Theor Appl Genet (2000) 100:1000–1009

L.W.D. van Raamsdonk · M. Vrielink-van Ginkel
C. Kik

Phylogeny reconstruction and hybrid analysis in *Allium* subgenus *Rhizirideum*

Received: 23 August 1999 / Accepted: 14 September 1999

Abstract Amplified fragment length polymorphisms (AFLPs) for assessing nuclear DNA diversity have been used for the reconstruction of the phylogeny and evolution of several sections of *Allium* subgenus *Rhizirideum*. A dataset of 355 characters for 33 accessions belonging to 20 species has been compiled. The band-sharing of five interspecific hybrids and of an F_2 population between *Allium cepa* and *Allium roylei* with their parents indicated a heterozygosity level between 6 and 14%, which allows the use of dominant markers such as AFLPs for phylogeny reconstruction. A majority rule consensus tree based on 56 most-parsimonius trees (CI = 0.528) revealed a separate clade for each of the sections, *Cepa*, *Rhizirideum* and *Schoenoprasum*, and one clade combining the sections *Oreiprason* and *Petroprason*. An unweighted pair group mean average (UPGMA)-based dendrogram showed the same subdivision. The trees and the 'Hybrid Distance' approach both supported the assumption of a hybrid origin for *A. roylei* with considerable subsequent secondary evolution. The establishment of three alliances in the section *Cepa* and the close relationship of sections *Oreiprason* and *Petroprason* are now confirmed. The predictions of the Soybean domestication scenario, i.e. selection of a crop from one progenitor with subsequent narrowing of the genetic diversity of the crop, which applies to the cultigens *A. cepa* and *Allium fistulosum*, is supported by the Hybrid Distance approach.

Key words *Allium* subgenus *Rhizirideum* · *Allium cepa* · *Allium roylei* · AFLP · Onion · Phylogeny reconstruction · Hybrid Distance · nDNA

Communicated by R. Hagemann

L.W.D. van Raamsdonk (✉) · M. Vrielink-van Ginkel · C. Kik
DLO – Centre for Plant Breeding
and Reproduction Research (CPRO-DLO),
P.O. Box 16, 6700 AA Wageningen, the Netherlands
e-mail: l.w.d.vanraamsdonk@rikilt.wag-ur.nl
Fax: +31-317-417717

Introduction

Phylogenetic reconstruction of evolutionary processes using cladistic methods is theoretically not suitable for the analysis of hybrid speciation. The cladistic approach is exclusively based on a branching pattern of speciation (Stuessy 1990). At the same time, the phenetic approach only gives a spatial representation of the diversity under study after the calculation of mutual distances. The statement that, until recently, no specific methods were available for the analysis of hybrid speciation (McDade 1997; Kardolus et al. 1998) is remarkable in the view of the fact that, at the diploid level, hybridisation is of major importance (Rieseberg and Wendel 1993). As an indication, a total of 23,000 species combinations (excluding orchids) was estimated from a total of 250 000 plant species (van Damme 1992).

The lack of approaches for the study of hybrid speciation became apparent in the study of *Allium* subgenus *Rhizirideum* and especially from the phylogenetic position of *Allium roylei*. Notwithstanding the crossing capability of *A. roylei* with *Allium cepa*, *A. roylei* combines characters of *A. cepa*, and its relatives, with characters that are predominantly found in species of other sections especially the sections *Rhizirideum* and *Schoenoprasum*. This situation indicates that the taxonomic position of *A. roylei* close to *A. cepa* remains disputable (van Raamsdonk et al. 1997). The relatively large phylogenetic distance between *A. roylei* and *A. cepa*, combined with their reasonable crossability, allows the production of F_2 populations with a sufficient amount of diversity suitable for mapping studies (van Heusden et al. 2000). The unique situation of the crossability of *A. roylei* with either *A. cepa* or *Allium fistulosum* resulted in the production of bridge cross-hybrids that combine parts of all three genomes (Khrustaleva and Kik 1998, 2000).

A preliminary study aimed at unravelling the exact phylogenetic position of *A. roylei* resulted in the development of a hybrid distance measure for analysing hybrid speciation. This measure is based on the assumption that hybrids share the allelic profiles of both parents. Consid-

ering dominant molecular markers, either the visible band of a particular molecular fragment or the non-visible 'null allele' may be passed from the parent to the hybrid. The fractions of bands from a parental molecular profile that is not found in the profile of the hybrid depends on the rate of heterozygosity of the used molecular markers in the parent. Assuming no heterozygosity in an artificial dataset, the phenetic measure 'Hybrid distance' was developed and used as a distance measure in an agglomerative cluster analysis (van Raamsdonk 1999). The levels of Hybrid Distance, as well as the dendrograms, confirmed the designated position of hybrids, backcross hybrids and hybrids with secondary evolution in the artificial dataset.

The analysis of biodiversity by means of amplified fragment length polymorphisms (AFLPs) has occasionally been employed, e.g. in rice (Zhu et al. 1998), in maize (Pejic et al. 1998) or in lettuce (Hill et al. 1996). The phylogeny of some potato species has also been reconstructed based on AFLP diversity by Kardolus et al. (1998). The homology of AFLP bands was cheked by Rouppe van de Voort (1997). Of a range of putative homologous fragments 19 out of 20 tested fragments appeared to have almost identical sequences. A similar result in onion and its relatives was obtained by Wilkie et al. (1993) with RAPDs: seven co-migrating bands, of which the fragments were amplified by random primers and subsequently sequenced from different species, were all homologous. Pejic et al. (1998) investigated the genetic similarity among maize inbred lines with known pedigrees using RFLPs, PCR with random primers (RAPDs), microsatellites and AFLPs, and found an accuracy of AFLPs comparable to that of RFLPs. The genetic-similarity trees after AFPLs, RFLPs and microsatellites were highly correlated, but the AFLP tree showed the highest correlation with the pedigree data (Pejic et al. 1998).

In the present study a set of AFLP markers will be used for studying the diversity in nuclear DNA (nDNA) subjected to cladistic, phenetic and Hybrid Distance analyses. This study includes equal numbers of accessions of the sections *Cepa, Rhizirideum* and *Schoenoprasum* in order to ensure a balance in the resulting phylogeny. Some species of section *Oreiprason* (three species), of section *Petroprason* (one) and of section (*Reticulato-bulbosa* (one) were added for comparison. *Allium tuberosum* will be used as outgroup. The aim of the present study is: (1) to analyse the amounts of heterozygosity in order to validate the use of a dominant marker system for phylogenetic studies, (2) to unravel phylogenetic relationships in *Allium* subgenus *Rhizirideum* and especially within and among the sections *Cepa, Rhizirideum* and *Schoenoprasum,* (3) to study relationships and domestication in section *Cepa,* and (4) to determine the phylogenetic position of *A. roylei*.

Materials and methods

Plant material

The species and accessions included in the analysis are listed in the Appendix. The sections *Cepa, Rhizirideum* and *Schoenoprasum* are chosen as a core of the set of taxa under study for the putative systematic position of *A. roylei* in one of these sections and because of the expected crossability relationships with *A. cepa*, whereas the included species and accessions are chosen randomly from these sections. Vouchers are deposited in either WAG, GAT, WAHO or K.

Five interspecific hybrids between five wild *Allium* species and *A. cepa*, kindly provided by Dr. J. Keller, IPK, Gatersleben, were also used.

A. cepa × *A. senescens*	IPK Gatersleben, All 1111,
A. cepa × *A. karelinii*	IPK Gatersleben, All 1092,
A. cepa × *A. globosum*	IPK Gatersleben, All 1052,
A. cepa × *A. hymenorrhizum*	IPK Gatersleben, All 1093,
A. cepa × *A. obliquum*	IPK Gatersleben, All 1088.

The background of these hybrids is documented by Keller et al. (1996).

DNA-isolation

Genomic DNA was isolated from leaf material with the mini-prep DNA-isolation method: approximately 0.25 g of fresh leaf material was collected in Eppendorf tubes, frozen and ground in liquid nitrogen and stored at –50°C. Seven hundred and fifty microliters of DNA-isolation buffer (IB) with $Na_2S_2O_5$ (3.8 g/l) was added to the leaf material. This mixture was incubated for 60 min at 65°C with occasional inverting of the tubes [IB = lysis buffer : extraction buffer : sarkosyl (5% w/v) = 2.5 : 2.5 : 1; lysis buffer = 0.2 M Tris-HCl pH = 7.5, 0.05 M EDTA, 2 M NaCl, 2% w/v CTAB; extraction buffer= 0.35 M sorbitol, 0.1 M Tris HCl pH 7.5, 5 mM EDTA]. The DNA was further purified by adding 750 µl of chloroform / isoamylalcohol (24 : 1), inverting the tubes (10–20 times) and centrifuging for 5 min at 14 000 rpm. After transfer of the supernatant to a new tube the DNA was precipitated by the addition of 400 µl of isopropanol (–20°C). The DNA could either be hooked out or had to be pelleted for 5 min at 14 000 rpm. The DNA samples were washed once with 70% ethanol and re-suspended in 100 µl of TE; 10 µl of the 100-µl DNA suspension was routinely used for AFLP reactions.

AFLP analysis

AFLP reactions for the mapping population were carried out as described by Vos et al. (1995) with the *Eco*RI/*Mse*I restriction enzyme combinations. A total of seven selective nucleotides were used (+3, +4). A two-step amplification procedure was conducted. In order to avoid a less-stringently acting selective nucleotide (on the third position from the 3′ end) the two pre-amplification primers had combined three selective nucleotides (+1, +2). The reaction mix after pre-amplification was diluted to 1/20 and 12.5 µl was used in the final amplification. Amplified fragments were separated on denaturating polyacrylamide gels.

Scoring of data

AFLPs of nuclear DNA were produced using four different primer combinations (Table 1). The primer combinations were chosen based on the maximum diversity in a selected set of species. In all cases a large diversity in the entire set of accessions was found. The variation in the autoradiograms was quantified in order to produce a dataset consisting of 355 characters representing the variable loci in the AFLP profiles.

Data analysis

All analyses are based on a matrix with zero (band absent) and one (band present) scores for each band position per accession. The phylogenetic reconstruction was carried out using the PAUP software package 3.1 (Swofford 1991). Data were treated as unor-

Table 1 Overview of the pre-amplification and amplification primers used

Pre-amplification primers	
E01:	5′ GAC TGC GTA CCA A**TT C**A 3′
M02C:	5′ GAT GAG TCC TGA GT**A A**CC 3′
Amplification primers	
E36:	5′ GAC TGC GTA CCA ATT CAC C 3′
E37:	5′ GAC TGC GTA CCA ATT CAC G 3′
E40:	5′ GAC TGC GTA CCA ATT CAC 3′
E42:	5′ GAC TGC GTA CCA ATT CAC 3′
M52A:	5′ GAT GAG TCC TGA GTA ACC CA 3′
M52T:	5′ GAT GAG TCC TGA GTA ACC CT 3′
M52C:	5′ GAT GAG TCC TGA GTA ACC CC 3′
M52G:	5′ GAT GAG TCC TGA GTA ACC CG 3′

dered and equally weighted. Consensus trees were based on all the most-parsimonius trees of equal length after a heuristic search of trees using the TBR and MULPARS options.

Two reciprocal measures between a hybrid *h* and a parental taxon *p* have been designed for the Hybrid Distance calculations:

$$Dh_h = \frac{s_h}{p + s_h}; \quad Dh_p = \frac{s_p}{p + s_p},$$

where p = number of double matches (1-1 combinations), s_h = sum of hybrid-parent = 0–1 combinations (i.e. a ‘0‘ in the hybrid and a ‘1‘ in the parent), and s_p = sum of hybrid-parent = 1–0 combinations. In the case of a hybrid that exclusively combines both parental character state profiles, the Dh_h measure equals zero. The value of the Dh_p measure depends on the relationship of both parents with each other. The distance measures Dh_h and Dh_p will be referred to as the Hybrid Distance. Application of the Hybrid Distance methodology to a combination of two species yields an indication of the band-sharing of one species with another. The principal difference between the Hybrid Distance and Jaccard’s distance a.o. is the difference between the reciprocal distances. Further information and testing of the Hybrid Distance is given by van Raamsdonk (1999).

For a range of phenetic methods a symmetric distance matrix is required, i.e. both reciprocal distances need to be equal. In the case of the Hybrid Distance both reciprocals are not equal due to obvious reasons. Since the Hybrid Distance is designed to find unexpected short distances in putative parent-daughter taxon combinations, a symmetric matrix can be designed using the shortest distance in each combination.

Phenetic analyses were carried out with the program package IRIS (van Raamsdonk 1988). UPGMA cluster analysis was based on the Jaccard’s Distance coefficient, and Single Linkage clustering was carried out based on the the Hybrid Distance matrix optimised for the shortest reciprocal distance in each combination.

Results

Phylogenetic and phenetic analyses

Scoring of the variation of AFLP fragments on four autoradiograms resulted in 355 characters per accession. Parts of the profiles of some species are presented in Fig. 1. The profile of *Allium oschaninii* deviates markedly from all other profiles because of the very low number bands in the profile. This situation might be due to an erroneous pre-amplification and, therefore, this profile is not included in further analyses.

The phylogenetic analysis with PAUP produced 56 most-parsimonious trees of 670 steps with a consistency index (CI) of 0.528 ($CI_{uninformative}$ = 0.471) and a retention index (RI) of 0.744. One of the most parsimonious trees is presented in Fig. 2. The 80% majority rule consensus tree is presented in Fig. 3. All three sections with a major share in the dataset, i.e. sections *Cepa*, *Schoenoprasum* and *Rhizirideum*, each possess a single separated clade in the tree. The species of sections *Oreiprason* and *Petroprason* are mixed in one clade. The tree after bootstrap analysis showed the same topology as the majority rule consensus tree. The same major division between sections was found after UPGMA cluster analysis based on a distance matrix calculated using the Jaccard distance (Fig. 4). The first division within section *Cepa* separates *A. fistulosum / Allium altaicum* (subsection *Phyllodolon*) from the other species (subsection *Cepa*) in both analyses. *Allium lineare* (section *Reticulato-Bulbosa*) in both trees is connected to species of the section *Rhizirideum*. Differences between the two trees are mainly found at the level of connection between the sections. The cladogram of Fig. 3 shows the position of

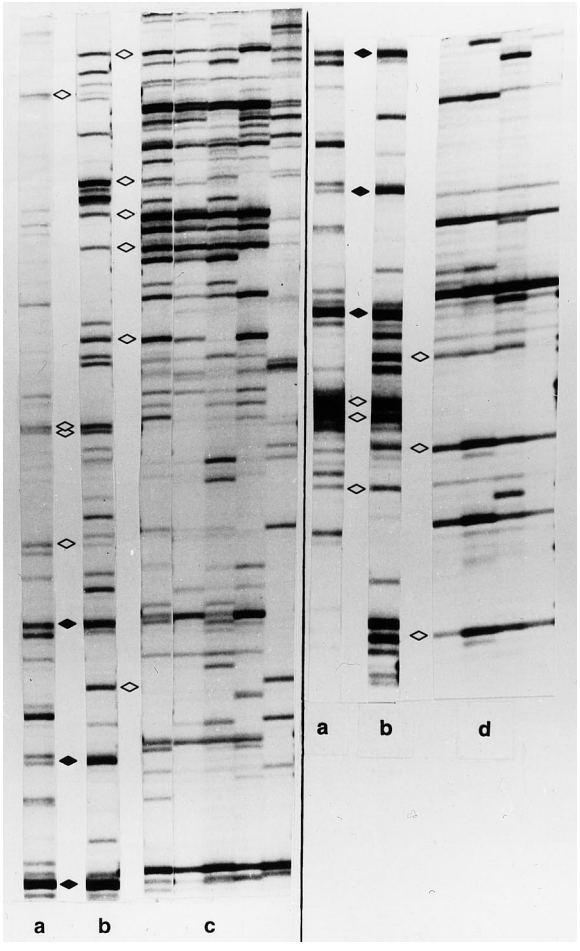

Fig. 1a–d Parts of AFLP profiles for selected species. **a** *A. vavilovii*, **b** *A. roylei*, **c** *A. nutans* (left four lanes) and *A. angulosum* (*right lane*), **d** *A. schoenoprasum* (left two lanes) and *A. altyncolicum* (right two lanes). The *connecting diamonds* indicate band sharing; *black diamonds* in the comparison between **a** and **b** point to identical bands in the profiles

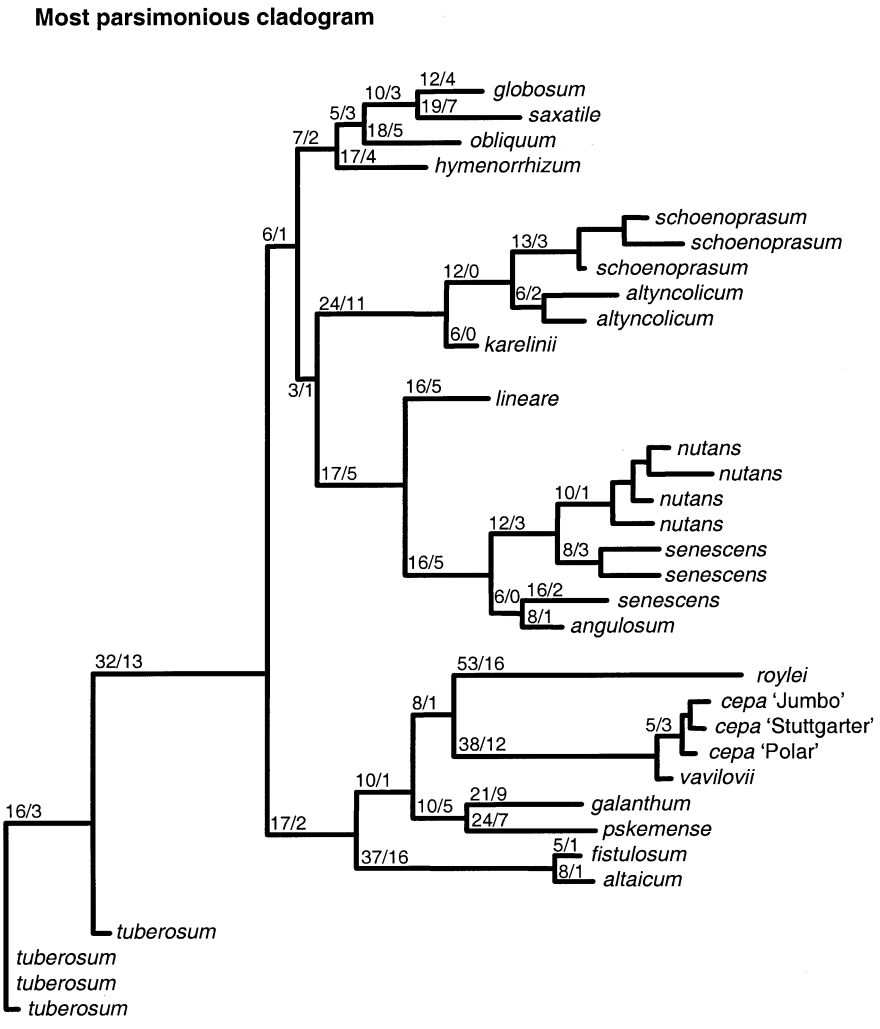

Fig. 2 One of the 56 most-parsimonious trees after phylogenetic analysis of 355 AFLP characters scored for 33 accessions of *Allium* subgenus *Rhizirideum*. Length: 670 steps, CI = 0.528, $CI_{uninformative}$ = 0.471, RI = 0.744. *Indices* (x/y) at each branch indicate the total number of bands supporting the branch (x) and the number of unique bands (y)

section *Cepa* in one clade against all other sections in one other main clade, except for the outgroup taxa of section *Butomissa*, whereas the cluster with section *Cepa* is primarily connected to a cluster with the representatives of section *Rhizirideum* in the dendrogram (Fig. 4). *A. tuberosum* occupies a position among the other clusters since a dedicated outgroup does not exist in phenetic analysis. The connecting level between the main clusters is highly comparable, which indicates that they are almost equally different from each other.

In both trees the position of *A. roylei* is among the representatives of section *Cepa*. However, the single branch of *A. roylei* is clearly the longest one in the cladogram of Fig. 2. *A. roylei* shares eight apomorphisms with the clade of *A. cepa / Allium vavilovii*, but among the 53 apomorphisms of its own branch four reversals are found. Of the 37 synapomorphisms (Fig. 2: 53 minus 16) determining the position of *A.roylei* in the cladogram, a total of 26 of these is also found in any branch of section *Rhizirideum* and eight in the clade of section *Schoenoprasum*. The highest number of synapomorphisms in a single branch are shared by *A. roylei* and *Allium pskemense* (five) and by *A. roylei* and two *Allium nutans* accessions (four and three respectively). Fifty six most-parsimonious trees, excluding *A. roylei*, had 614 steps, a CI of 0.550 ($CI_{uninformative}$ = 0.496) and a RI = 0.766. The majority rule consensus tree of these 56 trees showed the same topology as presented in Figs. 2 and 3, but consequently without *A. roylei*.

Hybrids

The profiles of the five hybrids were screened for band sharing with both their parents. Most bands present in the hybrids can be traced back to one or both of the parents. Some hybrid profiles, together with those of the parents, are presented in Fig. 5. The hypothesis is that hybrids share the bands of both parents. Bands absent in the hybrid may point to heterozygosity for that fragment in the parent and to the situation that the null allele is passed through, or this may be due to the fact that the exact parental individual has not been used for the AFLP screening. The amounts of shared or not-shared bands are presented in Table 2. The percentage of the *A. cepa* bands that are not found in the profiles of the hybrids

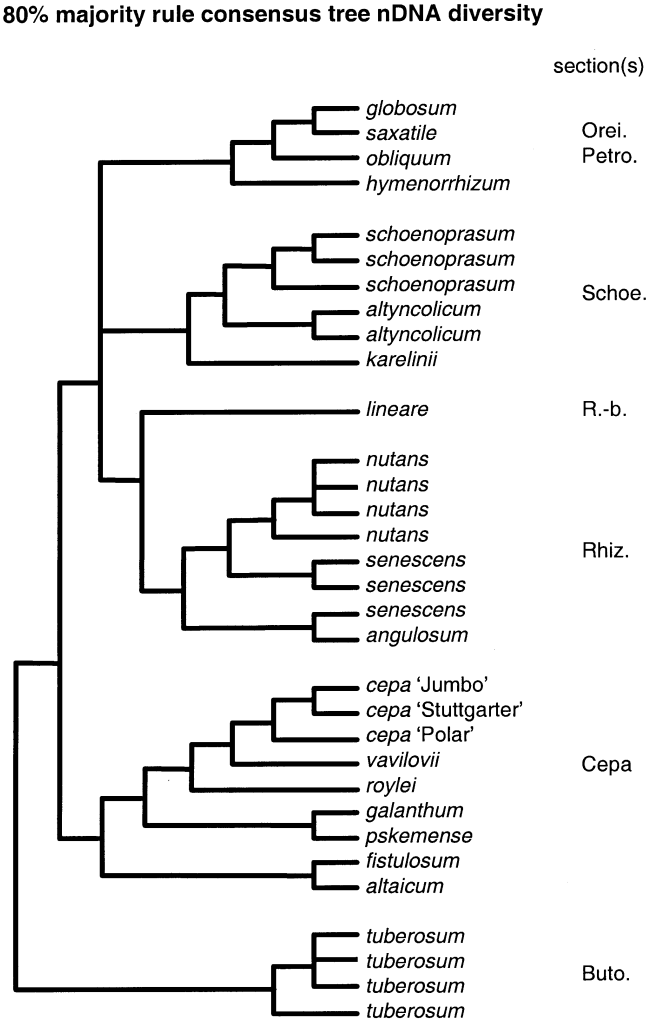

Fig. 3 Eighty percent majority rule consensus tree of the 56 most-parsimonious trees showing the evolutionary position of 33 accessions of *Allium* subgenus *Rhizirideum*. Section names are abbreviated: *Cepa*: section *Cepa*; *Rhiz*: section *Rhizirideum*; *Schoe*: section *Schoenoprasum*; *Orei*: section *Oreiprason*; *Petro*: section *Petroprason*; *R-b*: section *Reticulato-bulbosa*; *Buto*: section *Butomissa*

ranges from 7.0 to 14.3%, whereas the five wild species as paternal parents did not pass 6.3–13.8% of their fragments to the hybrids. These percentages are an indication of the heterozygosity of the respective species. The number of bands unique to the hybrids, i.e. not found in the profiles of either parent, is very low. The finding of these bands can be explained by the fact that the comparison between the hybrids and the parents as depicted in Table 2 was not carried out using the exact parent individuals; rather, instead of these individuals, representatives of the same population were used.

Hybrid Distance analysis

Analysis using Hybrid Distance results in a measure which indicates the fraction of bands of an accession that is also found in the profile of another accession (''band sharing''). The design of the Hybrid Distance measure includes two reciprocal figures for the same combination of two accessions. Pooling of the values per taxon results in distances ranging from 0.03 between the cultivars of *A. cepa* to 1.0 between *A. tuberosum* and a representative of section *Schoenoprasum*, the latter case indicating that the two profiles do not have bands in common. An overview of the calculated Hybrid Distances is shown in Table 3. Species within (sub-)sections are separated from each other at distances of between 0.08 and 0.60. Distances between sections are found at levels from 0.75 to 1.0.

There is a large amount of band sharing between the two included cultigens (*A. cepa* and *A. fistulosum*) and their presumed immediate ancestors (*A. vavilovii* and *A. altaicum*, respectively). The distance from the cultigen to the wild ancestor (*cepa-vavilovii*: 0.03–0.07, and *fistulosum-altaicum*: 0.09) is shorter than the reciprocal value (*vavilovii-cepa*: 0.10–0.12, and *altaicum-fistulosum*: 0.14), which indicates that the cultigens show very few bands that are absent in their respective ancestors. Large differences between reciprocal distances were found between *A. fistulosum* and *Allium galanthum / A. pskemense*,

Table 2 Numbers of shared bands. For each cross the paternal parent (pp) is given with the number of bands found in the hybrid and both parents together, the number of bands shared by the hybrid and its maternal parent, found in the maternal parent only, the number of bands shared by the hybrid and its paternal parent, found in the paternal parent only and bands visible in the hybrid only. The percentages in the columns with the number of bands found in the parents only indicate the degree of heterozygosity in either parent

Paternal parent (pp)	All (hybr. and parents)	*Cepa* and hybr.	*Cepa* only		pp. and hybr.	pp. only		hybr. only
Cross with cv Polar as mother								
A. senescens	9	45	9	14.3%	38	8	13.8%	1
Crosses with cv Stuttgarter Riesen as mother								
A. karelinii	5	48	4	7.0%	17	3	10.7%	1
A. globosum	11	41	5	8.8%	31	6	12.5%	1
A. hymenorrhizum	8	44	5	8.8%	31	6	13.3%	2
A. obliquum	8	43	6	10.5%	23	2	6.3%	1

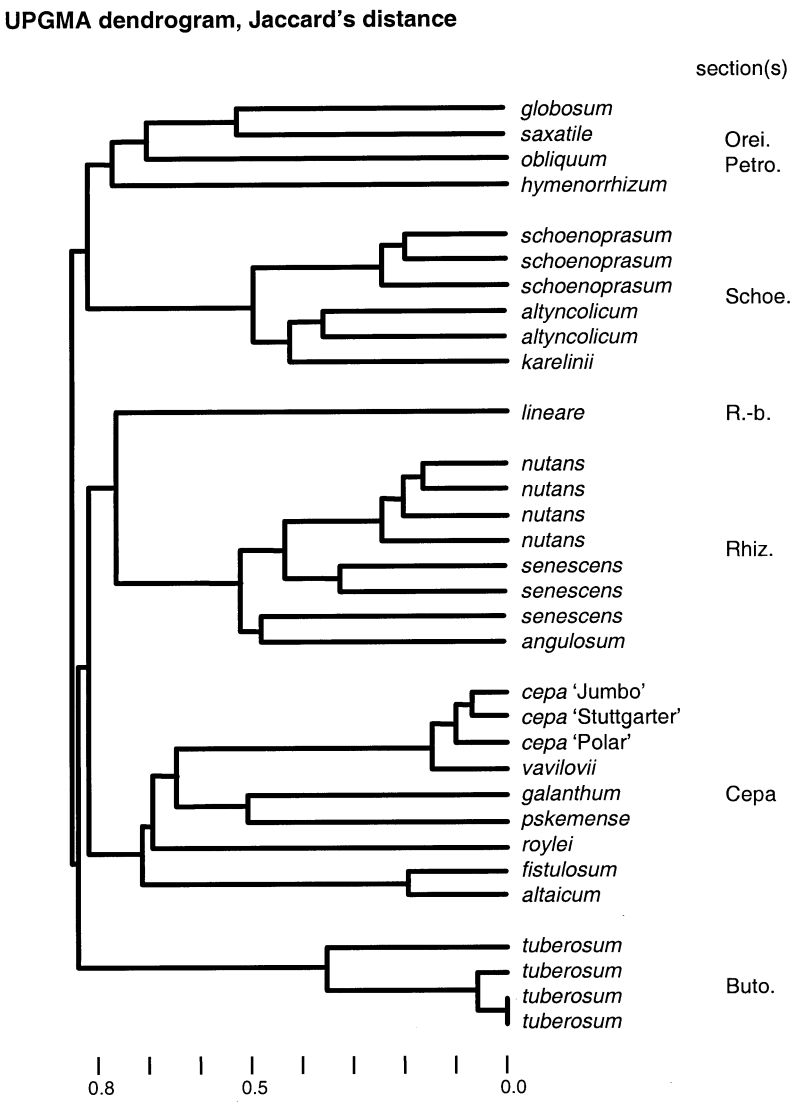

Fig. 4 UPGMA dendrogram based on Jaccard's coefficient showing the systematic position of 33 accessions of *Allium* subgenus *Rhizirideum*. Section names are abbreviated as in Fig. 3. Cophenetic correlation coefficient R = 0.989

Table 3 Hybrid distances between taxa of *Allium* subgenus *Rhizirideum*. The figures on the diagonal (in **bold**) represents the distances between accessions within species or between accessions and species within sections. The other figures indicate the distances between the indicated taxa. The figures in *italics* indicate levels between those of the distances within and between sections, i.e. the distances larger than 0.60 and shorter than 0.75

A. cepa	**0.03–0.07**	0.03–0.07	0.47–0.56	*0.62–0.64*	*0.67–0.69*	*0.65–0.67*	0.74–0.87	0.83–0.90	0.78–0.83
A. vavilovii	0.10–0.12	**0**	0.49–0.54	*0.64*	*0.69*	*0.67*	0.76–0.87	0.86–0.92	0.85–0.86
A. galanthum / pskemense	0.57–0.61	0.57	**0.34–0.41**	*0.63–0.64*	*0.64–0.66*	*0.65–0.72*	0.76–0.92	0.83–0.94	0.89–0.94
A. fistulosum	*0.60–0.61*	*0.61*	0.52–0.56	**0**	0.09	*0.63*	0.75–0.90	0.86–0.94	0.78–0.80
A. altaicum	*0.69–0.71*	*0.69*	*0.60–0.62*	0.14	**0**	*0.72*	0.81–0.90	0.83–0.94	0.75–0.80
A. roylei	*0.61–0.62*	0.60	0.56–0.60	*0.63*	*0.66*	**0**	*0.68–0.77*	0.79–0.90	0.89–0.92
Rhizirideum	0.81–0.92	0.80–0.89	0.75–0.92	0.78–0.90	0.81–0.91	0.73–0.86	**0.09–0.49**	0.79–0.96	0.90–0.98
Schoenoprasum	0.87–0.94	0.88–0.94	0.82–0.95	0.90–0.97	0.88–0.95	0.87–0.94	0.82–0.96	**0.08–0.51**	0.89–1.00
A. tuberosum	0.80–0.88	0.86–0.90	0.88–0.96	0.80–0.84	0.78–0.80	0.87–0.93	0.88–0.97	0.83–1.00	**0.0–0.35**
	cepa	*vavilovii*	*gal / psk*	*fistulosum*	*altaicum*	*roylei*	*Rhizi.*	*Schoe*	*tub.*

and in most pairs of taxa with *A. roylei* involved (Table 3). *A. roylei* shows a shorter distance to every other taxon (row in Table 3) than the other taxa do to *A. roylei* (column in Table 3) except for the comparison with *A. tuberosum*. Within section *Cepa* there is a gradual difference between three groups after pooling of the reciprocal results. The distance of the *cepa/vavilovii* group with the *galanthum/pskemense* group ranges from 0.47 to 0.61, whereas the distance with the *fistulosum/altaicum* group is 0.60–0.71. The distance between the *galanthum/pskemense* and

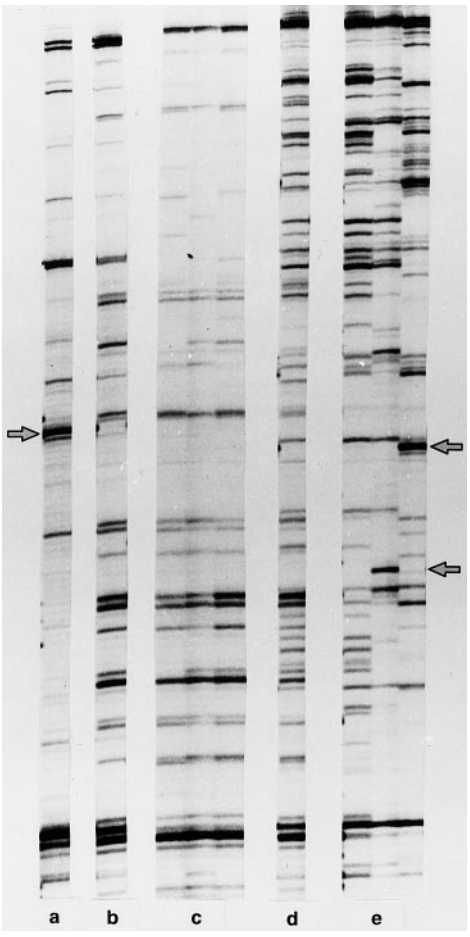

Fig. 5a–e Parts of AFLP profiles for selected hybrids with their parental species. **a** *A. karelinii*, **b** hybrid between *A. cepa* and *A. karelinii*, **c** *A. cepa* (cultivars 'Jumbo', 'Stuttgarter' and 'Polar'), **d** hybrid between *A. cepa* and *A. senescens*, **e** *A. senescens* (three accesions). *Arrows* indicate bands that are present in the parental species but absent in the hybrid

the *fistulosum/altaicum* groups is 0.52–0.63. These figures indicate that the *fistulosum/altaicum* group shares the lowest number of bands with the other two groups.

The percentages of bands found in a parent and not in the hybrid can be directly translated to Hybrid Distances, considering the structure of the formulae. Hence, the Hybrid Distances for the artificial hybrids to each of their parents range from 0.06 to 0.14. These distances are larger than found between the cultivars of *A. cepa* and are comparable to the distances between the cultigens and their ancestors.

The position of *A. roylei* is intermediate. This species is separated from sections *Cepa* and *Rhizirideum* at distances (from 0.56 to 0.77) that are intermediate between the distance levels within section and between sections, respectively. These distances indicate that *A. roylei* shares a considerable number of bands with species of both sections, as is also shown in Fig. 1 and indicated for Fig. 2.

Discussion

Within-accession diversity and hybrid analysis

The AFLP profiles are almost exclusively based on the variation in the nuclear genomes of the species. This statement is based on the fact that the nuclear genome is about 33 978-times as large as the organeller genomes [15 290 megabase pairs for the *A. cepa* genome (Arumuganathan and Earle 1991) and approximately 450 kb for the cp and mt genomes together], which means that less than 1 out of 30 000 fragments might be amplified from an organellar genome. Moreover, all 262 bands found in *A. cepa* and absent from *A. roylei* segregated in the F_2 population with *A. cepa* as mother (van Heusden et al. 2000). Inheritance of a band from the mother to an F_2 population by means of an organelle would result in 100% of the F_2 individuals showing the band instead of a segregating pattern. These figures indicate that it is highly unlikely to find any measurable effect on the presented AFLP results caused by organellar diversity.

The estimated amount of within-accession variation, as concluded from the non-shared bands in the parent-hybrid combinations (Table 2), indicates that a vast majority of the bands represent identical alleles at both homologous loci. Taking into account that the currently presented AFLP variation is based on nuclear genome diversity this situation indicates a high level of homozygosity. The estimations of heterozygosity taken from the hybrid population between *A. cepa* and *A. roylei* (see Van Heusden et al. 2000) are 8% and 12.6%, respectively. The very low numbers of bands unique to the hybrids, due to the fact that the profile comparison was carried out at the level of populations, indicates that relatively low levels of variation within wild populations can be expected in the parental species employed. Although this assumption does not allow any conclusion about the rate of heterozygosity, low intrapopulational levels of diversity are a logical result of a low rate of heterozygosity. This situation is supported by other studies as discussed by Van Raamsdonk (1999). The absence or presence of bands in a profile is apparently not seriously influenced by individual allele differences, which is also indicated by the almost complete consensus of the clades (Fig. 2) and clusters (Fig. 3) with the currently employed taxonomic division in sections, groups and species of the subgenus *Rhizirideum*, as based on morphology, cytogenetics or chemical compounds (El-Gadi and Elkington 1977; Hanelt 1990; Hanelt et al. 1992; van Raamsdonk and de Vries 1992; Hanelt and Fritsch 1994).

The Hybrid Distance approach appears to be useful for analysing levels of diversity (Table 3). The distances among species within sections and among sections in the current results are generally at a higher level compared to the results of van Raamsdonk (1999). The previous results were based on a dataset with 107 characters, which is considerably lower than the current 355

characters and on a smaller range of species. Especially, the distances of *A. roylei* with other species and sections are higher in the present results than in the situation with a smaller dataset, but remain lower than the distances between sections. The use of basic distances, as presented in Table 3, might be more powerful than the construction of a dendrogram based upon these data (cf. McDade 1997) since information is lost at every step in the calculation of a dendrogram. The Hybrid Distance approach appears to be useful for supporting the designation of a subsection *Phyllodolon* in the section *Cepa* which includes *A. fistulosum* and *A. altaicum* (Hanelt et al. 1992). The lower distances of *A. roylei* with most other taxa, rather than vice versa (row-wise vs column-wise comparisons in Table 3), indicate a relatively high fraction of bands in the *A. roylei* profile shared with other taxa. This is an important aspect for discussing the origin of this species.

Hybrid Distance is especially designed for analysing hybrids and hybrid speciation. The distances between the hybrids and their parents calculated according to the Hybrid Distance formula are logically equal to the amounts of heterozygosity in the parents (Table 2). Since no secondary evolution of the artificial hybrids took place, there is no further influence on the distances between the hybrids and both of the respective parents. Although, until recently, no dedicated tools for detecting and analysing hybrids were stated to occur (McDade 1997; Kardolus et al. 1998), ordination methods resulted in intermediate positions of hybrids in a range of studies (e.g. Adams 1982; Brochmann 1987). An intermediate position between the parents with a small but detectable maternal effect with respect to morphological traits was also found with artificial hybrids between species of *Allium* section *Cepa* (van Raamsdonk and de Vries 1992). Another approach to analysing the spatial difference between hybrids and their parents was proposed by Wells (1980). Intermediacy after an ordination analysis of the diversity of wild species does not yield a proof for hybrid speciation. Both the ordination method and the Wells hybrid analysis can be applied successfully for identifying hybrids only when a reasonable assumption about hybrid origin is available. Provided that a hypothesis for the nature of hybrids is available (McDade 1997), the Hybrid Distance approach may fill the gap in the range of analytical techniques. The hypothesis that the molecular profile of a hybrid shares the bands of the profiles of both the parents, dependent on the rate of heterozygosity, could be used as a basis for the Hybrid Distance approach.

Phylogeny

Several molecular diversity studies based on the nuclear DNA as well as the organellar DNA of the genus *Allium*, or parts of it, have been published. For a proper comparison the current nDNA diversity pattern will be discussed exclusively together with the results of other nDNA studies. The phenetic study of Dubouzet et al. (1997), including a range of different species of the subgenus *Rhizirideum*, is now well-supported by the present results. The fact that several important species, such as *A. vavilovii* as a progenitor of *A. cepa* and *A. roylei*, are excluded from their analysis and that only a UPGMA dendrogram is provided (Dubouzet et al. 1997) implies that evolutionary conclusions can not be drawn from their study. Comparison with the results of Wilkie et al. (1993) and of Bradeen and Havey (1995) is difficult because of the smaller set of species used in these studies, which might influence the individual position of accessions (e.g. Sanderson and Donoghue 1989). The current study is the first phylogenetic analysis with a balanced set of species of the sections *Cepa*, *Schoenoprasum* and *Rhizirideum* of the subgenus *Rhizirideum*. A firm basis is provided for the systematics and phylogeny of these sections.

The choice of *A. tuberosum* as an outgroup (Havey 1992; van Raamsdonk et al. 1997) is strongly supported by the long branch in Fig. 2 that separates *A. tuberosum* from all other species. A remote position of *A. tuberosum* and a connection of a representative of the section *Reticulato-Bulbosa* (*A. lineare* in our study) with the species of the section *Rhizirideum*, as far as can be measured by phenetic distance, can also be seen in the results of Dubouzet et al. (1997), although their single accession of the section *Reticulato-Bulbosa* represents a different species (*Allium cyaneum*).

The combination of species of the sections *Oreiprason* and of *Petroprason* in one cluster (Fig. 3) is congruent with the close resemblance of the species of these two sections in both morphology and geographical distribution (Hanelt et al. 1992) as well as the phenetic analysis of nDNA (Dubouzet et al. 1997).

Relationships and domestication in section *Cepa*

The three alliances in section *Cepa* as proposed by Hanelt (1990) are found in the current results in both the cladogram and the phenogram (Figs. 3 and 4). The three groups, i.e. *A. fistulosum* and *A. altaicum*, *A. galanthum* and *A. pskemense*, and *A. cepa*, *A. vavilovii* and *A. oschaninii*, are also found after morphological analysis (van Raamsdonk and de Vries 1992), after analysis of a supra-nuclear dataset (van Raamsdonk et al. 1997), and after nDNA analysis (Bradeen and Havey 1995). Friesen et al. (1999) also detected a close resemblance between *A. altaicum* and *A. fistulosum* after RAPD nDNA analysis and a monophyletic origin of the cultigen *A. fistulosum*. In the phenogram of Dubouzet et al. (1997) *A. fistulosum* and *A. altaicum* do not form a cluster together, which is also the case for *A. galanthum* and *A. pskemense*. The separation of both species pairs is different from our results (Figs. 3 and 4). The re-establishment of subsection *Phyllodolon* (*A. fistulosum* and *A. altaicum*; Hanelt et al. 1992) against subsection *Cepa* (all other species of section *Cepa*) is supported by the trees pre-

sented in Figs. 3 and 4 as well by the gradual differences between the species and alliances expressed by means of Hybrid Distance.

The very short distances among the cultivars of *A. cepa* and the relatively short distances between the cultigens *A. cepa* and *A. fistulosum* and their immediate ancestors fits with the Soybean scenario of domestication, wich describes the domestication of crops that are selected from one progenitor with a subsequent narrowing of the genetic diversity of the crop (van Raamsdonk 1995; Friesen 1999). It can be predicted from this scenario that all bands found in the crop will also be present in the wild progenitor and, hence, that the distance from the crop to the progenitor should be zero. In the case of *A. cepa* and *A. fistulosum* distances of 0.03–0.09 to the respective ancestors are found. The current data are based on dominant markers, which means that the distances discussed depend on the rate of heterozygosity. The distances from 0.03 to 0.09 are lower than the estimated rates of heterozygosity. Other co-dominant marker systems, such as RFLPs or isozyme data, would result in lower distances. So far, the hypothesis that the Hybrid Distance approach might be useful for the analysis of domestication scenarios can be confirmed by the present results. The application of Hybrid Distance to other crops and domestication scenarios (van Raamsdonk 1995, 1999) and to other marker systems will provide more insight into its usefulness.

Origin of *A. roylei*

The position of *A. roylei* in either section *Rhizirideum* or in section *Cepa* has been previously discussed (van der Meer and de Vries 1990; de Vries et al. 1992). Analysis of a preliminary dataset revealed an intermediate position of *A. roylei* and confirmed the usability of the Hybrid Distance approach in comparison with, and in addition to, cladistic analysis, cluster analysis and analysis with the SplitsTree approach (van Raamsdonk 1999). The present results covering an extensive range of nDNA diversity shows a position of *A. roylei* in the section *Cepa*, which is in agreement with crossability results (van Raamsdonk et al. 1992). However, a considerable number of synapomorphisms occur between *A. roylei* and species of the section *Rhizirideum*. This might indicate a hybrid origin with a considerable process of subsequent secondary evolution, as is also found after analysis by means of the Hybrid Distance (Table 3). An evolutionary background that combines parts of the genomes of different species might explain the possibility that *A. roylei* bridges the genomes of *A. cepa* and *A. fistulosum* (Khrustaleva and Kik 1998, 2000). Extended crossability analysis with other species is necessary. The study of cpDNA diversity in the same set of species would allow an extensive comparison with the present results and with those of Havey (1992), Linne von Berg et al. (1996) and Mes et al. (1997).

Acknowledgements We thank our colleagues from CPRO-DLO, Dr. A.W. van Heusden and Dr. J.H. Sandbrink, for their support of this study. Dr. J. Keller (IPK, Gatersleben, Germany) kindly provided the hybrid accessions. Dr. A.W. van Heusden provided the basic data of the *A. cepa* / *A. roylei* hybrid population. This research was supported by the EC-FAIR Programme (FAIR CT95-465) and by the Dutch Ministry of Agriculture, Nature Management and Fisheries.

References

Adams RP (1982) A comparison of multivariate methods for the detection of hybridization. Taxon 31:646–661

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Pl Mol Biol Rep 9:208–218

Bradeen JM, Havey MJ (1995) Randomly amplified polymorphic DNA in bulk onion and its use to assess inbred integrity. J Am Soc Hort Sci 120:752–758

Brochmann C (1987) Evaluation of some methods for hybrid analysis, exemplified by hybridization in *Argyranthum* (Asteraceae). Nordic J Bot 7:609–630

Damme JMM van (1992) Hybridisation between wild and transgenic plants. In: Weverling J, Hekstra GP (eds) Ecological effects of genetically modified organisms, Netherlands Ecological Society, Arnhem, pp 81–92

Dubouzet JG, Shinoda K, Murata N (1997) Phylogeny of *Allium* L. subgenus *Rhizirideum* (G. Don ex Koch) Wendelbo according to dot-blot hybridization with randomly amplified DNA probes. Theor Appl Genet 95:1223–1228

El-Gadi A, Elkington TT (1977) Numerical taxonomic studies on species in *Allium* subgenus *Rhizirideum*. New Phytol 79:183–201

Friesen N, Pollner S, Bachmann K, Blattner FR (1999) RAPDs and noncoding chloroplast DNA reveal a single origin of the cultivated *Allium fistulosum* from *A. altaicum* (Alliaceae). Am J Bot 86:554–562

Hanelt P (1990) Taxonomy, evolution and history. In: Brewster JL, Rabinowitch HD (eds) Onions and allied crops, volume 1: biochemistry, food science, and minor crops. CRC Press Inc., Boca Raton, Florida, pp 1–26

Hanelt P, Fritsch R (1994) Notes on some infrageneric taxa in *Allium* L. Kew Bull 49:559–564

Hanelt P, SchultzeMotel J, Fritsch R, Kruse J, Maass HI, Ohle H, Pistrick K (1992) Infrageneric grouping of *Allium* – the Gatersleben approach. In: Hanelt P, HAmmer K, Knüpffer H (eds) The genus *Allium* – taxonomic problems and genetic resources. Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany, pp 107–123

Havey MJ (1992) Restriction enzyme analysis of the chloroplast and nuclear 45s ribosomal DNA of *Allium* sections *Cepa* and *Phyllodolon* (Alliaceae). Pl Syst Evol 183:17–31

Heusden AW van, Ooijen JW van, Vrielink-van Ginkel R, Verbeek WHJ, Wietsma WA, Kik C (2000) Genetic mapping in an interspecific cross in *Allium* with amplified fragment length polymorphism (AFLP) markers. Theor Appl Genet 100:118–126

Hill M, Witsenboer H, Zabeau M, Vos P, Kesseli R, Michelmore R (1996) PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. Theor Appl Genet 93:1202–1210

Kardolus JP, Eck HJ van, Berg RG van den (1998) The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy. Pl Syst Evol 210:87–103

Keller ERJ, Schubert I, Fuchs J, Meister A (1996) Interspecific crosses of onion with distant *Allium* species and characterization of the presumed hybrids by means of flow cytometry, karyotype analysis and genomic in situ hybridization. Theor Appl Genet 92:417–424

Khrustaleva LI, Kik C (1998) Cytogenetical studies in the bridge cross *Allium cepa* × (*A. fistulosum* × *A. roylei*). Theor Appl Genet 96:8–14

Khrustaleva LI, Kik C (2000) Introgression of *Allium fistulosum* into *A. cepa* mediated by *A. roylei*. Theor Appl Genet 100: 17–26
Linne von Berg G, Samoylov A, Klaas M, Hanelt P (1996) Chloroplast DNA restriction analysis and the infrageneric grouping of *Allium* (Alliaceae). Plant Syst Evol 200:253–261
McDade L (1997) Hybrids and phylogenetic systematics. III. Comparison with distance methods. Systematic Bot 22:669–683
Meer QP van der, Vries JN de (1990) An intraspecific cross between *Allium roylei* Stearn and *Allium cepa* L., and its backcross to *A. cepa*. Euphytica 47:29–31
Mes THM, Friesen N, Fritsch RM, Klaas M, Bachmann K (1997) Criteria for sampling in *Allium* based on chloroplast DNA PCR-RFLPs. Systematic Bot 22:701–712
Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M (1998) Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor Appl Genet 97: 1248–1255
Raamsdonk LWD van (1988) IRISsys: system for information regrouping to identify structures, package for multivariate data analysis, release 3.1. Manual (in Dutch). CPRO-DLO Wageningen
Raamsdonk LWD van (1995) The cytological and genetical mechanisms of plant domestication as exemplified by four crop models. Bot Rev 61:367–399
Raamsdonk LWD van (1999) A new measure for analysing hybrid speciation. In: Raamsdonk LWD van, Nijs JCM den (eds) Evolution in man-made habitats, Proc VIIth IOPB symposium (in press)
Raamsdonk LWD van, Vries T de (1992) Biosystematic studies in *Allium* L. section *Cepa*. Bot J Linn Soc 109:131–143
Raamsdonk LWD van, Wietsma WA, Vries JN de (1992) Crossing experiments in *Allium* L. section *Cepa*. Bot J Linn Soc 109:293–303
Raamsdonk LWD van, Smiech MP, Sandbrink JM (1997) Introgression explains incongruency between nuclear and chloroplast DNA-based phylogenies in *Allium* section *Cepa*. Bot J Linn Soc 123:91–108
Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: Harrison RG (ed) Hybrid zones and the evolutionary process, Oxford University Press, New York, pp 70–109
Rouppe van de Voort J, Zandvoort van P, Eck HJ van, Folkertsma RT, Hutten RCB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol Gen Genet 255:438–447
Sanderson MJ, Donoghue MJ (1989) Patterns of variation in levels of homoplasy. Evolution 43:1781–1795
Stuessy TF (1990) Plant taxonomy. Columbia University Press, New York
Swofford D (1991) Phylogenetic analysis using parsimony, version 3.1
Vos P, Hogers R, Bleeker M, Reijans M, Lee T van de, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
Vries JN de, Wietsma WA, Jongerius MC (1992) Introgression of characters from *Allium roylei* Stearn into *A. cepa* L. In: Hanelt P, Hammer K, Knüpffer H (eds) The genus *Allium* – taxonomic problems and genetic resources. Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany, pp 321–325
Wells H (1980) A distance coefficient as a hybridization index: an example using *Mimulus longiflorus* and *M. flemingii* (Scrophulariceae) from Santa Cruz, California. Taxon 29:53–65
Wilkie SE, Isaac PG, Slatter RJ (1993) Random amplified polymorphic DNA (RAPD) markers for genetic analysis in *Allium*. Theor Appl Genet 86:497–504
Zhu J, Gale MD, Quarrie S, Jackson MT, Bryan GJ (1998) AFLP markers for the study of rice biodiversity. Theor Appl Genet 96:602–611

Appendix

List of plant material used

Name	Source and origin	CPRO ref. no.
Section *Cepa*:		
A. cepa 'Jumbo' 'Stuttgarter' 'Polar'		96147
A. vavilovii	Grown from 83010: H.B. Chorog, wild origin	95081
A. oschaninii	H.B. Budapest	78227
A. galanthum	Grown from 79147: USDA Beltsville	82550
A. pskemense	H.B. Alma-Ata	65448
A. fistulosum	Grown from 76201: H.B. Odessa	84236
A. altaicum		91126
Section *Rhizirideum*:		
A. roylei	Grown from 79150: USDA Beltsville C 502 originally from India	95001
A. senescens	All Union Res. Inst. Veg. USSR	89010
	IPK Gatersleben s.n.	93005
	IPK: All 1122	96159
A. nutans	All Union Res. Inst. Veg. USSR	89008
	All Union Res. Inst. Veg. USSR	89009
	All Union Res. Inst. Veg. USSR	89011
	IPK Gatersleben s.n.	93012
A. angulosum	IPK: Tax 256	96149
Section *Schoenoprasum*:		
A. schoenoprasum	IPK Gatersleben s.n.	93007
	IPK: All 911	96158
	IPK: Tax 1605	96152
A. altyncolicum	IPK: Tax 42	96150
	IPK Gatersleben s.n.	93008
A. karelinii	IPK: Tax 536	96154
Section *Oreiprason*:		
A. globosum	IPK: All 1123	96151
A. saxatile	IPK: Tax 632	96157
A. hymenorrhizum	IPK: Tax 750	96153
Section *Petroprason*:		
A. obliquum	IPK: Tax 94	96156
Section *Reticulato-bulbosa:*		
A. lineare	IPK: Tax 2335	96155
Section *Butomissa:*		
A. tuberosum	Origin: Thailand	
	All Union Res. Inst. Veg. USSR	89006
	All Union Res. Inst. Veg. USSR	89007
	IPK: Tax 1830	96148