Theor Appl Genet (1990) 79:360–368

THEORETICAL AND APPLIED GENETICS

Phylogeny of wild and cultivated *Solanum* species based on nuclear restriction fragment length polymorphisms (RFLPs)

T. Debener, F. Salamini and C. Gebhardt

Max-Planck-Institut für Züchtungsforschung, Carl-Von-Linné-Weg 10, D-5000 Köln 30, FRG

Received October 23, 1989; Accepted November 2, 1989
Communicated by G. Wenzel

Summary. Phylogenetic relationships between 14 wild and 3 cultivated *Solanum* species, including the European potato, *Solanum tuberosum* ssp. *tuberosum*, were inferred using DNA restriction fragment length polymorphisms (RFLPs) as discriminating characters. Phenetic trees were obtained based on distance matrices as well as on parsimony methods, which were not significantly altered either by the computational method used, or by the individual plant genotypes or RFLP markers chosen for the analysis. The reliability of the tree topologies was assessed and, as expected, it increased with the number of polymorphic restriction fragments scored. The individual genotypes within each species, the different species themselves, and the main branches of the trees were clearly separated from each other. The least reliable parts in the trees were the positions of closely related species within the main clusters. *S. tuberosum* spp. *tuberosum* formed one group closely related with *S. tuberosum* spp. *andigena*, *S. stenotomum*, and *S. canasense*. This was well separated from a second group formed by *S. sparsipilum*, *berthaultii*, *kurtzianum*, *gourlayi*, *vernei*, *spegazzinii*, *chacoense*, and *megistacrolobum*. Two further branches were formed by *S. demissum* and *S. acaule*, and by *S. polyadenium*, *S. pinnatisectum*, *S. stoloniferum*, and *S. etuberosum*. The phenetic trees presented here supported the description of relationships among *Solanum* species based on biosystematic studies, with the exception of the placement of *S. chacoense* and *S. megistacrolobum*.

Key words: RFLP – *Solanum* – Phylogenetic analysis

Introduction

The cultivated potato (*S. tuberosum* ssp. *tuberosum*) developed with all probability in historical times from *S. tuberosum* spp. *andigena*, which was introduced to Europe from the Andean region of South America in the second half of the sixteenth century (Salaman 1985). The cultivated potato is related to a large number of *Solanum* species of Mexico and South America and it can easily be hybridized, particularly to the species of the latter group. For this reason, the approximately 180 wild and 8 cultivated tuber-bearing species of the genus *Solanum* form a valuable gene reservoir for different breeding purposes (Correll 1962; Hawkes 1978; Hawkes and Hjerting 1969, 1989; Ross 1986). Indeed, since the beginning of this century some of them have been successfully used as sources of resistance genes and other traits to improve the cultivated potato (Ross 1986; Foldo 1987), and most modern varieties have in their pedigree one or more wild *Solanum* species (Stegemann and Schnick 1985). The phylogenetic relationships among *Solanum* species are based on biosystematic studies and, more recently, on chemotaxonomy (Desborough and Peloquin 1969; Wietschel and Reznik 1980a, b) and on polymorphisms of chloroplast DNA (Buckner and Hyde 1985; Hosaka et al. 1984; Hosaka 1986; Hosaka and Hanneman 1988a, b). Questions as to the origin of *S. tuberosum* or the phylogenetic relationships of different series within the section Petota are still being discussed (Brücher 1974; Hawkes 1978; Gottschalk 1984; Hosaka 1986). Difficulties arise not only from the capacity of available data to serve as discriminating characters, they are also caused by the methodology adopted. For instance, the "truth" of a phylogenetic relationship cannot be assessed experimentally, and the correct computation of phylogenetic trees is also disputed among schools differing in their view of the evolutionary process (Nei 1987; Felsenstein 1988; Cronquist 1987).

Nuclear RFLPs are a new and very suitable instrument for phylogenetic studies, as demonstrated in a number of species (Song et al. 1988a, b; Havey and Muehlbauer 1989). When restricted to closely related

Table 1. List of species and accessions used for W1 and D1

Series	Species	Accession no.	Ploidy level ($n=12$)	Abbreviation [a]
Etuberosa	*S. etuberosum* Lindl.	28476	2n	etb a
		53007	2n	etb b
Longipedicellata	*S. stoloniferum* Schlecht. et Rche.	007229	4n	sto a
		007230	4n	sto b
Pinnatisecta	*S. pinnatisectum* Dun.	008168	2n	pnt a
Polyadenia	*S. polyadenium* Greenm.	008182	2n	pld a
Demissa	*S. demissum* Lindl.	010022	6n	dms a
Acaulia	*S. acaule* Bitt.	018627	4n	acl a
		016835	4n	acl b
		024555	4n	acl c
Commersoniana	*S. chacoense* Bitt.	008030	2n	chc a
		016979	2n	chc b
Megistacroloba	*S. megistacrolobum* Bitt.	008113	2n	mga a
		027262	2n	mga b
Tuberosa (wild)	*S.* × *berthaultii* Hawkes	010063	2n	ber a
	S. canasense Hawkes	008012	2n	can a
		008110	2n	can b
	S. gourlayi Hawkes	007180	2n	grl a
		018529	2n	grl b
	S. kurtzianum Bitt. et Wittm.	017576	2n	ktz a
		017585	2n	ktz b
	S. sparsipilum Juz et Buk.	015455	2n	spl a
	S. spegazzinii Bitt.	016906	2n	spg a
		016929	2n	spg b
		018326	2n	spg c
	S. vernei Bitt. et Wittm.	015451	2n	vrn a
		017536	2n	vrn b
		EBS 2664	2n	vrn c
Tuberosa (cult.)	*S. stenotomum* Juz. et Buk.	018478	2n	stn a
		027166	2n	stn b
		CIP 701070	2n	stn c
	S. tuberosum ssp. *andigena* Juz. et Buk.	007462	4n	adg a
		024677	4n	adg b
		028078	4n	adg c
	S. tuberosum ssp. *tuberosum* L.	Bintje	4n	BI
		H79.134/44	2n	13
		H80.696/4	2n	40

[a] According to Huaman and Ross (1985)

species (Nei 1987), they may indeed provide a consistent picture of the evolutionary relationships between species. In this work, relationships among a series of *Solanum* species were examined by comparing restriction fragment patterns, using high resolution polyacrylamide gel electrophoresis and RFLP probes developed from *S. tuberosum* spp. *tuberosum* (Gebhardt et al. 1989). Different computational methods were employed and special emphasis was given to tests on the reliability of the results. Agreements and discrepancies on the results with the current phylogeny of *Solanum* species are discussed.

Materials and methods

Plant material

Seeds of 34 accessions of 15 wild and 2 cultivated *Solanum* species were obtained from the germ plasm collection of the Bundesforschungsanstalt für Landwirtschaft (FAL) at Braunschweig, FRG, courtesy of J. R. Hoekstra. The European potato *S. tuberosum* ssp. *tuberosum* was represented by the tuber-propagated variety Bintje (from Scharnhorst, Außenstelle MPI für Züchtungsforschung) and the diploid breeding line 13 (H 79.134/44, MPI für Züchtungsforschung). The 2n breeding line 40 (H 80.696/4, MPI für Züchtungsforschung) is an interspecific hybrid between diploid *S. tuberosum* spp. *tuberosum* and

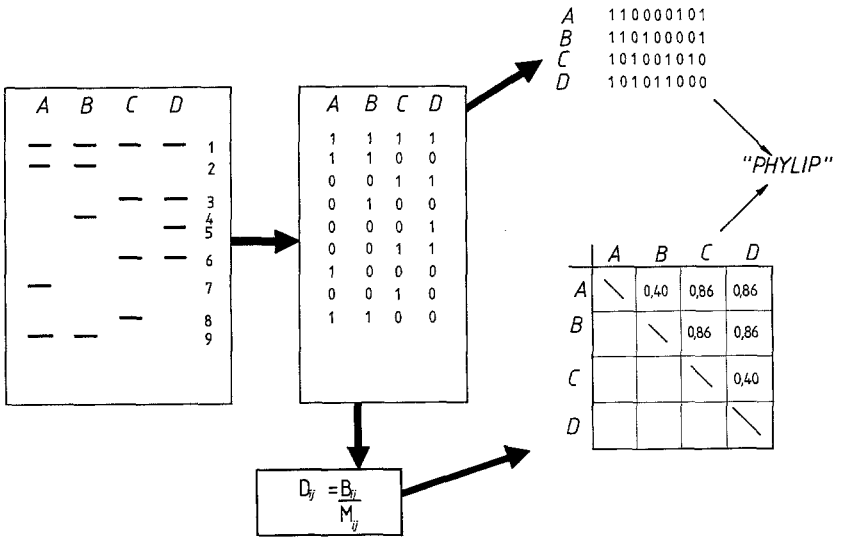

Fig. 1. Scheme for processing fragment positions on autoradiographs into 1–0 matrices, and generating input files for parsimony and distance matrix analysis with the PHYLIP program package. *A*, *B*, *C*, and *D* are individual genotypes; fragment nos. 2, 3, 6, and 9 are "informative", fragment nos. 1, 4, 5, 7, and 8 are "non-informative". D_{ij} = distance between genotypes *i* and j, B_{ij} = number of different fragments between *i* and *j*, M_{ij} = total number of fragment positions in *i* and *j*

Table 2. List of probes used for W1 and D1

W1		D1	
Probe name	Linkage group	Probe name	Linkage group
CP 62	1	CP 11	1
rbcS 2 (a)	2 [a]	PC 116	2
rbcS 2 (b)	2	GP 26	2
rbcS 2 (c)	2	GP 25	3
rbcS 1 (a)	3 [a]	CP 65	4
rbcS 1 (b)	3	WX	7
rbcS 1 (c)	3	CP 18	8
CP 65	4	GP 24	8
GP 21	5	CP 50	8
CP 44	6	CP 61	8
CP 61	8	GP 27	9
GP 27	9	CP 52	9
		CP 66	10
		CP 57	11
		GP 83	11
		GP 41	12
		c0164	N.D. [b]

[a] (a), (b), (c) are three non-overlapping probes from the same genomic region (T. Debener, in preparation)
[b] N.D. – not determined

an unknown accession of *S. spegazzinii*. Names, numbers, ploidy levels, and abbreviations of all accessions used are listed in Table 1. Leaf and shoot material was harvested from plants originated from single seeds and propagated via cuttings or tubers (a progeny derived from a single seed is here referred to as genotype). Plant material was freeze-dried and stored at −70° or −20 °C.

Probes

Anonymous single- or low-copy-number cDNA and genomic sequences of potato and probes corresponding to the potato genes of ribulose-1,5-bisphosphate carboxylase, small subunit (Wolter et al. 1988; Debener et al. in preparation) were used to detect RFLPs. Identification numbers and map positions of probes (Gebhardt et al. 1989) are given in Table 2.

RFLP detection

DNA isolation, restriction digests, gel electrophoresis, blotting, probe preparation, and hybridization were as described by Gebhardt et al. (1989), except that for the data set W1, denaturing polyacrylamide gradient gels (3%–10%) were used, which increased the resolution to about 2 bases difference in fragment size. Three to four microgram of genomic DNA per sample were digested with the restriction enzymes AluI, HaeIII, DraI, DdeI, Sau3A, TaqI, and RsaI, respectively, according to the supplier's instructions (Boehringer, Gibco-BRL).

Analysis of data

The presence or absence of specific restriction fragments was scored on autoradiograms and transformed into a 1 (present) and 0 (absent) matrix over all genotypes and all fragment positions scored (Fig. 1). A computer program written for an Atari 1040 ST generated input files for the PHYLIP program package developed by Felsenstein (1987). PHYLIP was run on a MicroVax 2 under VMS version 4.4. The restriction fragment patterns were computed in two different forms of information (Fig. 1): the 1-0 sequences were used directly as input for the "discrete character state" programs carrying out various forms of the so-called "parsimony analysis" (Felsenstein 1987). Two programs, METRO and PENNY, making use of different parsimony algorithms were applied. The pairwise distances D_{ij} (Fig. 1) between all individual genotypes (=Operational Taxonomic Units or OTU) were calculated using the complement of the Jaquard index (Jackson et al. 1989), and entered into the distance matrix programs FITCH and KITSCH, respectively. The sum of squares (=sq) was provided with the output of FITCH and KITSCH, and displayed the overall differences between the phenogram and the original data matrix (Felsenstein 1987). The average percent standard deviation (=s) was provided as output with the method of Fitch and Margoliash (1967) ($P=2.0$).

Similar to the "jackknife methods" (Felsenstein 1988; Diaconis and Efron 1987), a program was written for a random data resampling procedure. New data sets were made by resampling 50% of the fragment positions from the original matrix in a random order without replacement. The phenograms obtained with the reduced matrices were fit to the phenogram from the

full data set, using option U of the FITCH program. The variability coefficients (=cv, percent standard deviation from the mean value) of the branch lengths, resulting from 8–10 independent runs, gave a relative measure for the reliability of the values obtained.

Results

Two series of experiments formed the data base for the construction of phenograms.

The W1 data set

The first data set W1 included 38 genotypes (OTUs) representing 12 *Solanum* species and one interspecific hybrid (line 40). Each species comprised two to five individual genotypes of one to three accessions. Genomic DNA was restricted with seven four-cutter enzymes, and the restriction fragments were separated on denaturing polyacrylamide gradient gels (3%–10%), in order to maximize the information content per genotype per probe. After transfer of the fragments onto nylon membranes, the filters were hybridized with 12 single- or low-copy-number probes mapping to at least eight different loci. Names and chromosomal location of probes are given in Table 2 (Gebhardt et al. 1989).

Restriction fragment patterns of 70 probe/enzyme combinations (5–7 enzymes per probe) were examined, resulting in 777 fragment positions from which a distance matrix was computed (Fig. 1). Using the program FITCH (options G, J; $P=2.0$), which is based on the method of Fitch and Margoliash (1967), a phenogram was obtained that is shown in Fig. 2. Because the calculations of genetic distances may depend on the algorithm used, on the input order, or on the number of the OTUs, the reliability of our phenograms was assessed using several approaches. First, the program FITCH was run ten times with a different input order of the OTUs. The resulting phenogram was not affected. Second, because it is known that the reliability of the analysis increases with decreasing number of OTUs, three data subsets of W1, comprising distinct branches of the phenogram of Fig. 2, were separately analyzed again with FITCH (option G; $P=2.0$). The three phenograms obtained considering the 16 OTUs of *S. sparsipilum, S. spegazzinii, S. vernei, S. chacoense,* and *S. gourlayi,* the 14 OTUs of *S. canasense, S. stenotomum, S. tuberosum* spp. *andigena,* and Bintje, and the 6 OTUs of *S. tuberosum* spp. *andigena* and Bintje were slightly different in branch length, but were identical in their topology compared with corresponding parts of the tree in Fig. 2. The only exception was the position of *S. stenotomum* b7 (stn b7), which is shown in Fig. 3. Third, different algorithms were applied to the same W1 data set. The method of Cavalli-Sforza and Edwards (1967) (FITCH options G, J; $P=0.0$) and

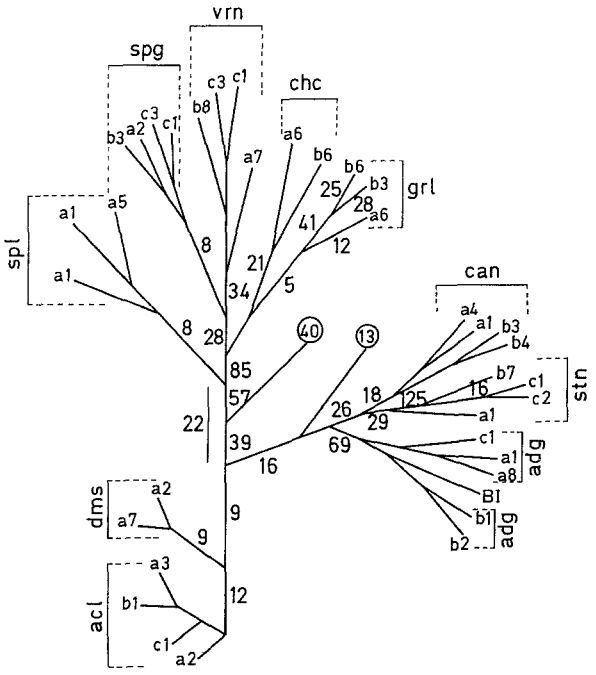

Fig. 2. Unrooted phylogenetic tree for data set W1. 38 OTUs were analyzed with FITCH (modified method of Fitch and Margoliash 1967; options G, J; $P=2.0$). sq=4, 6, s=5, 7. Ten runs were made with different input orders of the OTUs. Species were represented by individually numbered plants of one to three accessions (*a, b, c*). cv values (computed from eight independent runs of the resampling procedure) are given for some internodes and branches, and are partially omitted for reasons of clearity

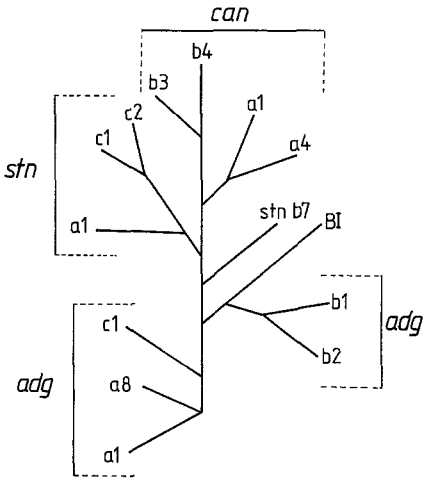

Fig. 3. Unrooted phylogenetic tree for a data subset of W1, containing only 14 OTUs. The data were analyzed with FITCH (options G, J; $P=2.0$). Genotype numbering as in Fig. 2

of an intermediate method (FITCH options G, J; $P=1.0$) resulted in phenograms similar to those of Fig. 2. Slight variations in the positions of *S. acaule* a3, of Bintje (BI) and line 13 within the *andigena-stenotomum* cluster, and of the branching point of *S. sparsipilum* were noticed (data not shown). The program KITSCH (option J;

$P=2.0$ and $P=0.0$), which performs distance matrix methods with the assumption of an evolutionary clock, gave a much higher sum of squares compared to FITCH but, again, the resulting phenogram differed only in the positions of Bintje (BI), lines 13 and 40 (data not shown). The general structure of the W1 phenogram was stable, regardless of the distance matrix method adopted.

Parsimony methods approach the construction of evolutionary trees with different assumptions (Nei 1987; Felsenstein 1987, 1988). The program METRO, carrying out Wagner parsimony (38 OTUs, 777 characters, annealing parameter = 0.001), found one most parsimonious tree with a length of 1809 steps, differing from the topology in Fig. 2 only in the location of one *S. acaule* genotype (acl a3), line 13, and the *S. sparsipilum* (spl) branch within the near neighbourhood (data not shown). Using only the informative 606 fragment positions gave similar results (fragments present in all genotypes and fragments present in one genotype only are considered non-informative; Fig. 1). Again, the main structure of the tree remained unaltered.

Fourth, by resampling at random 50% of the fragment positions from the complete W1 data set, eight new data sets were created. The topologies obtained with the reduced data sets and the FITCH program were fit to the topology determined with the full data set. The resulting variability coefficients (cv) are indicated in Fig. 2 for some of the internodes determining the structure of the phenogram. The internodes separating the main branches and the species *S. acaule* (acl), *S. demissum* (dms), *S. sparsipilum* (spl), *S. spegazzinii* (spg), *S. vernei* (vrn), *S. chacoense* (chc), and *S. gourlayi* (grl) appear more reliable (i.e., have lower cv values) than the internodes within the *S. tuberosum* spp. *andigena* (adg), BI, *S. stenotomum* (stn), *S. canasense* (can) species cluster and within the species (e.g., grl, stn). Comparatively reliable also were the end branches, with an average cv of 16%. The observed variability in the positioning of the interspecific hybrid 40 and of the branch point of *S. sparsipilum* (spl) was also reflected in the high cv values of the attached internodes.

Having assessed the consistency of the phenogram for the W1 data set, the following conclusions were drawn. The intraspecific variability was large compared to the interspecific variability, as indicated by fairly similar branch lengths within a species and between species. Nevertheless, all individuals belonging to the same species were clustered before they joined the clusters of other species, and all individuals of the same accession clustered before they joined the other individuals of the same species (with the exception of *S. acaule* accession a). Species borders and – less reliably – accession borders were, therefore, reflected in the phenogram. In the phenogram there are three distinct branches. A closely related cluster is formed by the cultivated potatoes *S. tuberosum* ssp. *andigena* and ssp. *tuberosum* (BI, 13), *S. stenotomum*, and the wild species *S. canasense* (series Tuberosa).

The wild species *S. sparsipilum, spegazzinii, vernei, gourlayi* of the series Tuberosa cluster with *S. chacoense* of the series Commersoniana. *S. demissum* (series Demissa) and *S. acaule* (series Acaulia) form a third branch, with *S. demissum* (dms) being closer to the Tuberosa than *S. acaule* (acl). The interspecific hybrid between *S. spegazzinii* and *S. tuberosum* spp. *tuberosum* (line 40) falls clearly between the clusters of the parental species.

The D1 data set

The second data set D1 was obtained from 34 genotypes of 18 *Solanum* species, including the 12 species of data set W1. Each species was represented by two individual genotypes of one or two accessions, with the exception of *S. tuberosum* spp. *tuberosum* (represented only by Bintje) and *S. etuberosum* with only one genotype. The latter was justified because *S. etuberosum* was the only species in which extremely small RFLP differences between genotypes were observed. Restriction fragments were separated on continuous 4% denaturing polyacrylamide gels. Twenty-four probe/enzyme combinations (one or two enzymes per probe) from 17 single- or low-copy-number probes mapping to at least 17 loci (Table 2) provided a matrix of 267 fragment positions. Compared to W1, the number of genotypes per species and the information content per probe was therefore reduced, but the number of species, probes, and loci was increased. W1 and D1 had only three probe/enzyme combinations and five genotypes in common. The W1 and D1 data sets for the five common genotypes were analyzed with FITCH and KITSCH (options G, J; $P=2.0$). The resulting phenograms had identical topologies with minor differences in branch lengths. The FITCH phenograms obtained for the five genotypes common to W1 and D1 are shown in Fig. 4. Parsimony analysis with METRO (annealing parameter = 0.001) and PENNY (using a different search algorithm than METRO) gave identical tree topologies (data not shown). The similarity of W1 and D1 presented in Fig. 4 demonstrated that largely independent experimental RFLP data sets lead to the same results when phenograms were computed.

The phenogram shown in Fig. 5 is the result of the analysis with FITCH (34 OTUs, options G, J, O; $P=2.0$) using the complete D1 data set. Since *S. etuberosum* is known to be most distantly related to the other species in the set (Hawkes 1978; Hosaka 1986), it was used as an outgroup (option 0).

As a consequence of the reduced amount of information in D1 compared to W1, the variability of the topologies obtained for D1 with different methods was higher. This was also reflected in higher cv values obtained in the

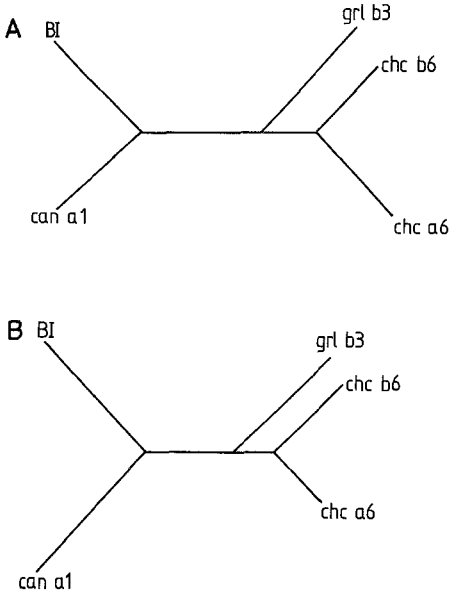

Fig. 4 A and B. Unrooted phylogenetic trees for five genotypes (OTUs) that were common to both data sets **A** W1 and **B** D1. With the exception of three probe/enzyme combinations, both data sets were based on different RFLP markers. Analysis was with FITCH (options G, J; $P=2.0$). Genotypes are numbered as in Fig. 2

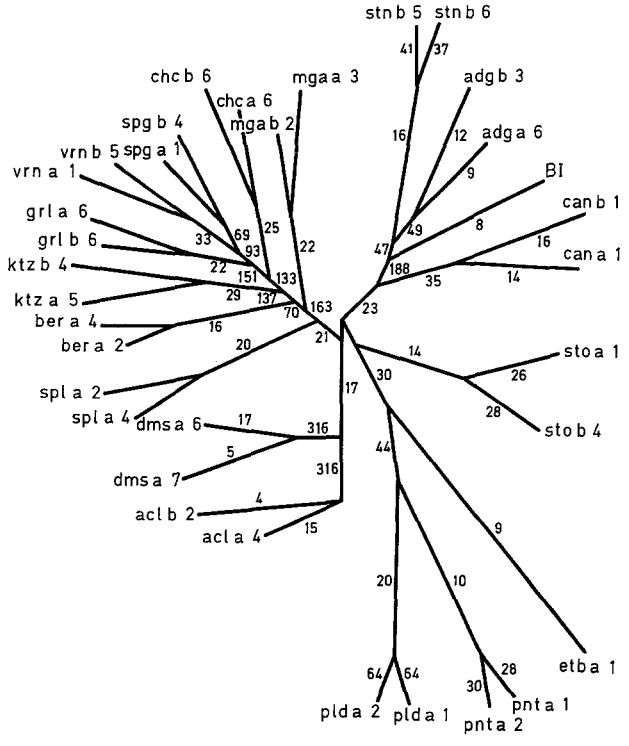

Fig. 5. Unrooted phylogenetic tree for the data set D1. 34 OTUs were analyzed with FITCH (options G, F; $P=2.0$, ten runs with different input order). sq = 3.98, s = 5.96. cv values of branch lengths were computed from ten independent runs of the resampling procedure, and are given only in part for reasons of clarity. Genotypes are numbered as in Fig. 2

resampling experiment, which were adopted as relative estimates for the reliability of the branch length values. Particularly the internodes separating the split points of different species within the main clusters showed large variation (e.g., the cv was 163% between *S. sparsipilum* and *S. megistacrolobum*), whereas the internodes leading to the end branches were relatively stable (e.g., cv = 9%, 10%, and 20% for *S. etuberosum*, *S. pinnatisectum*, and *S. polyadenium*, respectively). Changing the input order of OTUs resulted in two phenograms differing slightly in the average percent standard deviation and topologies. Figure 5 shows the phenogram with the smaller s and the topology that was more similar to the W1 phenogram of Fig. 2.

In the phenogram obtained by the method of Cavalli-Sforza and Edwards (1967) (FITCH options G, J; $P=0.0$), the positions of *S. chacoense* (chc) and *S.kurtzianum* (ktz) in Fig. 5 were exchanged. Parsimony analysis with METRO (34 OTUs, 267 characters, annealing parameter = 0.001) produced 11 similar most parsimonious trees, having the main clusters as shown in Fig 5 but differing in the relative positions of Bintje (BI), *S. stenotomum* (stn), *S. megistacrolobum* (mga), *S. chacoense* (chc) and, in particular, of *S. sparsipilum* (spl) and *S. stoloniferum* (sto). In some cases, *S. sparsipilum* (spl) fell in the group of the cultivated Tuberosa and *S. stoloniferum* (sto) clustered with the wild Tuberosa. Reducing the data set from the total 267 to 180 "informative" characters gave no significant alteration. when the D1 and the W1 phenograms were analyzed, it was noticed that the main clusters of the W1 phenogram reappear in the D1 phenogram. The additional wild *Solanum* species included in D1, *S. berthaultii* and *S. kurtzianum* and, remarkably, *S. megistacrolobum* (belonging to the series Megistacroloba, according to the classical taxonomy) (Hawkes 1978), fell into the *sparsipilum, spegazzinii, vernei, chacoense, gourlayi* group, although the relationships between some of the species within the group were unreliable based on their high cvs. *S. stoloniferum, pinnatisectum, polyadenium,* and *etuberosum* formed an additional cluster.

Discussion

According to Nei (1987), about 10^{44} possible tree topologies exist for 34 genotypes or OTUs. The problem of finding from this enormous number those which fit best the experimental data can be approached with different algorithms based on different evolutionary theories (Nei 1987; Felsenstein 1988). The analysis of two largely independent RFLP data set with several distance matrix and parsimony methods resulted in very few and similar topologies, differing mainly in replacements within the near neighbourhood. Particularly obvious was the consistency

among methods in the W1 data set, despite the fact that one requirement of parsimony analysis – the independence of character states, in our case fragment positions – was not fulfilled for an unknown fraction of the 777 fragment positions. Due to the experimental design, equivalent restriction fragments were detected with the same probe and different restriction enzymes which, therefore, cannot be considered independent. Differences in results caused by the input order of OTUs or by their large number also had no significant influence on the topologies, nor did the choice of genotypes representing a species or the choice of the set of probes.

Differences, however, were noted in the reliability of branch and internode lengths within and between the W1 and D1 topologies, as quantified by the resampling procedure. The W1 phenogram, constructed with the more informative data set of 777 fragment positions, was more reliable than D1, which had only 267 fragment positions. Indeed, the 11 minimal trees obtained with parsimony analysis for D1 showed stronger deviations among each other. Within the W1 and D1 phenograms the species were always clearly separated, although the intraspecific variability, even within accessions, was high compared to the interspecific variability (with the exception of *S. etuberosum*). The level of resolution obtained using nuclear RFLPs was in any case much higher for closely related species than in chloroplast DNA polymorphisms (Hosaka et al. 1984; Hosaka 1986; Hosaka et al. 1988; Hosaka and Hanneman 1988 a, b). These were restricted to a few differences between close relatives and, in some cases, failed to resolve clearly distinct species.

The internodes separating the major clusters in the W1 and D1 phenograms were much more reliable than the internodes determining the split points within the clusters. A comparison concerning the reliability of tree topologies obtained by RFLP studies in other plant species (Song et al. 1988 a, b) is, however, difficult because in the cited studies, the phenograms were based on a single computational method and did not include indices of reliability. The tree topologies derived from nuclear RFLP analysis were compared to the system of classification based on biosystematic studies. Hawkes (1989) grouped the species of the subsections Potatoe and Estolonifera into 21 series according to their putative relatedness, inferred from a wide range of morphological data (Table 1). Our phenograms correlated with this system in that all species of the series Tuberosa were clustered in two related groups, with the wild species *S. spegazzinii* (spg), *S. vernei* (vrn), *S. sparsipilum* (spl), *S. gourlayi* (grl), *S. berthaultii* (ber), and *S. kurtzianum* (ktz) in one, and *S. canasense* (can), *S. stenotomum* (stn), *S. tuberosum* spp. *andigena* (adg), Bintje, and line 13 in the other. The latter group supports the theory that the wild *S. canasense* is one ancestor of *S. stenotomum* which, in turn, is most likely one of the ancestors of *S. andigena* (Bukasov 1966; Hawkes 1978, 1990; Hosaka 1986). The cultivated European potato was represented in our studies by the old variety Bintje, which was released well before potato breeders started to intentionally introgress wild germ plasm in cultivated potatoes (Ross 1986; Stegemann and Schnick 1985), and by the diploid line H 79.134/44 (13), in the pedigree of which a certain amount of *andigena, demissum, acaule* and possibly other wild germ plasm was introgressed early in the pedigree (Ross 1986; H. Uhrig, personal communication). These genotypes belong to the same species as *S. tuberosum* spp. *andigena* and, therefore, were expected to cluster closely with the *andigena, stenotomum, canasense* group. The interspecific nature of line 40 was clearly recognized by our phenograms, which placed it in an intermediary position between the parental clusters (Fig. 2).

The other series of Hawkes were represented in our study by one species each. *S. demissum* and *S. acaule* of the series Demissa and Acaulia, respectively, formed a well-separated branch, with *S. demissum* being closer to the Tuberosa than *S. acaule*. This is in agreement with Hawkes (1972, 1978, 1989) and Correll (1962). *S. stoloniferum, polyadenium, pinnatisectum,* and *etuberosum* were clearly separated from each other and from the Tuberosa branches. *S. etuberosum,* the species morphologically most distinct from all others, was located at the most distant point in the phenogram, which also agrees with results from chemotaxonomical studies (Wietschel and Reznik 1980 a). The position of *S. etuberosum* is also well correlated to the recent separation of the series Etuberosa from the subsection Potatoe into the new subsection Estolonifera (Hawkes 1989).

The same topology for *S. etuberosum, S. pinnatisectum, S. polyadenium,* and *S. stoloniferum* was obtained by Hosaka et al. (1984). In contrast to earlier classifications, RFLP data assigned *S. chacoense* (formerly series Commersoniana) to the wild Tuberosa cluster. This position of *S. chacoense* was very stable in all RFLP phenograms computed by us so far. Hawkes and Hjerting (1989) mentioned that *S. chacoense* and the close relative, *S. tarijense,* were somehow distinct from the rest of the series Commersoniana, and they recently placed them into the series Yungasensa (Hawkes 1989, 1990). It is furthermore hypothesized that *S. chacoense* is a possible ancestral species to a group of series called the primitive Rotata, which comprises the series Megistacroloba, Cuneolata, and parts of the series Tuberosa and Conicibaccata. This is in agreement with the positions of both *S. chacoense* and *S. megistacrolobum* in the RFLP-derived phenograms. As *S. chacoense* is an extremely polymorphic species with a wide ecological distribution (Hawkes and Hjerting 1989), the analysis of more accessions might help in finding a more precise taxonomic position for this species.

It is also worth mentioning that in our phenograms the distance between the two Tuberosa branches was comparable to the distances between Tuberosa and other series and was, therefore, longer than one would expect from the current classification. An interesting observation concerned the two Mexican species, *S. pinnatisectum* and *S. polyadenium,* which clustered with the Chilean species *S. etuberosum* before they joined *S. demissum,* another Mexican species. The geographical distribution of species was therefore not reflected in the phenograms.

In summarizing this work, we have shown that nuclear RFLPs provide an unlimited source of characters discriminating individual genotypes, accessions, subspecies, and species of the genus *Solanum*. As such, they should be considered as an excellent tool for studying population genetics, taxonomy, and evolution of the potato. The analysis of genetic distances presented here is an example. The reliability of the results could be increased by enlarging the data sets, although this might be limited by the computer capacity. Even then the phylogenetic relationships within closely related clusters might not be resolved without a certain degree of ambiguity. Asking more specific phylogenetic questions – restricting, e.g., the analysis to species within a cluster, and including several accessions per species and more conserved RFLP markers – seems a promising strategy for further studies of evolution within the genus *Solanum*.

Acknowledgements. The authors would like to thank Prof. H. Ross, Dr. H. Uhrig, and Dr. P. Degens for helpful discussions, Prof. J. G. Hawkes for providing us with preprints of his manuscripts, B. Walkemeier for technical assistance, and M. Pasemann for preparing the manuscript. This work was supported by the Bundesministerium für Forschung und Technologie (BMFT) under Project No. BCT 03902-1.06.

References

Brücher H (1974) Über Art-Begriffe und Art-Bildung bei *Solanum* (Sect. Tuberarium). Beitr Biol Pflanz 50:393–429

Buckner B, Hyde BB (1985) Chloroplast DNA variation between the common cultivated potato (*Solanum tuberosum* ssp. *tuberosum*) and several South American relatives. Theor Appl Genet 71:527–531

Bukasov SM (1966) Die Kulturkartoffel und ihre wildwachsenden Vorfahren. Z Pflanzenzuecht 55:139–164

Cavalli-Sforza L, Edwards L (1967) Phylogenetic analysis: models and estimation procedures. Evolution 32:550–570

Correll DS (1962) The potato and its wild relatives. Texas Research Foundation, Renner/Tx

Cronquist A (1987) A botanical critique of cladism. Bot Rev 53:1–52

Desborough S, Peloquin SJ (1969) Disc electrophoresis of tuber proteins from *Solanum* species. Phytochemistry 5:727–733

Diaconis A, Efron BM (1987) Computer intensive methods in statistics. Sci Am 248:116–130

Felsenstein J (1987) PHYLIP (Phylogeny Inference Package) version 3.0 manual. University of Washington Press, Seattle/WA

Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. Annu Rev Genet 22:521–565

Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155:279–284

Foldo NE (1987) Genetic resources: their preservation and utilization. In: Jellis GJ, Richardson DE (eds) The production of potato varieties: technical advances. Cambridge University Press, Cambridge, pp 10–27

Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. Theor Appl Genet 78:65–75

Gottschalk W (1984) The origin of the potato: an open problem. Nucleus 27:37–44

Havey MJ, Muehlbauer FJ (1989) Variability for restriction lengths and phylogenies in lentil. Theor Appl Genet 77:839–843

Hawkes JG (1972) Evolutionary relationships in wild tuber-bearing *Solanum* species. Symp Biol Hung 12:65–69

Hawkes JG (1978) Biosystematics of the potato. In: Harris PM (ed) The potato crop. Chapman and Hall, London, pp 15–69

Hawkes JG (1989) Nomenclatural and taxonomic notes on the infrageneric taxa of the tuber-bearing *Solanums* (*Solanaceae*). Taxon 38:489–492

Hawkes JG (1990) Evolution of the cultivated potatoes and their wild tuber-bearing relatives. Kulturpflanz (in press)

Hawkes JG, Hjerting JP (1969) The potatoes of Argentina, Brazil, Paraguay and Uruguay. Clarendon Press, Oxford

Hawkes JG, Hjerting JP (1989) The potatoes of Bolivia. Clarendon Press, Oxford

Hosaka K (1986) Who is the mother of the potato? – restriction endonuclease analysis of chloroplast DNA of cultivated potatoes. Theor Appl Genet 72:606–618

Hosaka K, Hanneman RE (1988a) The origin of the cultivated tetraploid potato based on chloroplast DNA. Theor Appl Genet 76:172–176

Hosaka K, Hanneman RE (1988b) Origin of chloroplast DNA diversity in the Andean potatoes. Theor Appl Genet 76:333–340

Hosaka K, Ogihara U, Matsubayashi M, Tsunewaki K (1984) Phylogenetic relationships between the tuberous *Solanum* species as revealed by restriction endonuclease analysis of chloroplast DNA. Jpn J Genet 59:349–369

Hosaka K, Zoeten GA de, Hanneman RE (1988) Cultivated potato chloroplast DNA differs from the wild type by one deletion: evidence and implications. Theor Appl Genet 75:741–745

Huaman Z, Ross RW (1985) Updated listing of potato species, names, abbreviations and taxonomy status. Am Potato J 62:629–641

Jackson DA, Somers KM, Harvey HH (1989) Similarity coefficients: measures of co-occurrence and association or simply measures of occurrence? Am Nat 133:436–453

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Ross H (1986) Potato breeding: problems and perspectives. Plant Breeding No. 37. Parey, Berlin Hamburg

Salaman RN (1985) The potatoes of America and their relation to the early european varieties. In: Hawkes JG (ed) The history and social influence of the potato. Revised impression. Cambridge University Press, pp 51–72

Song KM, Osborn TC, Williams PH (1988a) Brassica taxonomy based on nucelar restriction fragment length polymorphisms (RFLPs). 1. Genome evolution of diploid and amphidiploid species. Theor Appl Genet 75:784–794

Song KM, Osborn TC, Williams PH (1988b) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 2. Preliminary analysis of subspecies within *B. rapa* (syn. *campestris*) and *B. oleracea*. Theor Appl Genet 76:593–600

Stegemann H, Schnick D (1985) Index europäischer Kartoffelsorten. Mitt Biol Bundesanst Land- und Forstwirtsch Berlin-Dahlem 227

Wietschel G, Reznik H (1980a) Die Flavonoidmuster der knollentragenden *Solanum*-Arten. 2. Die Flavonoid-Glycoside der Arten aus Series I–XVI. Z Pflanzenphysiol 99:79–88

Wietschel G, Reznik H (1980b) Die Flavonoidmuster der knollentragenden *Solanum*-Arten. 3. Die Flavonoid-Glycoside der Arten aus Series XVII. Z Pflanzenphysiol 99:149–158

Wolter FP, Fritz CC, Willmitzer L, Schell J, Schreier PH (1988) rbcS genes in *Solanum tuberosum:* Conservation of transit peptide and exon shuffling during evolution. Proc Natl Acad Sci USA 85:846–850