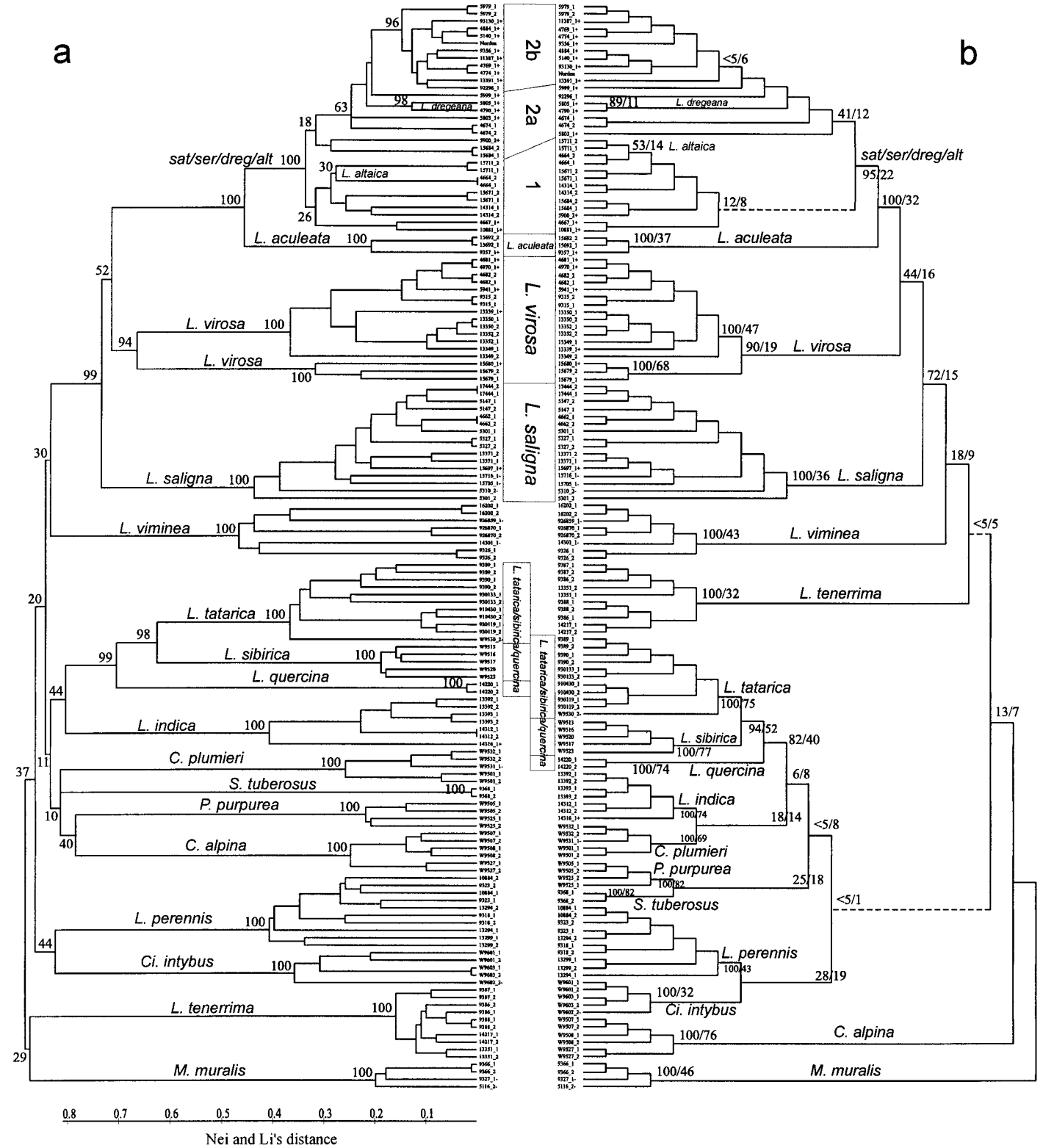

Fig. 2. (a) UPGMA phenogram based on Nei and Li's distance. Numbers on branches are bootstrap values. (b) Cladogram resulting from successive weighting of the strict consensus tree from 10 000 random-addition sequences with TBR branch swapping and ''multrees'' switched off. Numbers on each branch are jackknife value (left of slash) and total number of AFLP bands supporting the branch (unweighted; right of slash). Dotted branches collapse in the strict consensus of the trees from the random-addition sequence searches. Two plants were used for each accession, indicated by _1 and _2. A plus sign indicates that the AFLP patterns were identical for the two plants, and only one of the plants is depicted in the phenogram (*sat/ser/dreg/alt* = species cluster containing *L. sativa, L. serriola, L. dregeana,* and *L. altaica*). The boxes between (a) and (b) indicate well-supported clusters present in both trees. Numbers 1, 2a, and 2b indicate different groups of species referred to in the text. 1 = *L. serriola/L. altaica,* 2a = *L. sativa/L. serriola/L. dregeana,* 2b = *L. sativa/L. serriola.*

species makes it an ideal outgroup for studies into *L. sativa, L. serriola,* and their closest relatives.

All *serriola*-like species together, i.e., including *L. aculeata,* form a homogeneous group of closely related species within subsect. *Lactuca.* This is indicated by the 100% jackknife and bootstrap support for this group in the present AFLP analysis (Fig. 2a,b), the 95% bootstrap support in a previous ITS-1 analysis, and the fact that all *serriola*-like species are fully interfertile (Koopman et al., 1998).

Position of L. saligna and L. virosa within subsect. Lactuca—The results of previous studies on plant morphology (De Vries and Van Raamsdonk, 1994), crossability (Thompson, Whitaker, and Kosar, 1941; Lindqvist, 1960; De Vries, 1990), SDS (sodium dodecyl sulphate) electrophoresis patterns of seed proteins (De Vries, 1996), isozyme analysis of foliar esterases (Roux, Chengjiu, and Roux, 1985), karyotype (Lindqvist, 1960; Koopman and De Jong, 1996), chromosome banding pattern (Koopman, De Jong, and De Vries, 1993), DNA content (Koopman and De Jong, 1996; Koopman, 2000), nuclear RFLPs (restriction fragment length polymorphisms) (Kesseli, Ochoa, and Michelmore, 1991), mtDNA RFLPs (Vermeulen et al., 1994), nuclear AFLPs (Hill et al., 1996), and ITS-1 sequences (Koopman et al., 1998) showed different possibilities for the position of *L. saligna* and *L. virosa* relative to the *serriola*-like species, as was discussed in Koopman et al. (1998).

The present results indicate that *L. virosa* is more closely related to the *serriola*-like species than is *L. saligna.* However, the position of *L. saligna* and *L. virosa* relative to the *serriola*-like species is not very reliable, as is indicated by the low bootstrap and jackknife supports on the branches separating *L. saligna* and *L. virosa.* The results are not consistent with the AFLP analysis of Hill et al. (1996), indicating that *L. saligna* is the closest relative of the *serriola*-like species and that *L. perennis* is even more closely related to the *serriola*-like species than is *L. virosa.* However, Hill et al. (1996) do not indicate support values for the relationships. Given the different positions of *L. saligna* and *L. virosa* in the study of Hill et al. (1996) and in the present study, and the lack of branch support for these positions, we conclude that the available AFLP data are inconclusive as to the position of *L. virosa* and *L. saligna* relative to the *serriola*-like species. This is also true for the position of *L. perennis.*

In the present study, *L. virosa* accessions CGN 15679 and 15680 form a separate clade with a 100% bootstrap support. The anomalous position of these accessions may indicate that they are a distinct infraspecific taxon within *L. virosa.* The distinct position of CGN 15679 and CGN 15680 is also reflected by the fact that their DNA content is 1.16 times that of the other *L. virosa* accessions (Koopman, 2000).

Species clusters/clades outside subsect. Lactuca—In a previous study using ITS-1 sequences, four clades of species were detected outside subsect. *Lactuca*: (1) *L. tatarica*/*L. sibirica*/*L. viminea,* (2) *L. perennis*/*C. plumieri,* (3) *L. tenerrima*/*S. tuberosus,* (4) *M. muralis*/*C. alpina.* Only one of these clades could be partially confirmed by our AFLP results. In all phenetic and cladistic analyses, the only well-supported cluster/clade outside subsect. *Lactuca* was one with *L. tatarica, L. sibirica,* and *L. quercina.* The relationship between *L. tatarica* and *L. sibirica* is consistent with the ITS-1 results and with the classification of Feráková (1977). The close relationship of *L. tatarica* and *L. sibirica* with *L. quercina* is not. In the ITS-1 phylogeny, *L. quercina* has its own distinct branch, while *L. tatarica* and *L. sibirica* occupy the same clade. In the classification of Feráková (1977), *L. quercina* is classified in section *Lactucopsis,* while *L. tatarica* and *L. sibirica* together make up section *Mulgedium.* The close relationship between *L. tatarica* and *L. sibirica* was confirmed by our recent crossing experiments, reported here for the first time. We conducted reciprocal crosses between four *L. tatarica* accessions and four *L. sibirica* accessions, and these crosses yielded viable seeds for six out of eight combinations of accessions. The F1 plants were vigorous and fully fertile, indicating a close genetic relationship between *L. tatarica* and *L. sibirica.* The possible close relationship of *L. tatarica*/*L. sibirica* with *L. quercina* needs more verification. Species relationships involving *L. viminea, L. indica, L. perennis, L. tenerrima, M. muralis, C. plumieri, C. alpina, S. tuberosus, P. purpurea,* and *C. intybus* could not be assessed in the present study, because the AFLPs were too variable to determine reliable relationships of these species.

Methodological considerations—In the present study, the data were analyzed both phenetically and cladistically. The validity of such analyses is sometimes disputed, although this dispute was not reflected in literature until now. Critics recognize two main sources of error in the cladistic analysis of AFLP data. Firstly, the fact that AFLPs are anonymous markers is a source of error. Because AFLP fragments are identified by their length and not by their base composition, nonidentical fragments of equal length will mistakenly be scored as identical. Secondly, the fact that AFLPs are scored dominantly is a source of error. AFLPs are usually scored as dominant characters, i.e., with only the character states present (1), and absent (0). In reality, at least part of the bands may represent codominant markers that have three character states, namely 0/0, 1/0, and 1/1. Both sources of error introduce homoplasies in the data set, possibly leading to erroneous tree topologies in cladistic analyses. In our opinion, the impact of these homoplasies on the conclusions regarding species relationships will be minor.

When we compare phenetic and cladistic analysis of AFLP data, there are two possible situations. Firstly, the topologies of the phenogram and the cladogram may be identical. In this case, the homoplasies were too minor to influence the topology of the cladogram. Consequently, they will not affect conclusions on species relationships. Secondly, the topologies of the phenogram and the cladogram may be different. In this case, the homoplasies significantly affected the topology of branches in the cladogram. However, because the differences are caused by homoplasies, there will be internal conflict in the data defining these branches. In branch support analyses such as bootstrapping or jackknifing, the presence of such conflicting data gives rise to low support values. These poorly supported branches will be discarded as uninformative when conclusions on species relationships are drawn. Therefore, in this case, too, the homoplasies in the AFLP data will not affect the conclusions on species relationships. In both cases, cladistic analysis of AFLP data will give rise to reliable phylogenetic conclusions, notwithstanding the validity of the theoretical objections.

The first case is illustrated by a study of Kardolus, Van Eck, and Van den Berg (1998) in which a cladogram and a phenogram of 16 wild *Solanum* species show highly similar topologies, even for moderately supported groups. In our lettuce study both cases are present: the well-supported parts of the phenogram and the cladogram show similar topologies, while

the differences in the remaining parts of the trees are poorly supported.

Conclusions—AFLPs proved to be suitable molecular markers to study the relationships among closely related species of *Lactuca* s.l. In phenetic analyses of a data subset, the combination of Jaccard's similarity coefficient with UPGMA clustering resulted in the highest cophenetic value. The results of a principal coordinates analysis of the subset were comparable to those of the UPGMA analysis. A data set comprising all accessions was analyzed phenetically as well as cladistically, and the well-supported parts of the trees were comparable for both types of analyses. The AFLP results corroborated the conclusions from a previous ITS-1 sequence study (Koopman et al., 1998) that the *serriola*-like species *L. sativa, L. serriola, L. dregeana,* and *L. altaica* cannot be reliably distinguished and are probably conspecific. *Lactuca dregeana* possibly escaped from cultivation. *Lactuca aculeata* is closely related to the other *serriola*-like species, but clearly different. The AFLP results were inconclusive as to the position of *L. saligna* and *L. virosa* relative to the *serriola*-like species, but the status of subsect. *Lactuca* (the *serriola*-like species together with *L. saligna* and *L. virosa*) as a recognizable group within *Lactuca* s.l. was supported in all analyses. In the previous ITS-1 study, a number of species clades outside subsect. *Lactuca* were identified. Among the relationships indicated by these clades, only the close relationship between *L. tatarica* and *L. sibirica* (together constituting *Lactuca* subsect. Mulgedium) was corroborated by the present AFLP results. The close relationship between these species was also corroborated by our crossability data.

LITERATURE CITED

ARENS, P., H. COOPS, J. JANSEN, AND B. VOSMAN. 1998. Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. *Molecular Ecology* 7: 11–18.

BREMER, K. 1994. Asteraceae: cladistics and classification. Timber Press, Portland, Oregon, USA.

DE VRIES, I. M. 1990. Crossing experiments of lettuce cultivars and species (*Lactuca sect. Lactuca,* Compositae). *Plant Systematics and Evolution* 171: 233–248.

DE VRIES, I. M. 1996. Characterization and identification of *Lactuca sativa* cultivars and wild relatives with SDS-electrophoresis (*Lactuca* sect. *Lactuca, Compositae*). *Genetic Resources and Crop Evolution* 43: 193–202.

DE VRIES, I. M., AND L. W. D. VAN RAAMSDONK. 1994. Numerical morphological analysis of lettuce cultivars and species (*Lactuca sect. Lactuca,* Asteraceae). *Plant Systematics and Evolution* 193: 125–141.

FERÁKOVÁ, V. 1977. The genus *Lactuca* L. in Europe. Univerzita Komenského Press, Bratislava, Czechoslovakia.

FRIETEMA DE VRIES, F. T. 1996. Cultivated plants and the wild flora. Ph.D. dissertation, Rijksherbarium/Hortus Botanicus, Leiden, The Netherlands.

FRIETEMA DE VRIES, F. T., R. VAN DER MEIJDEN, AND W. A. BRANDENBURG. 1994. Botanical files on lettuce (*Lactuca sativa*). *Gorteria* supplement 2.

HILL, M., H. WITSENBOER, M. ZABEAU, P. VOS, R. KESSELI, AND R. MICHELMORE. 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. *Theoretical and Applied Genetics* 93: 1202–1210.

HILLIS, D. M., AND J. P. HUELSENBECK. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* 83: 189–195.

IWATSUKI, K., T. YAMAZAKI, D. E. BOUFFORD, AND H. OHBA. 1995. Flora of Japan, vol. IIIb. Kodansha, Tokyo, Japan.

KARDOLUS, J. P., H. J. VAN ECK, AND R. G. VAN DEN BERG. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (*Solanaceae*). *Plant Systematics and Evolution* 210: 87–103.

KARSTEN, M. C. 1951. The Old Company's Garden at the cape and its superintendents. Maskew Miller Limited, Cape Town, South Africa.

KESSELI, R. V., O. OCHOA, AND R. W. MICHELMORE. 1991. Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*L. sativa*). *Genome* 34: 430–436.

KOOPMAN, W. J. M. 2000. Identifying lettuce species (*Lactuca subsect. Lactuca,* Asteraceae): a practical application of flow cytometry. *Euphytica* 116: 151–159.

KOOPMAN, W. J. M., AND J. H. DE JONG. 1996. A numerical analysis of karyotypes and DNA amounts in lettuce cultivars and species (*Lactuca subsect. Lactuca,* Compositae). *Acta Botanica Neerlandica* 45: 211–222.

KOOPMAN, W. J. M., J. H. DE JONG, AND I. M. DE VRIES. 1993. Chromosome banding patterns in lettuce species (*Lactuca sect. Lactuca, Compositae*). *Plant Systematics and Evolution* 185: 249–257.

KOOPMAN, W. J. M., E. GUETTA, C. C. M. VAN DE WIEL, B. VOSMAN, AND R. G. VAN DEN BERG. 1998. Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences. *American Journal of Botany* 85: 1517–1530.

LINDQVIST, K. 1960. Cytogenetic studies in the *serriola* group of *Lactuca. Hereditas* 46: 75–151.

MACE, E. S., C. G. GEBHARDT, AND R. N. LESTER. 1999b. AFLP analysis of genetic relationships in the tribe Datureae (Solanaceae). *Theoretical and Applied Genetics* 99: 634–641.

MACE, E. S., R. N. LESTER, AND C. G. GEBHARDT. 1999a. AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). *Theoretical and Applied Genetics* 99: 626–633.

MCGUIRE, P. E., E. J. RYDER, R. W. MICHELMORE, R. L. CLARK, R. ANTLE, G. EMERY, R. M. HANNAN, R. V. KESSELI, E. A. KURTZ, O. OCHOA, V. E. RUBATZKY, AND W. WAYCOTT. 1993. Genetic resources of lettuce and *Lactuca* species in California: an assessment of the USDA and UC collections and recommendations for long-term security. Report No. 12, Genetic Resources Conservation Program, Division of Agriculture and Natural Resources, University of California, Davis, California, USA.

MILBOURNE, D., R. MEYER, J. E. BRADSHAW, E. BAIRD, N. BONAR, J. PROVAN, W. POWELL, AND R. WAUGH. 1997. Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. *Molecular Breeding* 3: 127–136.

NEI, M., AND W.-H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA* 76: 5269–5273.

ROHLF, F. J. 1993. NTSYS-pc: numerical taxonomy and multivariate analysis system, version 1.80. Applied Biostatistics, Setauket, New York, USA.

ROUX, L., Z. CHENGJIU, AND Y. ROUX. 1985. Caractérisation de *Lactuca sativa* L. et des espèces apparentées par électrofocalisation des estérases. *Agronomie* 5: 915–921.

SWOFFORD, D. L. 1999. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer, Sunderland, Massachusetts, USA.

THOMPSON, R. C., T. W. WHITAKER, AND W. F. KOSAR. 1941. Interspecific genetic relationships in *Lactuca. Journal of Agricultural Research* 63: 91–107.

VAN DE PEER, Y., AND R. DE WACHTER. 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Computer Applications in the Biosciences* 10: 569–570.

VAN ECK, H. J., J. ROUPPE VAN DER VOORT, J. DRAAISTRA, P. VAN ZANDVOORT, E. VAN ENCKEVORT, B. SEGERS, J. PELEMAN, E. JACOBSEN, J. HELDER, AND J. BAKKER. 1995. The inheritance and chromosomal localization of AFLP markers in a non-inbred potato offspring. *Molecular Breeding* 1: 397–410.

VAN SOEST, L. J. M. 1997. Report of expedition to Uzbekistan in 1997: itinerary, collected materials and data. Agricultural Research Department, Centre for Plant Breeding and Reproduction Research, Centre for Genetic Resources, Wageningen, The Netherlands.

VELASCO NEGUERUELA, A. 1981. Sobre *Lactuca livida* Boiss. & Reuter. *Anales del Jardín Botánico de Madrid* 38: 316.

VERMEULEN, A., B. DESPREZ, D. LANCELIN, AND H. BANNEROT. 1994. Relationships among *Cichorium* species and related genera as determined by analysis of mitochondrial RFLPs. *Theoretical and Applied Genetics* 88: 159–166.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.

VOSMAN, B., P. ARENS, W. RUS-KORTEKAAS, AND M. J. M. SMULDERS. 1992. Identification of highly polymorphic DNA regions in tomato. *Theoretical and Applied Genetics* 85: 239–244.

ZOHARY, D. 1991. The wild genetic resources of cultivated lettuce (*Lactuca sativa* L.). *Euphytica* 53: 31–35.