Theor Appl Genet (2002) 104:146–156

C.E. McGregor · R. van Treuren · R. Hoekstra
Th.J.L. van Hintum

Analysis of the wild potato germplasm of the series *Acaulia* with AFLPs: implications for *ex situ* conservation

Received: 6 November 2000 / Accepted: 20 April 2001

Abstract The wild potato germplasm of the series *Acaulia* maintained at the Centre for Genetic Resources, The Netherlands, currently consists of 314 accessions. This collection comprises seed samples of the species *Solanum acaule* (ssp. *acaule*, ssp. *aemulans*, ssp. *palmirense* and ssp. *punae*) and *Solanum albicans* collected from South America. In order to validate taxonomic classification, to investigate the extent of redundancy and to study the distribution of genetic diversity across the collection area, the entire collection was analysed with two AFLP primer pairs on two plants per accession. Within the entire sample a total number of 130 polymorphic bands were scored for the two primer pairs. An UPGMA cluster analysis grouped the majority of plants according to the species and subspecies. A total number of 16 misclassifications were identified, including four cases that did not seem to belong to the series *Acaulia*. Two accessions were found to consist of plants of different AFLP clusters. AFLP data also allowed the taxonomic classification of the subspecies of 97 accessions that previously were described as *S. acaule* only. For 126 accessions the two individuals studied displayed identical AFLP profiles. Forty six of these 126 accessions shared their profiles with both or single plants of other accessions. These were all tested for identical profiles for a third primer pair, resulting in 15 duplication groups consisting of a total number of 22 accessions and 14 single plants. Analyses of molecular variance (AMOVA) were performed to examine the distribution of genetic variation. Comparison of geographic distances between the collection site of plants and the number of AFLP polymorphisms revealed no consistent relationship between geographic distance and genetic diversity. AFLP analysis appeared to be an efficient method to verify taxonomic classification and to identify redundancies in the wild germplasm of the series *Acaulia*. Implications of the results for the *ex situ* conservation of wild potato germplasm are discussed.

Keywords AFLP[1] · Collection management · Genetic resources · Potato · Redundancy · Taxonomy

Communicated by B.S. Gill

C.E. McGregor · R. van Treuren (✉) · R. Hoekstra
Th.J.L. van Hintum
Centre for Genetic Resources, The Netherlands (CGN),
Plant Research International B.V.,
Wageningen University and Research Centre, P.O. Box 16,
6700 AA Wageningen, The Netherlands
e-mail: R.vanTreuren@plant.wag-ur.nl
Tel.: +31-317-477078, Fax: +31-317-418094

Introduction

Solanum acaule Bitter is one of the most widely distributed wild potato species. Its habitat is the high, cold grasslands in the Puna, at an altitude of 2,600 to 4,600 m, and stretches from Ecuador to north-west Argentina (Ochoa 1990; Hosaka and Spooner 1992; Kardolus 1998). Although taxonomic treatment of the series *Acaulia* has gone through several changes (Brücher 1959; Correll 1962; Hawkes 1963 1978 1990) it is now generally accepted that it consists of two species: the hexaploid (2n=72) *Solanum albicans* (Ochoa 1983) and the tetraploid (2n=48) *S. acaule* (Ochoa 1990; Hosaka and Spooner 1992). Within *S. acaule*, three subspecies are recognised, ssp. *acaule*, ssp. *aemulans* and ssp. *punae* (Kardolus 1998; Hawkes 1990), although Ochoa (1990) and Hosaka and Spooner (1992) suggested that ssp. *punae* is synonymous with ssp. *acaule*. Using RFLP markers, Hosaka and Spooner (1992) found that the ssp. *aemulans* separated into two genetically different groups, one from the province Jujuy (J-*aemulans*) and one from the province La Rioja (L-*aemulans*) in Argentina. Kardolus (1998) recently described an additional *S. acaule* subspecies, ssp. *palmirense* (2n=72), collected in the province Chimborazo in Ecuador, which had a high DNA homology with the tetraploid ssp. *punae* (2n=48).

S. acaule is important for potato breeding because of its resistance to potato virus X, potato leaf roll virus,

[1] AFLP™ is a trademark of Keygene N.V. (Wageningen, The Netherlands)

Potato Spindle Tuber Viroid (PSTV) and cyst nematodes (Ross 1986), and because of its tolerance to frost, heat and drought. Two cultivated potato species have *S. acaule* in their ancestry. *Solanum juzepczukii*, a highly frost-tolerant triploid, is a natural hybrid of *S. acaule* and *Solanum stenotomum*, while *Solanum curtilobum*, a frost-tolerant pentaploid, is a hybrid of *Solanum tuberosum* and *S. juzepczukii* (Hawkes 1990). These two cultivated species taste bitter, because of glycoalkaloids and are generally used in the Andes region to make chuño, a freeze-dried product. *S. acaule* is one of the wild species that are frequently incorporated into European cultivars (Hawkes 1990). Large amounts of *S. acaule* germplasm have been collected and are kept in genebanks around the world. Currently *S. acaule* makes up 11% of the wild potato accessions in world potato genebanks (Huamán et al. 2000). The Dutch-German potato collection at the Centre for Genetic Resources, The Netherlands (CGN), contains about 2,000 accessions of 110 wild species, including 314 accessions of the series *Acaulia* (Centre for Genetic Resources, The Netherlands 2000).

Maintaining and evaluating germplasm is expensive and there are limits to the number of samples that can be handled effectively (Marshall and Brown 1975). This is even worse in the case of potato germplasm because of the costly virus tests that are required prior to regeneration. Germplasm collections invariably contain duplicate accessions, which may have arisen due to a variety of causes. Populations of wild potato species are frequently collected across small geographic distances, and sampling missions by different collectors are often directed to identical areas (e.g. type localities). Because the distribution of genetic variation in the natural areas is often unknown, collections may contain duplicate accessions. Eliminating these duplicates will reduce the size of the collection, with obvious advantages, not only for genebank curators, but also for users. Traditionally, genebank accessions have been described using passport data and morphological characteristics. However, passport data are often unreliable (Hintum et al. 1996) and environmental conditions may influence morphological characteristics. Characterisation is made even more difficult when dealing with a series such as *Acaulia* where not many qualitative characters can be scored (Kardolus 1998).

However, molecular markers can nowadays be applied to assist characterisation of the germplasm and the identification of redundancies (Kresovich et al. 1994; Virk et al. 1995; Ford-Lloyd et al. 1997; Verma et al. 1999). Obviously, the only way to determine with absolute certainty whether two individuals are identical is to compare all their heritable material. For practical reasons this is impossible, but potential redundancies can be identified by high-resolution molecular-marker analysis. Various molecular-marker methods are available nowadays, each with their own advantages and disadvantages. Since redundant germplasm can be identified with higher certainty as resolution increases (Virk et al. 1995), the ideal system should be capable of investigating large numbers of genomic fragments, and thus have a high information content (McGregor et al. 2000; Powell et al. 1996). The AFLP technique developed by Vos et al. (1995) is such a technique and generally produces between 50 and 100 scorable fragments per PCR reaction. This technique has been widely used in various crops, including cultivated potato (Powell et al. 1996; Milbourne et al. 1997; McGregor et al. 2000) and its wild relatives (Kardolus et al. 1998).

In the present study, AFLPs have been used to characterise the CGN's entire collection of the series *Acaulia*. Results are evaluated in relation to: (1) taxonomic classification, (2) the extent of redundancy, and (3) the distribution of genetic diversity across the collection area. Implications for *ex situ* conservation are also discussed.

Material and methods

Plant material and DNA extraction

The 314 accessions of the series *Acaulia* of the CGN's potato collection are summarised in Table 1, and their approximate collection sites are presented in Fig. 1. More detailed information is pro-

Fig. 1 Geographic distribution of the collection sites of the examined accessions of the series *Acaulia* in South America. *alb=S. albicans*, *pal=S. acaule* ssp. *palmirense*, *pne=S. acaule* ssp. *punae*. The two main regions in Argentina where *S. acaule* ssp. *aemulans* was collected are indicated by J-aem (province Jujuy) and *L-aem* (province La Rioja) Collection sites that are not encircled belong to *S. acaule* (subspecies unknown) and *S. acaule* ssp. *acaule*. Cluster B represents the provinces Ayopaya, Cercado, Chapare and Quillacollo of the department Chochamba where ten accessions described as *S. acaule* or *S. acaule* ssp. *acaule* were collected

Table 1 Summary table of the 314 examined accessions of the series *Acaulia* of the Dutch-German potato collection maintained at CGN. Accessions are grouped by taxonomic status, as described prior to the present study. Country of origin and collection area (dept.=department, prov.=province) are denoted, together with the number of accessions involved

Country	Collection area	Taxonomic status						
		S. acaule	*S. acaule* ssp. *acaule*	*S. acaule* ssp. *aemulans*	*S. acaule* ssp. *palmirense*	*S. acaule* ssp. *punae*	*S. albicans*	Total number of accessions
Argentina	Prov. Jujuy	6	93	8				107
Argentina	Prov. La Rioja	2	1	9				12
Argentina	Prov. Salta	7	39	2				48
Argentina	Prov. Tucumán		7					7
Argentina	Unknown	3	2					5
Bolivia	Dept. Chuquisaca		1					1
Bolivia	Dept. Cochabamba	6	8					14
Bolivia	Dept. La Paz	29	10					39
Bolivia	Dept. Oruro	14	1					15
Bolivia	Dept. Potosí	16	5					21
Bolivia	Dept. Tarija	2						2
Bolivia	Unknown		2					2
Ecuador	Prov. Chimborazo				1			1
Peru	Dept. Ancash						4	4
Peru	Dept. Arequipa		1					1
Peru	Dept. Ayacucho					1		1
Peru	Dept. Cajamarca						4	4
Peru	Dept. Cusco	1	1					2
Peru	Dept. Huancavelica					1		1
Peru	Dept. Huánuco					1		1
Peru	Dept. Junín					1		1
Peru	Dept. La Libertád						1	1
Peru	Dept. Lima	1	1				1	3
Peru	Dept. Puno	1	6					7
Peru	Unknown	3	3					6
Unknown	Unknown	7	1					8
Total		98	182	19	1	4	10	314

vided by CGN (2000). Throughout the paper accessions are denoted by CGN number. In case accessions have not been given a CGN number yet, the previous genebank number, the BGRC (Braunschweig Genetic Resources Center) number, is being used. Leaf material from two plants of each of the 314 accessions were harvested from seedlings grown in a glasshouse, except for a single accession (CGN 20672) that due to poor germination yielded only one plant. Samples were stored at –70°C, freeze dried, and ground into a powder using glassbeads and a retch-shaking mill. Total genomic DNA was extracted using the CTAB method, as described by Fulton et al. (1995). DNA concentrations were estimated using agarose-gel electrophoresis and lambda DNA as a standard.

AFLP analysis

The AFLP procedure (Vos et al. 1995) was carried out as described by Arens et al. (1998) with slight modifications. Briefly, total genomic DNA (300–400 ng) was digested with *Eco*RI and *Mse*I, followed by ligation of the adapters. Pre-amplification was performed using a single adenine (A) selective nucleotide for each primer. For the selective amplification, an *Eco*RI primer, with two selective nucleotides, was labelled with ^{33}P and used in combination with a *Mse*I primer with 3 selective nucleotides. For both the pre-amplification and selective amplification the following amplification profile was used: an initial cycle of 94°C for 30 s, 65°C for 30 s, 72°C for 1 min, followed by 12 touchdown cycles in which the annealing temperature was reduced by 0.7°C per cycle. The annealing temperature was then kept constant at 56°C for the remaining 23 cycles. Amplification products were separated on a 6% polyacrylamide gel, and visualised by exposing an X-ray film to the dried gel.

Twenty three primer combinations were tested for their ability to produce reproducible AFLP profiles that could be scored unambiguously. Reproducibility of the primer combinations was tested by comparing the AFLP profiles of four tissue samples collected from the same individual. The primer combinations E+AC/M+AGC and E+AG/M+ACG were selected for the screening of all plants in the present study. Individuals with identical AFLP profiles for these two primer combinations were additionally screened for primer combination E+AA/M+ACC. This combination was specifically chosen for its ability to generate a large number of bands in order to increase resolution for the identification of possibly identical plants.

Data analysis

Two individuals (one from each of the accessions BGRC 18546 and BGRC 24549) displayed very faint AFLP profiles on the autoradiogram and were disregarded from analysis, thereby reducing the total sample to 625 plants. Polymorphic bands generated by primer combinations E+AC/M+AGC and E+AG/M+ACG were scored as present (1) or absent (0). No monomorphic bands were scored. Pair-wise genetic distances between plants were calculated based on the coefficient of Dice (1945) and graphically represented by a dendrogram using the UPGMA (unweighted pair-group method, arithmetic average) clustering algorithm. These analyses were carried out with the NTSYS-pc package (Rolf 1993).

Plants with identical profiles for primer combinations E+AC/M+AGC and E+AG/M+ACG were also screened for primer combination E+AA/M+ACC. These profiles were simply scored as identical (+), or not identical (–), between plants. With the exception of a single band that proved to be irreproducible when tested

on four replicate samples of the same individual, plants were considered identical when no band polymorphisms could be detected.

In order to determine whether genetic variation differs significantly among geographic regions, analyses of molecular variance (AMOVA) were carried out using genetic distances based on the simple matching coefficient. Per sub(species) molecular variance components were computed at different hierarchical levels, i.e. among provinces, among departments within provinces, and within departments. Variance components were tested for significance by a non-parametric re-sampling approach using 1,000 permuted data sets. AMOVA analyses were performed with version 1.55 of the software package WINAMOVA (Excoffier et al. 1992). In addition, geographical distances using recorded longitudes and latitudes were determined between all collection sites on a (sub)species basis. Correlation coefficients between geographical distances and genetic distances based on Dice coefficient were then calculated using the Genstat 5 software package (release 4.1). Frequency distributions of the number of AFLP polymorphisms between plants for different classes of geographical distance were constructed using a tailor-made VisualBasic program.

Results

AFLP polymorphism

Among the 314 accessions examined, corresponding to 625 plants, a total number of 130 polymorphic bands was scored, 50 for primer combination E+AC/M+AGC and 80 for combination E+AG/M+ACG. An example of the autoradiogram of primer pair E+AC/M+AGC is shown in Fig. 2. *S. acaule* (Groups A to F) and *S. albicans* (Group G) could clearly be distinguished based on their AFLP profiles. A total number of 11 discriminative bands were observed between these species. Within *S. acaule*, three discriminative bands were found for ssp. *palmirense*, but none for the other subspecies. However, discriminative power substantially increased by examining sets of AFLP bands. These results indicated the power of AFLPs to discriminate different taxonomic levels among wild potato germplasm of the series *Acaulia* that are more difficult to distinguish morphologically (see also Fig. 2).

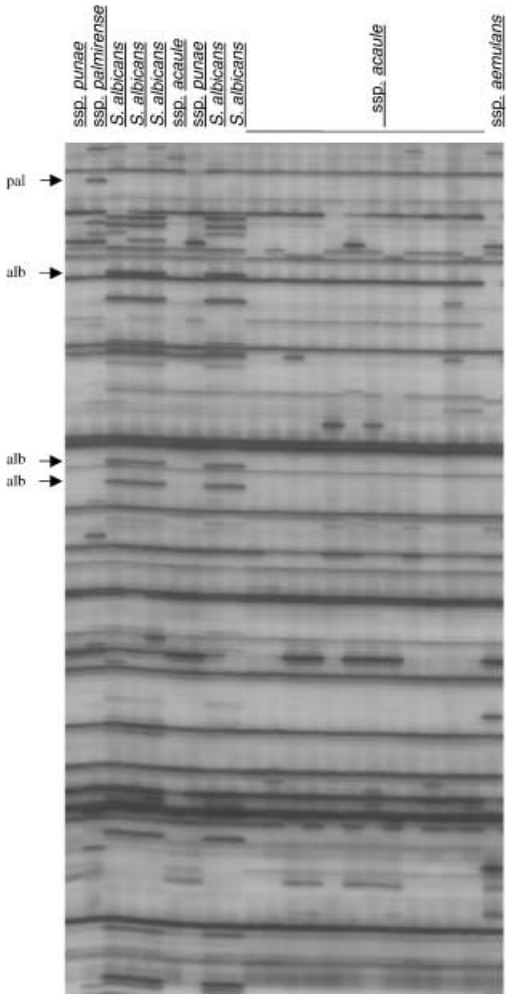


Fig. 2 Part of an autoradiogram of AFLP fragments based on primer pair E+AG/M+ACG, showing the profiles of the different species and subspecies (denoted at the top) of the series *Acaulia*. A number of clear, discriminative bands between (sub)species are indicated by *arrows* at the left side of the autoradiogram (*alb=S. albicans*, *pal=S. acaule* ssp. *palmirense*)

Taxonomic classification

The UPMGA clustering algorithm grouped the majority of plants according to the species and subspecies passport data (Fig. 3). Cluster A consists of 89 accessions described as *S. acaule* (subspecies unknown) and 172 accessions described as *S. acaule* ssp. *acaule*. Five accessions described as *S. acaule* ssp. *aemulans* and one accession of *S. acaule* ssp. *punae* were also assigned to cluster A. One plant from accession CGN 17928, described as *S. acaule* ssp. *acaule*, and one from accession BGRC 27151, described as *S. acaule* cluster in group A, while the other examined plant from these accessions was assigned to groups G and H respectively. Four accessions of *S. acaule* (BGRC 7957, 28019 and 28039) and six of *S. acaule* ssp. *acaule* (BGRC 7971, 28018, 28022, 28026, 28040, 28493 and 31196) grouped together in cluster B. Interestingly, all ten accessions originated from the department Cochobamba in Bolivia (see Fig. 1). Because 178 out of the 182 *S. acaule* ssp. *acaule* accessions examined clustered in group A or B, both these groups are considered S. *acaule* ssp. *acaule* clusters. Of the total number of 98 accessions of *S. acaule*, 93 clustered within the two *S. acaule* ssp. *acaule* groups. Cluster C is considered to be the *S. acaule* ssp. *palmirense*/ssp. *punae* group, consisting of the single *S. acaule* ssp. *palmirense* accession and three out of the four accessions of *S. acaule* ssp. *punae*. This cluster was found to be closely related to the *S. acaule* ssp. *acaule* clusters (A and B).

Clusters D and E are considered *S. acaule* ssp. *aemulans* as 13 out of the 19 accessions fell within either one of these clusters. The seven *S. acaule* ssp. *aemulans* accessions of cluster D were all collected in the province Jujuy (J-*aemulans*) while those of cluster E all originated from the province of La Rioja (L-*aemulans*), both in

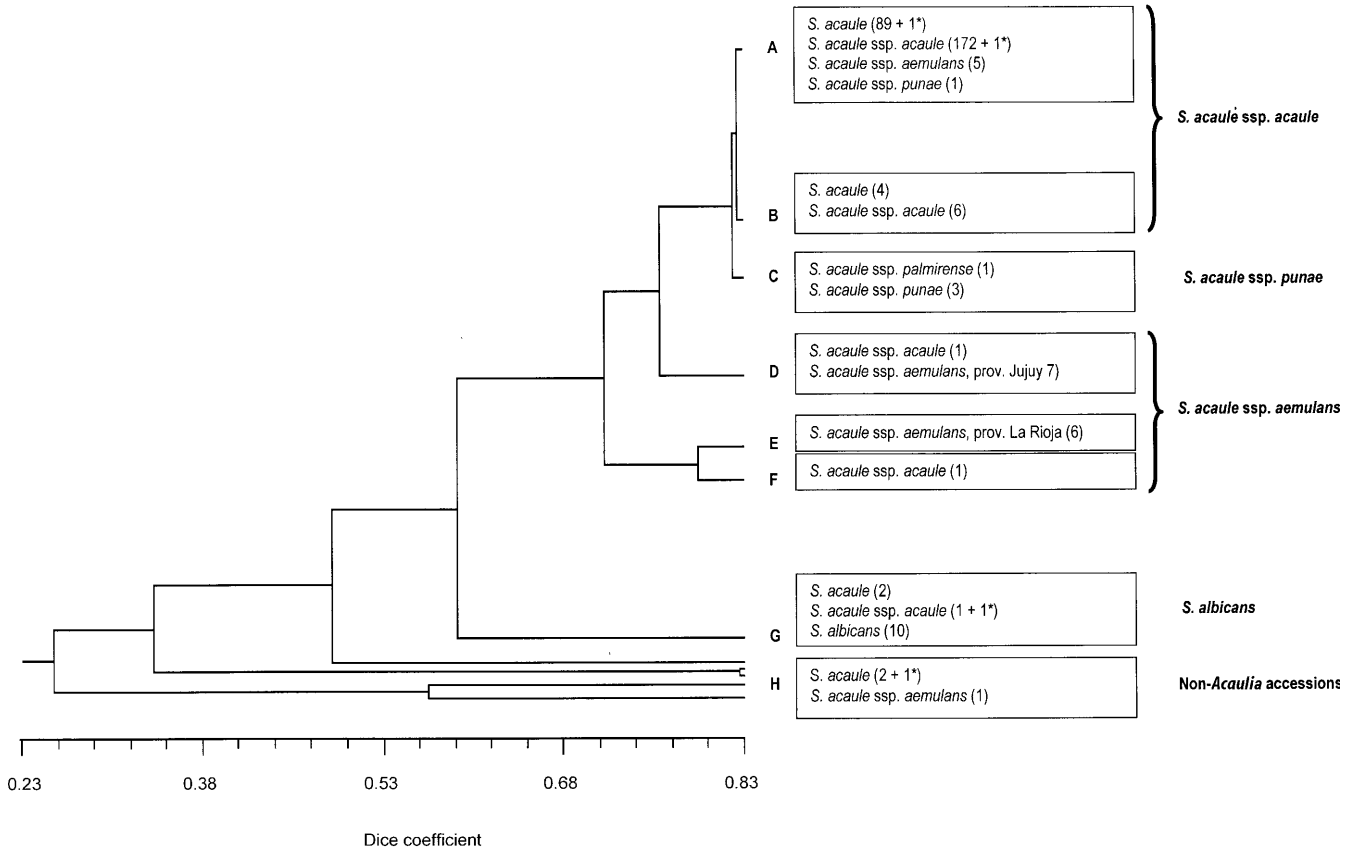


Fig. 3 UPGMA cluster analysis of the examined accessions of the series *Acaulia* based on combined data for primer combinations E+AC/M+AGC and E+AG/M+ACG. Because of the large number of plants involved, the dendrogram is cut off at Dice coefficient 0.83. Eight groups, denoted by *A–H*, are distinguished in the figure and for each group the taxonomic status of the material included in the group is indicated. Between parentheses the number of accessions involved is presented, together with the number of single plants (*-marked) in case the two individuals of an accession were clustered in different groups

Argentina. These two provinces are separated by 550 km (see Fig. 1). Group D also contains one accession that according to the passport data was also collected in the province Jujuy, but is described as *S. acaule* ssp. *acaule*. A single accession (BGRC 17112) described as *S. acaule* ssp. *acaule* (group F) clusters close to group E. This accession originated from the department Famatina in the province of La Rioja in Argentina, just like all the *S. acaule* ssp. *aemulans* accessions of group E.

All ten *S. albicans* accessions cluster together in group G, which is therefore considered to be the *S. albicans* group. Two accessions described as *S. acaule* and one described as *S. acaule* ssp. *acaule* were also assigned to group G, as well as a single plant of accession CGN 17928 described as *S. acaule* ssp. *acaule*.

Two accessions (BGRC 7968 and BGRC 27181) and one single plant from accession BGRC 27151, all described as *S. acaule*, as well as accession BGRC 17632, described as *S. acaule* ssp. *aemulans*, together formed group H. Compared to groups A–G, the accessions of group H displayed very divergent AFLP profiles, suggesting that they do not belong to the series *Acaulia*. Group H is therefore considered the "unknown" group, demanding further investigation for taxonomic classification.

Extent of redundancy

AFLP analysis with primer pairs E+AC/M+AGC and E+AG/M+ACG revealed 126 accessions with no intra-accession variation. Among these accessions 80 displayed unique AFLP profiles, while the remaining 46 accessions shared their profiles with other accessions or single plants of accessions. In order to increase resolution, these individuals were all subjected to AFLP analysis with primer pair E+AA/M+ACC, which during the testing of primer pairs was found to generate a higher number of fragments. Based on identical profiles for the three AFLP primer pairs investigated, 15 duplication groups consisting of a total number of 22 accessions and 14 single plants were identified (Table 2). Except for one case, duplication groups consist of individuals belonging to the same species, and the majority of identified duplicates appeared to be sampled from the same collection area.

Distribution of genetic diversity

Prior to analysis of the distribution of genetic diversity, the taxonomic status of accessions and single plants was

Table 2 Duplication groups of accessions and/or single plants (*-marked) based on identical AFLP profiles for all three primer pairs studied. Accessions are denoted by CGN code or BGRC number. Taxonomic status, as described prior to the present study is indicated, together with country of origin and collection area (dept.=department, prov.=province)

Accession	Taxonomic status	Country	Collection area
BGRC 7946	*S. acaule* ssp. *acaule*	Argentina	Prov. Tucuman, Dept. Tafi
BGRC 24558	*S. acaule* ssp. *acaule*	Argentina	Prov. Tucuman, Dept. Tafi
BGRC 24559	*S. acaule* ssp. *acaule*	Argentina	Prov. Tucuman, Dept. Tafi
BGRC 24560	*S. acaule* ssp. *acaule*	Argentina	Prov. Tucuman, Dept. Tafi
BGRC 24561	*S. acaule* ssp. *acaule*	Argentina	Prov. Tucuman, Dept. Tafi
CGN 17929	*S. acaule* ssp. *acaule*	Argentina	Prov. Salta, Dept. La Poma
BGRC 17145	*S. acaule* ssp. *acaule*	Argentina	Prov. Salta, Dept. La Poma
BGRC 17083	*S. acaule* ssp. *acaule*	Argentina	Prov. Jujuy, Dept. Santa Catalina
BGRC 24550	*S. acaule* ssp. *acaule*	Argentina	Prov. Jujuy, Dept. Santa Catalina
CGN 17917*	*S. acaule* ssp. *acaule*	Argentina	Prov. Jujuy, Dept. Santa Catalina
BGRC 17078*	*S. acaule* ssp. *acaule*	Argentina	Prov. Jujuy, Dept. Santa Catalina
CGN 17920*	*S. acaule* ssp. *acaule*	Argentina	Prov. Jujuy, Dept. Santa Catalina
CGN 17919*	*S. acaule* ssp. *acaule*	Argentina	Prov. Jujuy, Dept. Yavi
BGRC 27200*	*S. acaule* ssp. *acaule*	Bolivia	Dept. Oruro, Prov. Cercado
CGN 18169*	*S. acaule*	Bolivia	Dept. Oruro, Prov. Cercado
CGN 18148	*S. acaule*	Bolivia	Dept. Oruro, Prov. Carangas
CGN 18153*	*S. acaule*	Bolivia	Dept. La Paz, Prov. Ingavi
CGN 18120	*S. acaule*	Bolivia	Dept. La Paz, Prov. Murillo
BGRC 27094*	*S. acaule*	Bolivia	Dept. La Paz, Prov. Murillo
BGRC 27030*	*S. acaule*	Bolivia	Dept. La Paz, Prov. Inquisivi
CGN 18224*	*S. acaule*	Bolivia	Dept. Cochabamba, Prov. Chapare
BGRC 28022	*S. acaule* ssp. *acaule*	Bolivia	Dept. Cochabamba, Prov. Chapare
BGRC 28493	*S. acaule* ssp. *acaule*	Bolivia	Dept. Cochabamba, Prov. Ayopaya
BGRC 27205*	*S. acaule* ssp. *acaule*	Bolivia	Dept. Potosi, Prov. Ibanez
BGRC 27221*	*S. acaule*	Bolivia	Dept. Potosi, Prov. Bustillo
CGN 18139*	*S. acaule* ssp. *acaule*	Peru	Dept. Puno, Prov. Chucuito
BGRC 10071*	*S. acaule* ssp. *acaule*	Peru	Dept. Puno, Prov. Puno
CGN 20666	*S. acaule* ssp. *acaule*	Peru	Dept. Lima, Prov. Canta
BGRC 7142	*S. acaule* ssp. *acaule*	Peru	Unknown
BGRC 15465	*S. acaule*	Unknown	Unknown
BGRC 15466	*S. acaule*	Unknown	Unknown
BGRC 17124	*S. acaule* ssp. *aemulans*	Argentina	Prov. Jujuy, Dept. Tilcara
BGRC 17125	*S. acaule* ssp. *aemulans*	Argentina	Prov. Jujuy, Dept. Tilcara
CGN 18028	*S. albicans*	Peru	Dept. Cajamarca, Prov. Cajamarca
CGN 20582	*S. albicans*	Peru	Dept. Cajamarca, Prov. Cajamarca
CGN 18029	*S. acaule* ssp. *acaule*	Bolivia	Unknown

re-classified based on groups A–G (see Fig. 3). Group H, consisting of non-*Acaulia* germplasm, was omitted from the analyses, as well as accessions of unknown geographic origin and accessions from the subspecies *palmirense* and *punae* of *S. acaule* because of the small number of accessions involved. To estimate variance components for different hierarchical levels in *S. acaule* ssp. *acaule*, separate nested AMOVAs were performed for the three countries sampled (Table 3). For *S. acaule ssp. acaule* the major part (approximately 60%) of the variation was found within collection areas, with the exception of samples collected in Peru, where the major part (approximately 65%) of the variation was found among provinces within departments. As expected based on the geographic distribution of the collection areas, the majority of the variation (66%) for *S. acaule* ssp. *aemulans* was found among provinces (Table 3b). For *S. albicans* 65% of the variation was found among departments (Table 3c).

In order to investigate how genetic diversity relates to geographic distance, pair-wise comparisons between plants were made for AFLP variation and geographic distance. Significant correlations were found between geographic distance and genetic distance for *S. acaule* ssp. *acaule* (r=0.118, P<0.001), *S. acaule* ssp. *aemulans* (r=0.880, P<0.001) and *S. albicans* (r=0.232, P<0.010). However, these results should be interpreted with caution since the proportion of the variance associated with the regression (r^2) is very low for *S. acaule* ssp. *acaule* (r^2=1.4%) as well as for *S. albicans* (r^2=5.4%). For *S. acaule* ssp. *aemulans* (r^2=77.4%) the higher correlation is mainly due to the two subgroups (province Jujuy vs province La Rioja) in the dataset. Pair-wise comparisons between altitude and AFLP variation were signifi-

Table 3 Results of nested AMOVAs based on the AFLP data obtained with primer pairs E+AC/M+AGC and E+AG/M+ACG. Analyses were performed separately for *S. acaule* ssp. *acaule* (A), *S. acaule* ssp. *aemulans* (B) and *S. albicans* (C), and were carried out at different hierarchical levels. For *S. acaule* ssp. *acaule*, separate analyses were made for Argentina, Bolivia and Peru. Sample sizes are indicated by *n*. *P*-values were derived from 1,000 permutation tests and denote the probability of observing larger variance components at random

(A) *S. acaule* ssp. *acaule*

Variance component	Argentina (n=311)	Variance component	Bolivia (n=177)	Peru (n=22)
Among provinces	0.55 (14.9%) *P*<0.001	Among departments	0.59 (19.4%) *P*<0.001	0.16 (8.2%) *P*<0.100
Among departments Within provinces	0.76 (20.5%) *P*<0.001	Among provinces Within departments	0.64 (21.1%) *P*<0.001	1.29 (65.8%) *P*<0.001
Within departments	2.40 (64.6%) P<0.001	Within Provinces	1.81 (59.5%) *P*<0.001	0.51 (26.0%) *P*<0.001

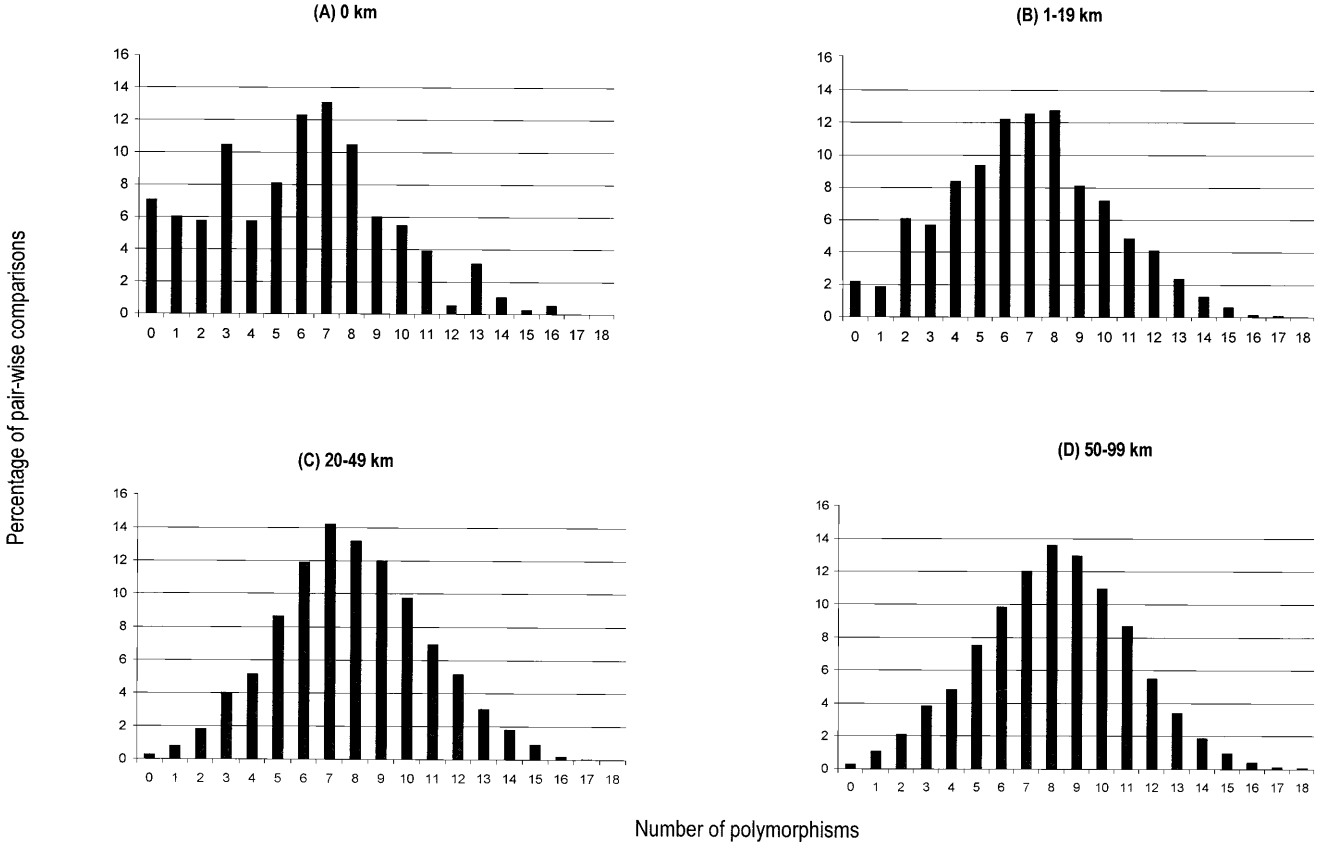
(B) *S. acaule* ssp. *aemulans*

Variance component	Argentina (n=30)
Among provinces	4.86 (65.0%) *P*<0.001
Among departments Within provinces	0.00 (0%)[a]
Within departments	2.62 (35.0%) *P*<0.001

(C) *S. albicans*

Variance component	Peru (n=20)
Among departments	1.22 (65.2%) *P*<0.001
Among provinces Within departments	0.20 (10.8%) *P*<0.050
Within provinces	0.45 (24.0%) *P*<0.001

[a] Samples for each Province (Jujuy and La Rioja) were collected in only one department (Tilcara and Famatina respectively)



cant for *S. acaule* ssp. *aemulans* (r=0.506, P<0.001), but not significant for *S. acaule* ssp. *acaule* (r=0.025) and *S. albicans* (r=0.089).

Examination of the distribution of the number of AFLP polymorphisms for different classes of geographic distance, revealed some interesting facts (Fig. 4). The number of polymorphisms between plants collected far apart followed a normal distribution (Fig. 4d). Identical AFLP profiles occurred, but only in a very low frequency. The histogram of the number of polymorphisms between plants collected at the same site seemed to be constituted of a mixture of the previously mentioned normal distribution and a substantial fraction of plants with no or only a few AFLP polymorphisms (Fig. 4a), suggesting the absence of a consistent relationship between geographic distance and genetic diversity.

Discussion

Taxonomic classification

Taxonomic classification as found in the present study was largely in line with the results obtained by Hosaka and Spooner (1992) using RFLP data, and Kardolus (1998) using both morphological and AFLP data. *S. acaule* ssp. *acaule* is more closely related to *S. acaule* ssp. *punae* than to *S. acaule* ssp. *aemulans*. Although only four accessions of ssp. *punae* could be tested, the fact that three of them cluster separately from ssp. *acaule* does not support the suggestion of Hosaka and Spooner (1992) that ssp. *punae* is synonymous with ssp. *acaule*. The single accession of ssp. *punae* that clusters with ssp. *acaule* is probably a misclassification. However, ssp. *acaule* and ssp. *punae* were not separated by a large genetic distance and therefore this apparent misclassification should be treated with caution (also see Kardolus 1998). The single accession of ssp. *palmirense*, first described by Kardolus (1998), groups loosely with the accessions of ssp. *punae*. Although this is the only accession of ssp. *palmirense* described so far, this result is in line with the opinion of Kardolus (1998) that this accession is more closely related to *S. acaule* than to *S. albicans*, notwithstanding the fact that both ssp. *palmirense* and *S. albicans* are hexaploid, and cannot refute its classification as a subspecies. The ten accessions described as *S. acaule* or *S. acaule* ssp. *acaule*, belonging to cluster B (Fig. 3), all originate from the provinces Ayopaya, Cercado, Chapare or Quillacollo of the department Chochamba in Bolivia (see Fig. 1). Four other accessions from the same department, but collected in different provinces (Arani and Arque), cluster within group A, suggesting that the germplasm collected from the provinces Ayopaya, Cercado, Chapare and Quillacollo represents a genetically differentiated subset.

The finding that accessions of *S. acaule* ssp. *aemulans* collected from the province La Rioja (L-*aemulans*) cluster separately from those collected in the province of Jujuy (J-*aemulans*) in Argentina corroborates the results of Hosaka and Spooner (1992). Also the finding that the accessions from Jujuy are more closely related to ssp. *acaule* than to the La Rioja accessions is in line with this study, although Kardolus (1998) did not observe any morphological differences between J-aemulans and L-aemulans. Okada and Clausen (1982) suggested that J-*aemulans* is a hybrid between *Solanum megistacrolobum* and ssp. *acaule*, although *S. megistacrolobum* is not found in that region (Hosaka and Spooner 1992). The elevation of ssp. *albicans* to species rank (Ochoa 1983) is clearly supported by the present study, as well as the fact that *S. albicans* is more closely related to ssp. *aemulans* than to ssp. *acaule*, *palmirense* or *punae* (Hosaka and Spooner 1992).

A total number of 16 apparent misclassifications were identified in the present study (Table 4). Three of them (BGRC 7966, CGN 17927 and CGN 20671) have already been under debate (Kardolus 1998), while the other 13 accessions are described as misclassification for the first time. The only available information about accession BGRC 18289, described as *S. acaule* ssp. *acaule*, is that it was collected in Bolivia. However, since this accession was found to be identical to accession CGN 18028, described as *S. albicans* collected in Peru (Table 2), and because it clusters together with the other accessions of *S. albicans*, this is undoubtedly an error. Accessions BGRC 17112 and 17118 have both been described as *S. acaule* ssp. *acaule*. Although they cluster with L-*aemulans* and J-*aemulans*, respectively, they are not tightly linked to the ssp. *aemulans* accessions, suggesting that these may be taxonomic intermediates. Three accessions (BGRC 7968, 17632 and 27181) did not cluster with the rest of the series *Acaulia* (Fig. 3: group H). When tissue material was collected for the present study, it was already noted that BGRC 17632 displayed a very distinct leaf morphology, although this accession has also been described as an intermediate between ssp. *acaule* and ssp. *aemulans* (J.B. Bamberg, personal communication). Recently, the accessions from cluster H were examined morphologically and supported our findings based on the AFLP data that these four accessions do not to belong to the series *Acaulia*. The correct taxonomic classification is given in the footnotes of Table 4 (R. van den Berg, personal communication).

It should be noted that only two plants per accession were analysed in the present study. In case accessions consist of individuals belonging to different taxonomic levels, the small sample size may underlie some of the discrepancies observed in taxonomic status. In the present study, two accessions (CGN 17928 and BGRC 27151) were indeed found to constitute a mix of differ-

◀ **Fig. 4** Frequency distribution of the number of AFLP polymorphisms in pair-wise comparisons of plants of *S. acaule* ssp. *acaule* based on primer combinations E+AC/M+AGC and E+AG/M+ACG. Results are presented separately for pairs that were collected at the same site (*A*) and for pairs with a geographical distance between the collection sites of 1–19 km (*B*) 20–49 km (*C*) and 50–99 km (*D*)

Table 4 Taxonomic status of accessions of the series *Acaulia* of the Dutch-German potato collection maintained at CGN, as described prior to the present study, and suggested-modification based on the AFLP profiles obtained with primer pairs E+AC/M+AGC and E+AG/M+ACG. Possible taxonomic intermediates are *-marked. Accessions are denoted by CGN code or BGRC number, and are designated "to be determined" in case no clustering was observed within the series *Acaulia* (see Fig. 3)

Prior taxonomic classification	Group (Fig. 3)	New taxonomic classification	Accession
S. acaule	G	*S. albicans*	BGRC 7955
S. acaule	G	*S. albicans*	BGRC 7956
S. acaule	A and H	*S. acaule* ssp. *acaule*/to be determined[a]	BGRC 27151
S. acaule	H	To be determined[b]	BGRC 7968
S. acaule	H	To be determined[c]	BGRC 27181
S. acaule ssp. *acaule*	A and G	S. *acaule* ssp. *acaule*/*S. albicans*	CGN 17928
S. acaule ssp. *acaule*	G	*S. albicans*	CGN 18029
S. acaule ssp. *acaule*	F	*S. acaule* ssp. *aemulans**	BGRC 17112
S. acaule ssp. *acaule*	D	*S. acaule* ssp. *aemulans**	BGRC 17118
S. acaule ssp. *aemulans*	A	*S. acaule* ssp. *acaule*	BGRC 7966
S. acaule ssp. *aemulans*	A	*S. acaule* ssp. *acaule*	BGRC 17116
S. acaule ssp. *aemulans*	A	*S. acaule* ssp. *acaule*	BGRC 17184
S. acaule ssp. *aemulans*	A	*S. acaule* ssp. *acaule*	BGRC 18496
S. acaule ssp. *aemulans*	A	*S. acaule* ssp. *acaule*	CGN 17927
S. acaule ssp. *aemulans*	H	To be determined[d]	BGRC 17632
S. acaule ssp. *punae*	A	*S. acaule* ssp. *acaule*	CGN 20671

[a] Later determined as *S. megistacrolobum*
[b] Later determined as *S. bukasovii*
[c] Later determined as *S. tuberosum ssp. andigena*
[d] Later determined as *S. kurtzianum*

ent taxonomic levels (Fig. 3, Table 4). This finding and the fact that small sample sizes were used imply that more accessions of mixed (sub)species may be present. These mixes may have occurred through sampling from plants with different taxonomic status during field collection missions, or during regeneration. Given this possibility, an increase of sample throughput by bulking multiple individuals (Hosaka and Spooner 1992; Kardolus 1998) does not seem the way to go since it may hamper the correct interpretation of results.

In addition to the identification of apparent misclassification, the AFLP data also allowed specification of the subspecies of the accessions that were described as *S. acaule* only. Out of these 98 accessions, 93 were classified as *S. acaule* ssp. *acaule*.

Extent of redundancy

Based on identical AFLP profiles for the three primer pairs, a total number of 15 duplication groups consisting of a total number of 22 accessions and 14 single plants were observed in the present study. A subset of the examined accessions were also part of two earlier studies. Hosaka and Spooner (1992) were unable to distinguish between accessions BGRC 28022 and CGN 18224, and between accessions BGRC 24559, 24560 and 24561, using RFLPs on 7–12 bulked seedlings per accession. Accessions CGN 17919 and 17920 were also found to be identical by Kardolus (1998), using isozymes on three plants per accession. For the majority of accessions, passport data supported the duplicate status, as plants from identified duplicates appeared to be collected from the same region or even from the same site (Table 2).

Most likely, the extent of redundancy has been underestimated in the present study because only plants with completely identical AFLP profiles were considered. However, the question is whether accessions should indeed be completely identical in order to consider them redundant. Although the technique is generally considered robust, it cannot be ruled out that some of the AFLP polymorphisms observed are actually artefact bands caused by methodological inconsistencies (Treuren 2001). One such band was observed for primer pair E+AA/M+ACC and subsequently was disregarded from analysis. However, reproducibility tests were performed on four replicate samples from just one individual. It cannot be ruled out, therefore, that artefact bands may occur at other positions, unless multiple samples are tested for each individual to be analysed. These problems may be circumvented by including a tolerance level for polymorphisms in the analyses (e.g. Arens et al. 1998). Heterogeneity within and among accessions may of course also be due to true genetic differences between individuals. Even though the wild species of the series *Acaulia* are self-fertilizing, intra-accession variation can be expected when studying increased sample sizes and analysing more primer pairs. Therefore, a more-reasonable question to address is whether two accessions are sufficiently different in order to consider them distinct, rather than to focus on the question whether two accessions are identical. This approach requires evaluation of the variation among accessions in relation to the variation within accessions (e.g. Phippen et al. 1997). Such analyses were not considered appropriate in the present study due to the small sample sizes used per accession.

Distribution of genetic diversity

Analyses of molecular variance showed that for *S. acaule* ssp. *acaule* collected in Argentina and Bolivia

significant genetic differentiation exists among the different collection areas, but that the major part (approximately 60%) of the variation is distributed within collection areas. Only for *S. acaule* ssp. *aemulans* was a high correlation found between geographic distance and genetic distance in pair-wise comparisons of plants. It should however be noted that this subspecies has been collected in two isolated areas (approximately 550 km apart) with no samples collected in the intervening area. For *S. acaule* ssp. *acaule*, histograms of the number of AFLP polymorphisms between plants showed a normal distribution for samples collected far apart (Fig. 4). The low frequency of identical AFLP genotypes observed among collection sites that are more than 20 km apart (Fig. 4c, d) probably represent individuals that do have a different genetic background and should not be considered duplicates. The smaller the distance between the collection sites, the larger the group with no or a few polymorphisms (Fig. 4a, b), together probably comprising duplicate genotypes. The fact that this group consists, not only of plants with identical AFLP profiles, but also of plants differing up to three bands, supports our notion that the extent of redundancy has probably been underestimated in the present study by focussing on identical AFLP profiles only. The fact that the main distribution of non-duplicates is nearly independent of geographic distance implies that genetic distance is unrelated to geographic distance between the collection sites. This is contrary to the conclusion of Hosaka and Spooner (1992) that "closely distributed accessions are also genetically close".

Implications for ex situ conservation

The application of molecular-marker technologies is becoming more and more common practise to study various aspects of plant genetic resources management (Bretting and Widrlechner 1995; Brown and Kresovich 1996). In the present study, AFLP analysis was used to validate taxonomic status, to determine the extent of redundancy and to describe the distribution of genetic diversity across the collection area of the wild potato germplasm of the series *Acaulia* maintained at CGN.

Assessing the taxonomic identity of germplasm and obtaining a clear understanding of the systematic relationships among a crop and its wild relatives is vital to the management of genetic resources as well as to the utilisation of accessions (Bretting and Widrlechner 1995). Misidentifications within the series *Acaulia*, as found in the present study, may have been due to difficulties in determining the taxonomic status of germplasm on the basis of morphology. In addition, discrepancies may have arisen during *ex situ* conservation, resulting for example from errors in documentation or in regeneration practises. In order to avoid the generation of artificial hybrids, taxonomic purity of accessions is important in the multiplication of germplasm. It cannot be ruled out that during multiplication hybridisation occurs through low frequencies of outcrossing, despite the fact that the wild germplasm of the series *Acaulia* is self-fertilising. Moreover, during the first years of the Dutch-German potato collection (1974–1988), regeneration of *Acaulia* germplasm was commonly practised by cross-pollination within accessions. In the present study, two accessions were observed that seemed to be intermediates between ssp. *acaule* and ssp. *aemulans*. If these accessions indeed represent hybrids, it remains to be determined whether these developed under natural or *ex situ* conditions.

Particularly in potato, regeneration of accessions is costly because virus tests are required in order to avoid the introduction of foreign pathogens into the country where they are conserved. In case of outcrossing wild potato species, regeneration is even more expensive because these often demand laborious hand-pollinations. Therefore, duplications are a nuisance to the genetic resources system because these do not contribute to the diversity existing in the collection, but nevertheless demand capacity for storage and maintenance. Due to increased collection sizes and decreased financial resources, genebanks have been stimulated to identify and remove redundancies from the collection to increase efficiency. In order to rationalize the collection of the series *Acaulia* by actually removing the potential duplicates identified in the present study, further insight in intra-accession variation is needed. Currently, a project is underway to investigate the potential duplication groups from the present study with AFLPs, using increased sample sizes.

Knowledge about the distribution of genetic diversity in the natural areas where the species of interest occur, may provide relevant information to develop sampling strategies that will maximise the probability of collecting genetically distinct samples. Based on the results of the present study, the probability of sampling identical or very similar genotypes of *S. acaule* ssp. *acaule* increases when sites are sampled less than 20 km apart. Despite the fact that even very distinct samples, albeit it in a low frequency, can be obtained from the same site, a distance of 20 km between collection sites could be used as a general guideline in future collection missions in order to maximise the diversity captured for *S. acaule* ssp. *acaule*.

Acknowledgements This work was supported by a fellowship, granted to C.E. McGregor, by the International Agricultural Centre on behalf of the Netherlands Ministry of Agriculture, Nature Management and Fisheries. The experiments comply with the current laws of the country in which the experiments were performed (The Netherlands).

References

Arens P, Coops H, Jansen J, Vosman B (1998) Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. Mol Ecol 7:11–18

Bretting PK, Widrlechner MP (1995) Genetic markers and plant genetic resource management. Plant Breed Rev 31:11–86

Brown SM, Kresovich S (1996) Molecular characterization for plant genetic resources conservation. In: Paterson H (ed) Genome mapping of plants. Academic Press, San Diego, pp 85–93

Brücher H (1959) Kritische Betrachtungen zur Nomenklatur Argentinischer Wildkartoffeln. V. Die Serie *Acaule*. Der Züchter 29:149–156

Centre for Genetic Resources, The Netherlands (2000) The Dutch-German Potato Collection. http://www.plant.wageningen-ur.nl/cgn/potato/

Correll DS (1962) The potato and its wild relatives: section Tuberarium of the genus *Solanum*. Texas Research Foundation, Renner, Texas

Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes:application to human mitochondrial DNA restriction data. Genetics 131:479–491

Ford-Lloyd BV, Jackson MT, Newbury HJ (1997) Molecular markers and the management of genetic resources in seed genebanks: a case study in rice. In: Callow JA, Ford-Lloyd BV, Newbury HJ (eds) Biotechnology and plant genetic resources, conservation and use. Biotechnology in Agriculture Series no. 19, CAB International, New York, pp 103–118

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13:207–209

Hawkes JG (1963) A Revision of the Tuber-bearing *Solanums* (2nd edn) Scottish Plant Breeding Station Record, pp 76–81

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

Hawkes JG (1990) The potato – evolution, biodiversity and genetic resources. Belhaven Press, London

Hintum ThJL van, Boukema IW, Visser DL (1996) Reduction of duplication in a *Brassica oleracea* germplasm collection. Genet Res Crop Evol 43:343–349

Hosaka K, Spooner DM (1992) RFLP analysis of the wild potato species, *Solanum acaule* Bitter (*Solanum* sect. *Petota*). Theor Appl Genet 84:851–858

Huamán Z, Hoekstra R, Bamberg JB (2000) The inter-genebank potato database and the dimensions of available wild potato germplasm. Am J Potato Res 77:353–362

Kardolus JP (1998) A biosystematic analysis of *Solanum acaule*. Ph D thesis, Wageningen Agricultural University, Wageningen, The Netherlands

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

Kresovich S, Lamboy WF, Li R, Ren J, Szewc-McFadden AK, Bliek SM (1994) Application of molecular methods and statistical analysis for discriminating accessions and clones of vetiver grass. Crop Sci 34:805–809

Marshall DR, Brown AHD (1975) Optimum sampling strategies in genetic conservation. In: Frankel OH, Hawkes JG (eds) Crop resources for today and tomorrow. Cambridge University Press, United Kingdom, pp 53–80

McGregor CE, Greyling MM, Lambert CA, Louw JH, Warnich L (2000) A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. Euphytica 113:135–144

Milbourne D, Meyer R, Bradshaw JE, Baird E, Bonar N, Provan J, Powell W, Waugh R (1997) Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. Mol Breed 3:127–136

Ochoa CM (1983) A new taxon and name changes in *Solanum* (sect. *Petota*) Phytologia 54:391–392

Ochoa CM (1990) The potatoes of South America: Bolivia. Cambridge University Press, Cambridge

Okada KA, Clausen AM (1982) Natural hybridization between *Solanum acaule* Bitt. and *S. megistacrolobum* Bitt. in the province of Jujuy, Argentina. Euphytica 31:817–835

Phippen WB, Kresovich S, Candelas FG, McFerson JR (1997) Molecular characterization can quantify and partition variation among genebank holdings: a case study with phenotypically similar accessions of *Brassica oleracea* var. *capitata* L. (cabbage) 'Golden Acre'. Theor Appl Genet 94:227–234

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2:225–238

Rolf FJ (1993) NTSYS-pc numerical taxonomy and multivariate analysis system, version 1.8. Exeter Software, Setauket, New York

Ross H (1986) Potato breeding – problems and perspectives. Verlag Paul Parey, Berlin Hamburg

Treuren R van (2001) Efficiency of reduced primer selectivity and bulked DNA analysis for the rapid detection of AFLP polymorphisms in a range of crop species. Euphytica 117:27–37

Verma SK, Khanna V, Singh N (1999) Random amplified polymorphic DNA analysis of Indian scented basmati rice (*Oryza sativa* L.) germplasm for identification of variability and duplicate accessions, if any. Electrophoresis 20:1786–1789

Virk PS, Newbury HJ, Jackson MT, Ford-Lloyd BV (1995) The identification of duplicate accessions within a rice germplasm collection using RAPD analysis. Theor Appl Genet 90:1049–1055

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