Theor Appl Genet (2006) 113:1515–1527
DOI 10.1007/s00122-006-0399-7

ORIGINAL PAPER

# Genetic analysis of the cultivated potato *Solanum tuberosum* L. Phureja Group using RAPDs and nuclear SSRs

**M. Ghislain · D. Andrade · F. Rodríguez · R. J. Hijmans · D. M. Spooner**

Received: 23 January 2006 / Accepted: 17 August 2006 / Published online: 14 September 2006


**Abstract** The *Solanum tuberosum* L. Phureja Group consists of potato landraces widely grown in the Andes from western Venezuela to central Bolivia, and forms an important breeding stock due to their excellent culinary properties and other traits for developing modern varieties. They have been distinguished by short-day adaptation, diploid ploidy ($2n = 2x = 24$), and lack of tuber dormancy. This nuclear simple sequence repeat (nSSR or microsatellite) study complements a prior random amplified polymorphic DNA (RAPD) study to explore the use of these markers to form a core collection of cultivar groups of potatoes. Like this prior RAPD study, we analyzed 128 accessions of the Phureja Group using nuclear microsatellites (nSSR). Twenty-six of the 128 accessions were invariant for 22 nSSR markers assayed. The nSSR data uncovered 25 unexpected triploid and tetraploid accessions. Chromosome counts of the 102 accessions confirmed these nSSR results and highlighted seven more triploids or tetraploids. Thus, these nSSR markers (except 1) are good indicators of ploidy for diploid potatoes in 92% of the cases. The nSSR and RAPD results: (1) were highly discordant for the remaining 70 accessions that were diploid and variable in nSSR, (2) show the utility of nSSRs to effectively uncover many ploidy variants in cultivated potato, (3) support the use of a cultivar-group (rather than a species) classification of cultivated potato, (4) fail to support a relationship between genetic distance and geographic distance, (5) question the use of any single type of molecular marker to construct core collections.

Communicated by F. J. Muehlbauer.

M. Ghislain (✉) · D. Andrade · F. Rodríguez
International Potato Center,
PO Box 1558, Lima, 12, Peru
e-mail: m.ghislain@cgiar.org

R. J. Hijmans
International Rice Research Institute,
Los Baños, Laguna, Philippines

*Present Address:*
F. Rodríguez · D. M. Spooner
USDA, Vegetable Crops Research Unit,
USDA-ARS, Department of Horticulture,
University of Wisconsin, 1575 Linden Drive,
Madison, WI, 53706-1590, USA

## Introduction

Genebanks have been assembled for cultivated species and their wild relatives to provide long-term availability of crop genetic diversity. However, many collections have become large and expensive to maintain, and others have accumulated redundant (duplicated) accessions (Frankel and Brown 1984). Funding constraints have stimulated studies to determine a subset of the entire collection that maintains most of the diversity of the entire collection (a "core collection;" Brown and Clegg 1983; Frankel 1984) that can be more intensively evaluated for breeding and other uses.

Many criteria have been used to construct a core collection, such as geographical distance and morphological data (Holbrook et al. 1993; Diwan et al. 1994; Basigalup et al. 1995; Huamán et al. 2000; Grenier et al. 2001), biochemical data (Grauke and Thompson 1995), or pedigree relationships (Van Hintum and Haalman 1994). Brown (1989) considered genetic marker assisted

selection as the preferred method to construct core collections, but it was initially considered too expensive and time-consuming to be effective (Gepts 1993). Increased efficiencies and concomitant cost reductions to obtain molecular data have improved dramatically, however, increasing their popularity to construct core collections. Molecular markers have several putative advantages to analyze diversity (Crawford 1990), including: (1) freedom from environmental and pleiotropic effects; (2) presence of more independent markers than morphological or biochemical data; (3) DNA characters can be more easily scored as discrete states as alleles or DNA base pairs (bp); (4) DNA characters have a much better chance of being associated to homologous traits; (5) most of the markers are selectively neutral. A large portion of the genome can be analyzed directly at the genetic level through DNA assays, while the other methods are based on phenotypic classes frequently corresponding to multiple genotypes (Gepts 1995). Molecular markers are useful for the management of ex situ collections to address genetic identification, redundancy, and genetic variation.

Several DNA marker candidates are of potential use to construct core collections (Spooner et al. 2005b). These include nuclear restriction site length polymorphisms (nRFLPs; Tanksley et al. 1989), random amplified polymorphic DNA (RAPDs; Williams et al. 1990), amplified fragment length polymorphisms (AFLPs; Vos et al. 1995) and nuclear simple sequence repeats (nSSRs, also referred to as nuclear microsatellites; Provan et al. 1996). Dominant markers, such as RAPDs and AFLPs, are relatively easy and inexpensive but are potentially less informative than co-dominant markers such as nRFLPs and nSSRs (Spooner et al. 2005b). More problematic, different molecular markers sometimes provide discordant results (Wendel and Doyle 1998). For these reasons we tested nSSRs as a complement to the RAPD data (Ghislain et al. 1999) for the construction of core collections in cultivated potato.

We use a cultivar-group classification of potato in our study. Until recently, the most commonly used taxonomic system (Hawkes 1990) recognized seven cultivated species and seven subspecies. Morphological phenetic studies of Huamán and Spooner (2002), combined with a summary of prior crossing and molecular data, showed only little support for these taxa and led them to reclassify cultivated landrace populations of potatoes (i.e., all "taxa" except modern cultivars of *Solanum tuberosum* L.) as a single species, *S. tuberosum*, with eight cultivar groups. One of these, Phureja Group (*Solanum phureja* Juz. and Bukasov subsp. *phureja*) was distinguished by short-day adaptation, diploid ploidy ($2n = 2x = 24$), and tubers lacking dormancy. They occur throughout the eastern slope of the Andes from western Venezuela to central Bolivia at elevation between 2,000 and 3,400 (Ochoa 1990).

## Materials and methods

### Plant materials, chromosome counts

One hundred and twenty-eight accessions out of the total CIP germplasm collection of 170 accessions of the Phureja Group were initially used for this study, with 70 diploid accessions of these with different nSSR genotypes (Table 1) used for final phenetic analysis (below). These 128 accessions included 85 from Colombia, 25 from Ecuador, 14 from Peru, and four from Bolivia. The identifications of these accessions were largely made by the potato taxonomists J.G. Hawkes, Z. Huamán, and C. Ochoa.

Somatic chromosome counts were made from materials planted from tubers of these collections, grown in greenhouses at the CIP station in Huancayo, Peru. Mitotic counts were obtained from root tips by the acetocarmine squash technique (Smith 1974).

### DNA extraction, microsatellite primers, and PCR conditions

Genomic DNA was obtained using standard protocols used at CIP (Ghislain et al. 1997). Its concentration was estimated by visually comparing the intensity of undigested genomic DNA with of the upper band (11,490 bp) of 1 μg of $\lambda$ DNA (Gibco-BRL, Gaithersburg, MD, USA) digested with *Pst*I (equivalent to 280 ng) and subjected to electrophoresis on a 1% agarose gel stained with ethidium bromide. We have used the potato genetic identification (PGI) kit described in Ghislain et al. 2004. The primers and amplification conditions used in the present study are listed in Table 2. PCR reactions were performed in a 20 μl volume containing 100 mM Tris–HCl, (Sigma, St. Louis, MO, USA) 20 mM $(NH_4)_2SO_4$ (Merck, Whitehouse Station, NJ, USA), 2.5 mM $MgCl_2$ (Merck), 0.2 mM of each dNTP (Pharmacia or Amersham, Phoenix, AZ, USA), 0.5 mM of each primer (forward and reverse, Genset or Operon Technologies and Gibco, Alameda, CA, USA), 1 U Taq polymerase (Gibco-BRL) and 10 ng of genomic DNA. The PCR reaction was overlaid with 50 μl of mineral oil. PCR was carried out in a PTC-100 thermocycler (MJ Research Inc., Watertown, MA, USA). The program used was the following: 3 min at 94°C, 2 min at annealing temperature (Ta), 1 min 30 s at 72°C,

**Table 1** The 70 accessions ultimately examined in this study for RAPD and nSSR phenetics (Fig. 4a, b)

| Map locality | CIP accession | Study code | Locality |
|---|---|---|---|
| 1 | 703512 | 49 | Colombia. Boyacá: Municipio Cerinza, Cerinza |
| 1 | 705808 | 87 | Colombia. Boyacá: Municipio Aquitania, Aquitania |
| 1 | 703524 | 121 | Colombia. Boyacá: Municipio Aquitania, Aquitania |
| 2 | 703572 | 16 | Colombia. Cundinamarca: Municipio Ventaquemada, Ventaquemada |
| 2 | 705798 | 71 | Colombia. Cundinamarca: Municipio Choconta, Choconta |
| 3 | 703597 | 38 | Colombia. Cauca: Municipio Silvia, Silvia |
| 4 | 703561 | 4 | Colombia. Cauca: Municipio Popayán, Popayán |
| 4 | 703566 | 9 | Colombia. Cauca: Municipio Silvia, Silvia |
| 4 | 705797 | 74 | Colombia. Valle del Cauca: no further data |
| 4 | 703562 | 85 | Colombia. Cauca: Municipio Totoró, Totoró |
| 4 | 703377 | 120 | Colombia. Cauca: Municipio Popayán, Popayán |
| 4 | 704228 | 125 | Colombia. Cauca: Municipio Silvia, Silvia |
| 4 | 704227 | 126 | Colombia. Cauca: Municipio Silvia, Silvia |
| 4 | 703570 | 127 | Colombia. Cauca: Municipio Silvia, Silvia |
| 5 | 703594 | 22 | Colombia. Cauca: Municipio Mercaderes, Mercaderes |
| 5 | 703567 | 26 | Colombia. Cauca: Municipio La Vega, La Vega |
| 5 | 703595 | 46 | Colombia. Cauca: Municipio Mercaderes, Mercaderes |
| 5 | 705825 | 83 | Colombia. Cauca: Municipio Mercaderes, Mercaderes |
| 6 | 703514 | 6 | Colombia. Nariño: Municipio Cumbal, Cumbal |
| 6 | 703541 | 8 | Colombia. Nariño: Municipio Pasto, Pasto |
| 6 | 703548 | 13 | Colombia. Nariño: Municipio Tuquerres, Tuquerres |
| 6 | 703539 | 20 | Colombia. Nariño: Municipio Tuquerres, Tuquerres |
| 6 | 703580 | 32 | Colombia. Nariño: Municipio Pasto, Pasto |
| 6 | 703586 | 36 | Colombia. Nariño: Municipio Pasto, Pasto |
| 6 | 705166 | 39 | Colombia. Nariño: no further data |
| 6 | 703545 | 40 | Colombia. Nariño: Municipio Tuquerres, Tuquerres |
| 6 | 705814 | 45 | Colombia. Nariño: Municipio Pasto, Pasto |
| 6 | 705164 | 53 | Colombia. Nariño: no further data |
| 6 | 705165 | 54 | Colombia. Nariño: no further data |
| 6 | 703579 | 58 | Colombia. Nariño: Municipio Tuquerres, Tuquerres |
| 6 | 706178 | 61 | Colombia. Nariño: no further data |
| 6 | 705154 | 64 | Colombia. Nariño: no further data |
| 6 | 705158 | 65 | Colombia. Nariño: no further data |
| 6 | 703506 | 68 | Colombia. Nariño: Municipio Contadero, Córdoba |
| 6 | 703549 | 75 | Colombia. Nariño: Municipio Cumbal, Cumbal |
| 6 | 705180 | 82 | Colombia. Nariño: no further data |
| 6 | 705802 | 90 | Colombia. Nariño: no further data |
| 6 | 705179 | 91 | Colombia. Nariño: no further data |
| 6 | 705174 | 93 | Colombia. Nariño: no further data |
| 6 | 705173 | 98 | Colombia. Nariño: no further data |
| 6 | 705168 | 101 | Colombia. Nariño: no further data |
| 6 | 705169 | 102 | Colombia. Nariño: no further data |
| 6 | 705801 | 107 | Colombia. Nariño: Municipio Contadero, Contadero |
| 6 | 705800 | 116 | Colombia. Nariño: no further data |
| 6 | 705163 | 117 | Colombia. Nariño: no further data |
| 7 | 705202 | 70 | Ecuador. Imbabura: no further data |
| 7 | 705201 | 108 | Ecuador. Imbabura: no further data |
| 7 | 705198 | 119 | |
| 8 | 703898 | 77 | Ecuador. Cotopaxi: C. Cotopaxi, Ca. De La Tacunga |
| 8 | 705327 | 86 | Ecuador. Cotopaxi: C. Saquisili, Mercado de Saquisili |
| 8 | 705319 | 109 | Ecuador. Tungurahua: C. Salcedo, Comuna Chambapongo |
| 8 | 705320 | 110 | Ecuador. Cotopaxi: C. Saquisili, Mercado de Saquisili |
| 9 | 705263 | 47 | Ecuador. Azuay: C. Gualaceo, Mercado de Gualaceo |
| 10 | 706220 | 89 | Ecuador. Azuay: C. Giron, Mercado de La Paz (Nieves) |
| 11 | 703767 | 35 | Peru. Piura: Provincia Huancabamba, Shapalla Km11 Huancabamba-Sapalache |
| 12 | 703308 | 17 | Peru. Lambayeque: Provincia Ferrenafe, Incahuasi |
| 13 | 703654 | 7 | Peru. Cajamarca: Provincia Cajamarca, Asunción |
| 14 | 704120 | 23 | Peru. Huánuco: Pachitea, Chaglla |
| 15 | 701570 | 122 | Peru. Junín: Provincia Jauja, Estación Experimental El Mantaro |

**Table 1** continued

| Map locality | CIP accession | Study code | Locality |
|---|---|---|---|
| 16 | 703800 | 1 | Peru. Puno: Provincia Sandia, Inincolla |
| 16 | 706023 | 21 | Peru. Puno: Provincia Sandia, Cuyocuyo |
| 16 | 703812 | 50 | Peru. Puno: Provincia Sandia, Cuyocuyo |
| 17 | 702287 | 19 | Bolivia. La Paz: Provincia Murillo, Mercado La Paz |
| Not mapped | 703515 | 14 | Colombia. no further data |
| Not mapped | 705806 | 15 | Colombia. no further data |
| Not mapped | 705079 | 24 | Colombia. no further data |
| Not mapped | 705804 | 33 | Colombia. no further data |
| Not mapped | 705803 | 72 | Colombia. no further data |
| Not mapped | 705805 | 105 | Colombia. no further data |
| Not mapped | 705796 | 118 | Colombia. no further data |

**Table 2** The 22 nuclear SSR primers used in this study, their chromosome location, annealing temperature, number of alleles, allele size, and polymorphic index content

| SCRI code | Chromosome | Annealing temperature (°C) | Number of alleles | Allele size (bp) | PIC-value |
|---|---|---|---|---|---|
| STM2030 | I | 55 | 1 | 180 | 0.0000 |
| STM1049 | I | 57 | 5 | 184–205 | 0.1570 |
| STM1064 | II | 55 | 1 | 192 | 0.0000 |
| STM2022 | II | 53 | 2 | 185–201 | 0.4722 |
| STM1058 | III | 55 | 1 | 116 | 0.0000 |
| STM1053 | III | 55 | 3 | 173–177 | 0.5343 |
| STM3023a | IV | 50 | 3 | 181–201 | 0.3109 |
| STM3023b | n.d. | 50 | 2 | 191–193 | 0.1490 |
| STM1031 | V | 55 | 3 | 265–269 | 0.4039 |
| STPoAc58 | V | 57 | 2 | 231–235 | 0.4909 |
| STM0019a | VI | 47 | 8 | 176–217 | 0.5939 |
| STM0019b | n.d. | 47 | 4 | 93–119 | 0.2968 |
| STM1052 | VII | 50 | 5 | 212–263 | 0.5301 |
| STM2013 | VII | 55 | 8 | 147–168 | 0.5716 |
| STM0031 | VII | 57 | 4 | 179–198 | 0.6639 |
| STWAX-2 | VIII | 53 | 5 | 227–244 | 0.6292 |
| STGBSS | VIII | 53 | 6 | 130–142 | 0.6400 |
| STM1016 | VIII | 53 | 6 | 247–262 | 0.6416 |
| STM1104 | VIII | 57 | 6 | 169–181 | 0.6730 |
| STM1017 | IX | 53 | 2 | 132–136 | 0.0253 |
| STM3012 | IX | 57 | 3 | 169–199 | 0.4823 |
| STM1106 | X | 55 | 6 | 142–164 | 0.6758 |
| STM0037 | XI | 53 | 4 | 77–95 | 0.2767 |
| STM0030 | XII | 53 | 7 | 130–152 | 0.6029 |

*n.d.* stands for not-determined

29 cycles of 1 min at 94°C, 2 min at Ta and 1 min 30 s at 72°C, with a final extension step of 5 min at 72°C. In some cases, a modified PCR program was used: 3 min at 94°; 16 double cycles of 1 min at 94°; 2 min at 60°, 1.5 min at 72°, and 1 min at 94°; 2 min at 50°, 1.5 min at 72°; and one final elongation cycle of 5 min at 72°.

## Nuclear SSR allele detection and scoring

Amplification products were separated using a denaturing polyacrylamide gel (6% acrylamide, Urea 7 M) stained with a silver staining protocol according to manufacturer's directions (Promega Corporation, Madison, WI, USA). The nSSR alleles were determined for size in bp and scored as present (1) or absent (0) on a denaturing polyacrylamide gel. Each nSSR allele was characterized by co-migration with a 1 bp ladder formed by a sequence of pUC18-forward primer. The scored band was either the upper band of each nSSR allele when visible as double band or the most intense one in the case of stutter bands.

## Data analyses

Molecular marker variance was calculated as $V_m = ((npq)/(n-1)$, where n = number of individuals of the population, and p and q = frequencies of presence and absence of a nSSR marker allele in the population.

The polymorphic index content (PIC) was calculated as $PIC = 1-\sum(p_i^2)$, where $p_i$ is the frequency of the ith allele detected in all individuals of the population (Smith et al. 1997).

We reanalyzed a 70-accession subset of the RAPD data (Ghislain et al. 1999) corresponding to the same accessions as our new nSSR data (see below). The following analyses were conducted using programs in NTSYS-pc Version 2.02k (Applied Biosystematics, Setauket, NY, USA). Similarities between accessions were estimated using the DICE, simple matching (SM) and Nei72 coefficients of similarity. Similarity matrices based on these three separate estimators (DICE, SM, Nei72), and different marker types (nSSR, RAPD) were compared using the Mantel matrix-correspondence test (Mantel 1967).

The geographic distribution of the accessions was mapped based on the locality data available (Fig. 1). Some accessions had locality data only to the level of the first-level administrative subdivisions (three records from Imbabura, Ecuador, and 15 from Nariño and one from Cauca, Colombia) and were mapped in the center of these departments, but only considering areas with an elevation of 2,000 m above sea level and higher. The maximum uncertainty (Wieczorek et al. 2003) due to this approach was about 100 km for the Colombian records and 45 km for the Ecuadorian records. We considered this acceptable given the scale of our study (2,500 km between the points furthest away). Pairwise geographic distances among all accessions were calculated using DIVA-GIS v5.2 (Hijmans et al. 2001). Instead of straight geographic distance, we could have considered measuring distances between localities using a path that is restricted to the Andes, and is not allowed to cross the Amazon low lands where there are no potatoes. However, because of the very strong North/South orientation of the Andes, such an adjustment could only marginally change the results and was therefore, not considered here. The correspondence of geographic distance to RAPD and nSSR distances was determined with the Mantel matrix-correspondence test.

Cluster analyses were conducted on similarity estimates using the unweighted pair-group method, arithmetic average (UPGMA), and neighbor joining (NJ) in SAHN. The "FIND" option was enabled to detect all possible trees. Cophenetic correlation coefficients were calculated for all combinations of similarity and tree building methods using the procedures COPH and MXCOMP. These coefficients indicate the correlation between a similarity matrix and the phenetic tree resulting from it after a cluster analysis, indicating goodness of fit of the cluster analysis to the similarity matrix. Clustering methods and similarity coefficients are described in Rohlf (1993). Bootstrap values were determined for the RAPD and nSSR trees from 500 random replicates using WinBoot (Yap and Nelson 1996). WinBoot does not handle missing data, present in a low frequency (3.4%) in the RAPD results. Bootstrap runs for RAPD data, therefore, were done twice, once with substituting missing values with a score of 0 and a second time with a substitution with 1; the lower of the two bootstrap values was considered as representative.

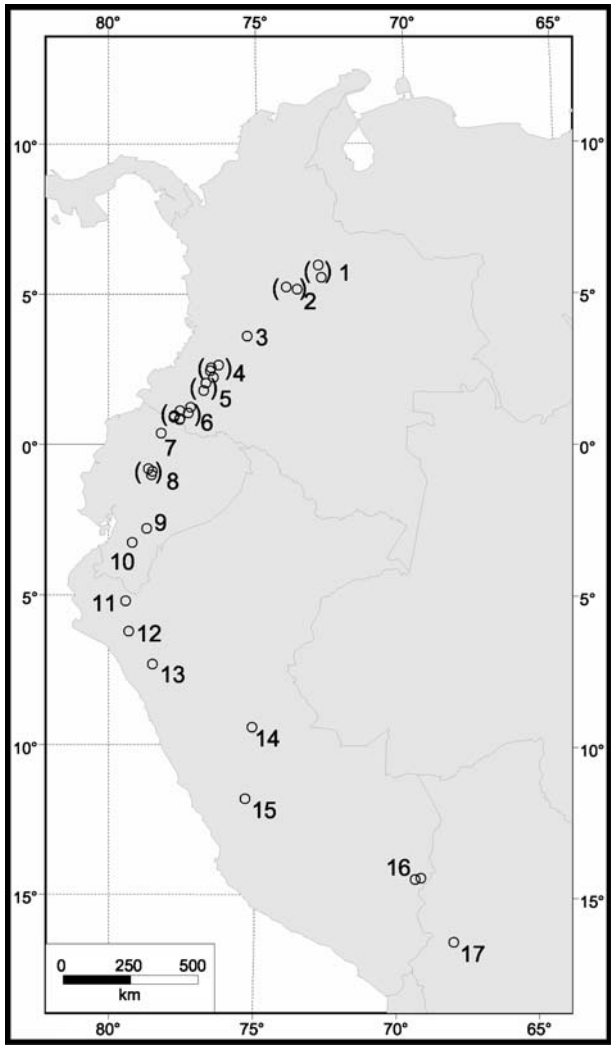


**Fig. 1** Generalized map localities for the 70 accessions examined for RAPD and nSSR variation in this study, corresponding to Figs. 4a, b. *Parenthesis* indicate multiple accessions at same locality

## Results

### Locus characterization of nSSR markers

Microsatellite markers used in the present study are claimed to be at single locus based on their genetic

linkage analysis (Milbourne et al. 1998; Ghislain et al. 2004). A complementary assay to confirm or refute their single locus status is the maximum number of alleles per genotype assayed in a large pool of genetically diverse germplasm with known ploidy. Although the nSSR set was developed with markers at single loci, we identified three markers with possibly two loci or more (STM0019, STM3023, and STM2013). Indeed, we found repeatedly 3–5 nSSR alleles per genotype of the Phureja Group accessions. In some cases nSSR allele sizes allowed us to distinguish the two loci. This was the case of STM0019, which generates two well-separated nSSR allele regions on polyacrylamide gels representing two loci named STM0019a locus and the STM0019b locus (Fig. 2). Therefore, nSSR allele identification is easy because their size range do not overlap (Table 2). The nSSR marker STM3023 is different in that size range overlaps of the alleles at the two loci, named STM3023a and STM3023b (Table 2). However, the nSSR alleles have different band phenotypes, which allow their recognition as of two different loci (Fig. 2). The third nSSR marker corresponding to at least two loci, STM2013, does not display identifiable differences but a higher number of alleles per genotypes known to be diploid by chromosome counting (Fig. 2).

## Nuclear SSR markers as ploidy identifiers

The nSSR analysis of the 128 genotypes of the Phureja Group collection revealed several genotypes not fitting the expected diploid ploidy. These accessions repeatedly displayed more than the two nSSR alleles expected for a marker at single locus (Fig. 3). The ploidy of these accessions was estimated by chromosome counting in root tips and corroborated the observation with the nSSR markers in all but one case (Table 3). These results validate the use of 21 nSSR markers of the PGI kit (excluding STM2013) as ploidy identifiers when assayed for diploid potato. Indeed, 68 of 71 genotypes identified as diploid by chromosome counting did not exceed two alleles for any of the 21 nSSR markers. Two of three remaining genotypes are thought to be scoring error while the third a labeling error and thus, eliminated from further analyses. For triploid and tetraploid genotypes, the 21 nSSR markers displayed more than two alleles in 23 cases of 28. Hence, of 74 genotypes displaying less than two alleles of 102 genotypes analyzed, the PGI kit predicted their diploid nature at 92%. Because the Phureja Group was defined partly on diploid ploidy, further analysis used only the 70 remaining accessions that were diploid and were distinct by nSSR markers.

## Redundancy in the Phureja Group collection

The use of the 22 nSSR markers on the 128 accessions of the Phureja Group collection produced 150 nSSR alleles. A total of 145 of these 150 alleles were polymorphic, and five monomorphic. One primer pair, STM2030, produced no variation on any of the 128

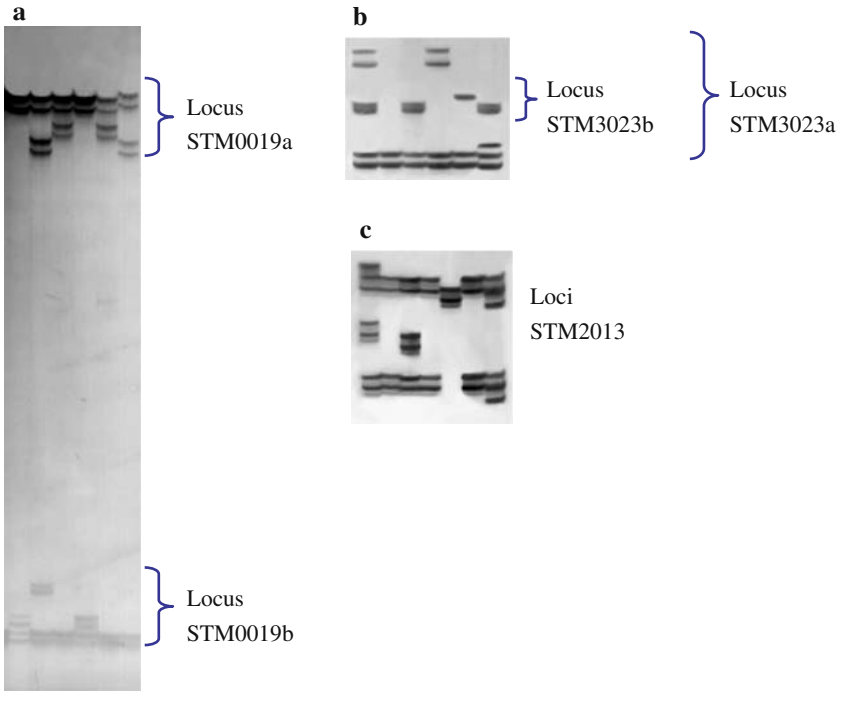


**Fig. 2** Multiloci nSSR markers: **a** STM0019 with two well separated nSSR allele regions on polyacrylamide gel (**a, b**). **b** STM3023 with clearly distinguishable nSSR allele phenotypes. **c** STM2013 with no distinctive features to identify locus specificity of the nSSR alleles

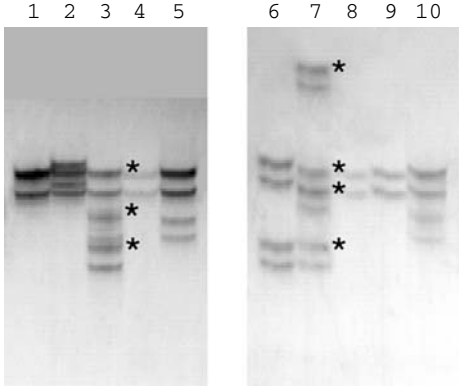


**Fig. 3** Putative ploidy inconsistencies revealed by exceeding number of nSSR alleles for single locus markers and 2× species. Example is shown for nSSR marker locus STM0019a of ten Phureja accessions of which accessions three and seven display three and four nSSR alleles (indicated by *asterisk*), respectively

accessions. Taking out the 26 polyploid accessions, the number of nSSR alleles dropped to 97 for the remaining 102 accessions. Two primer pairs (STM1064 and STM1058) produced no variation in the remaining 102 accessions (Table 2). The remaining 19 primer pairs uncovered 13 groups of individuals (total of 26) with identical nSSR alleles which were considered as duplicate because the probability of this happening by chance were between $5.2 \times 10^{-14}$ and $2.5 \times 10^{-5}$% (Table 4).

## Phenetic results

### *RAPD results*

In order to compare nSSR and RAPD marker systems, a new RAPD data set containing these 70 accessions was extracted from the original RAPD data set (Ghislain et al. 1999). Only 52 RAPD markers of the 79 originals were informative on this sample. The best cophenetic correlation coefficient was produced from the combination of DICE/UPGMA ($r = 0.618$), subjectively interpreted by Rohlf (1993) as a poor fit of the similarity matrix to the resulting tree. Successively lower *r*-values were produced by the combination Nei72/UPGMA (0.591), SM/UPGMA (0.590), SM/NJ (0.389), DICE/NJ (0.349), and NEI72/NJ (0.339). DICE (or the similar Jaccard's coefficient, that both ignore 0/0 matches for dominant markers) typically is considered an appropriate estimator for RAPDs. A single DICE/UPGMA tree was produced (Fig 4a). Bootstrap support for this tree was poor (mainly in the inner nodes) to as high as 82%. There are no well-defined geographic clusters.

**Table 3** Coincidence between allele number of 21 SSR markers and ploidy of potato landraces of the Phureja Group

| SSR alleles | Chrom counts | Microsatellite loci | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | STM2022 | STM1053 | STM1058 | STM3023a | STPoAc58 | STM0019a | STM0019b | STM0031 | STM1052 | STGBSS | STM1016 | STWAX-2 | STM1104 | STM3012 | STM1106 | STM0037 | STM0030 | | | | |
| | | II | III | III | IV | V | VI | n.d. | VII | VII | VIII | VIII | VIII | VIII | IX | X | XI | XII | | | | |
| 1 or 2 | 24 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 70 | 71 | 71 | 71 | 71 | 71 | 71 | 69 | 70 | 71 | | | | |
| | 36 | 10 | 10 | 10 | 10 | 8 | 8 | 10 | 8 | 9 | 9 | 7 | 8 | 9 | 10 | 10 | 9 | 10 | | | | |
| | 48 | 14 | 16 | 17 | 17 | 10 | 12 | 17 | 12 | 8 | 11 | 9 | 13 | 14 | 17 | 15 | 15 | 14 | | | | |
| 3 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | | | | |
| | 36 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 2 | 1 | 1 | 3 | 2 | 1 | 0 | 0 | 1 | 0 | | | | |
| | 48 | 4 | 2 | 1 | 1 | 7 | 5 | 1 | 6 | 6 | 7 | 8 | 5 | 4 | 1 | 2 | 3 | 3 | | | | |
| 4 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| | 36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| | 48 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 4 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | | | | |

Genotypes invariants by SSR marker analyses (26) or with ambiguous ploidy determination (3) are not included

**Table 4** List of genotypes of the Phureja Group that could not be resolved with the 22 nSSR markers of the potato genetic identification kit

| CIP accession | Study code | Locality | SSR duplicates | By chance | Final list of Phureja Group |
|---|---|---|---|---|---|
| 705803 | 72 | Colombia. no further data | Group 1 | $5.2 \times 10^{-14}$ | Yes |
| 703508 | 31 | Colombia. Nariño: Municipio Cumbal, Cumbal | Group 1 | | |
| 703528 | 42 | Colombia. Boyacá: Municipio Cerinza, Cerinza | Group 2 | $3.1 \times 10^{-6}$ | No (ploidy: 4×) |
| 705159 | 28 | Colombia. Cundinamarca: Municipio Chocontá, Chocontá | Group 2 | | |
| 705172 | 99 | Colombia. Boyacá: no further data | Group 2 | | |
| 705821 | 30 | Colombia. Nariño: Municipio Tuquerrers, Tuquerres | Group 2 | | |
| 703506 | 68 | Colombia. Nariño: Municipio Contadero, Contadero | Group 3 | $1.1 \times 10^{-6}$ | Yes |
| 703509 | 18 | Colombia. Nariño: Municipio Cumbal, Cumbal | Group 3 | | |
| 705198 | 119 | Ecuador. Imbabura: no further data | Group 4 | $1.5 \times 10^{-5}$ | Yes |
| 706180 | 103 | Colombia. Boyacá: no further data | Group 4 | | |
| 705327 | 86 | Ecuador. Cotopaxi: C. Saquisili, Mercado de Saquisili | Group 5 | $1.3 \times 10^{-5}$ | Yes |
| 703581 | 56 | Colombia. Nariño: Municipio Ipiales, Ipiales | Group 5 | | |
| 705157 | 10 | Colombia. Cundinamarca: Municipio Chocontá, Chocontá | Group 5 | | |
| 705161 | 67 | Colombia. Valle del Cauca: no further data | Group 5 | | |
| 705197 | 52 | Ecuador. Imbabura: no further data | Group 5 | | |
| 705217 | 73 | Ecuador. Carchi: no further data | Group 5 | | |
| 706169 | 76 | Ecuador. Imbabura: no further data | Group 5 | | |
| 707110 | 43 | Colombia. Cundinamarca: Municipio Chocontá, Chocontá | Group 5 | | |
| 703377 | 120 | Colombia. Valle del Cauca: Municipio Popayan, Popayan | Group 6 | $6.9 \times 10^{-6}$ | Yes |
| 705810 | 84 | Colombia. Boyacá: Municipio La Uvita, La Uvita | Group 6 | | |
| 704228 | 125 | Colombia. Nariño: no further data | Group 7 | $9.6 \times 10^{-6}$ | Yes |
| 703600 | 115 | Colombia. Antioquia: no further data | Group 7 | | |
| 705156 | 25 | Colombia. Cundinamarca: no further data | Group 7 | | |
| 705175 | 92 | Colombia. Boyacá: no further data | Group 7 | | |
| 707109 | 59 | Colombia. Cundinamarca: Municipio Chocontá, Chocontá | Group 7 | | |
| 705319 | 109 | Ecuador. Tungurahua: C. Salcedo, Mercado de Salcedo | Group 8 | $2.8 \times 10^{-7}$ | Yes |
| 705178 | 95 | Colombia. Boyacá: no further data | Group 8 | | |
| 703524 | 121 | Colombia. Boyacá: Municipio Aquitania, Aquitania | Group 9 | $9.6 \times 10^{-6}$ | Yes |
| 703565 | 55 | Colombia. Valle del Cauca: Municipio Silvia, Silvia | Group 9 | | |
| 705823 | 114 | Colombia. Nariño: Municipio Encamo, Encamo | Group 9 | | |
| 705162 | 11 | Colombia. Boyacá: no further data | Group 10 | $2.5 \times 10^{-5}$ | No (ploidy: 4×) |
| 705813 | 104 | Colombia. Nariño: Municipio Pasto, Pasto | Group 10 | | |
| 706179 | 29 | Colombia. Cundinamarca: no further data | Group 10 | | |
| 705154 | 64 | Colombia. Cundinamarca: no further data | Group 11 | $7.6 \times 10^{-7}$ | Yes |
| 703510 | 34 | Colombia. Valle del Cauca: no further data | Group 11 | | |
| 703547 | 5 | Colombia. Nariño: Municipio Tuquerrers, Tuquerres | Group 12 | $3.0 \times 10^{-7}$ | No (ploidy: 3×) |
| 705262 | 51 | Ecuador. Azuay: C. Gualaceo, Mercado de Gualaceo | Group 12 | | |
| 705163 | 117 | Colombia. Cundinamarca: Municipio Chocontá, Chocontá | Group 13 | $6.2 \times 10^{-6}$ | Yes |
| 703589 | 111 | Colombia. Nariño: Municipio Pupiales, Pupiales | Group 13 | | |

The column labeled "By chance" represents the probability that these accessions would have the same SSR alleles by chance alone. Those included in the final list of the 70 genotypes of the Phureja Group are indicated in the last column

*Nuclear SSR results*

The best cophenetic correlation coefficient was produced from the combination of SM/UPGMA ($r = 0.821$), subjectively interpreted by Rohlf (1993) as a good fit of the similarity matrix to the resulting tree. Successively lower *r*-values were produced by the combination SM/NJ (0.814), Nei72/UPGMA (0.800), SM/NJ (0.733), DICE/NJ (0.594), and NEI72/NJ (0.583). The SM coefficient is often considered an appropriate estimator for dominant nSSR markers. A total of 192 DICE/UPGMA trees were produced. Figure 4b presents a strict consensus tree of these 192 trees. Bootstrap support for this tree was poor (mainly in the inner nodes) to as high as 99%. Similarly, there are no well-defined geographic clusters.

Tree comparisons

There is extremely low concordance between the RAPD and nSSR results, and either of these to geographic distance. This is evident by: (1) the extremely low-Mantel matrix-correspondence values (0.21) that compared the DICE similarity matrix of the RAPD data to the SM matrix of the nSSR data, 0.12 for RAPD/geography, and 0.37 for SSR/geography; (2) the extremely low-Mantel matrix-correspondence

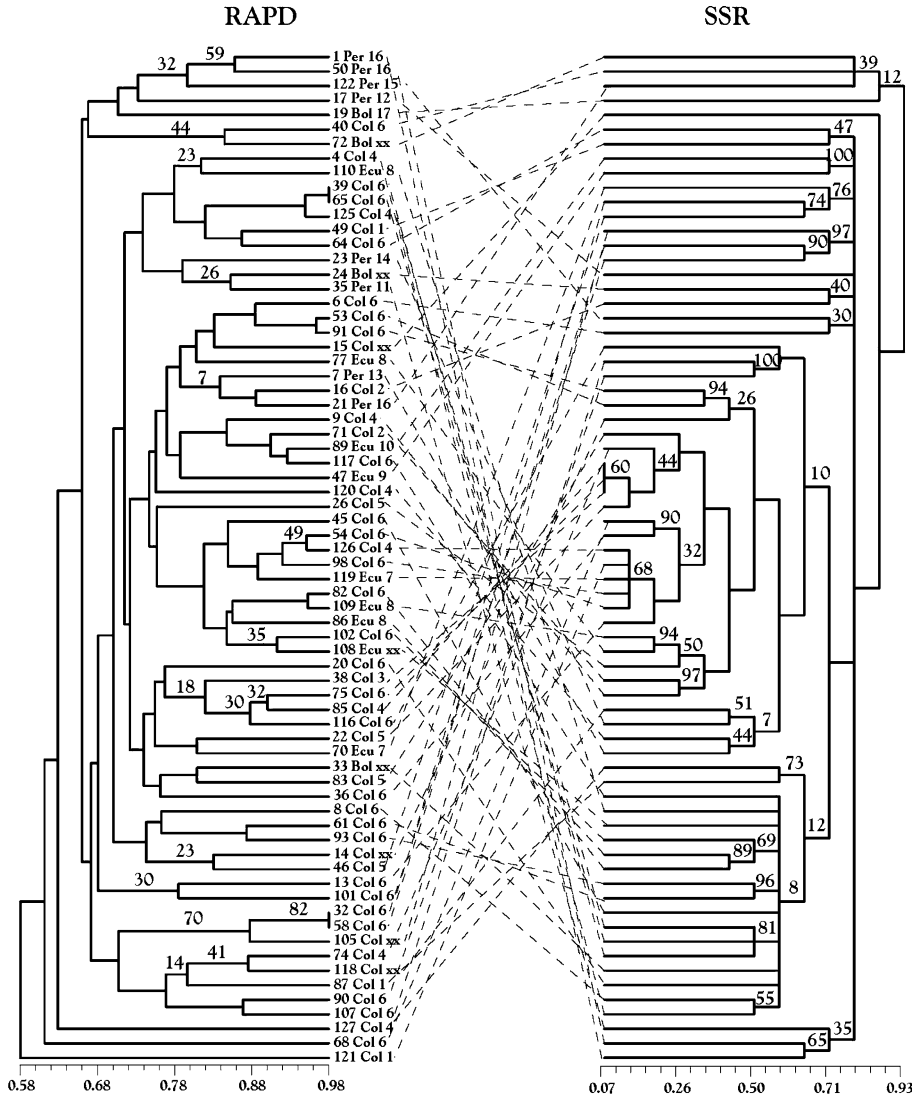


**Fig. 4** **a** Unweighted pair-group method, arithmetic average tree generated from RAPD data analyzed with DICE similarity coefficient (a single tree was produced) of 70 accessions of the cultivated potato *Solanum tuberosum* Phureja Group. Numbers above the branches represent bootstrap values. The three codes following the branches are accession numbers corresponding to Table 1, followed by a three-letter country code (*Col* Colombia, *Ecu* Ecuador, *Per* Peru, *Bol* Bolivia), followed by a map locality (xx = no locality data beyond country) corresponding to Fig. 1. **b** Consensus UPGMA tree generated from nSSR data of the same 70 accessions in Fig 4a, analyzed with SM similarity coefficient; the numbers in *circles* above the branches also represent bootstrap values. *Dotted lines* connect the identical accessions in the RAPD and nSSR results, showing discordance between RAPD and nSSR results

values of the corresponding trees of these similarity matrices of 0.29 for RAPD/nSSR, 0.17 for RAPD/geography, and 0.37 for SSR/geography; these low values across these marker types contrast with high values (0.93 and higher) comparing all combinations of similarity matrices and tree building methods within marker type; (3) a visual comparison of correspondence between the RAPD and nSSR trees (Figs. 4a, b), and (4) the very few branches that are common to both studies (designated by asterisks) on the trees.

## Use of nSSRs for constructing a core collection

In order to select a core collection of the 128 accessions of the Phureja Group, we applied the methodology described previously using RAPD markers (Ghislain et al. 1999). From each stratum under a certain similarity value, a genotype is selected considering maximum allele retention. Sampling is based on the phenogram built by similarity values not considering geographical origin. Due to the reduction from 128 to 70 accessions, the former core sample of 20 accessions using RAPD markers is not valid anymore as only five accessions of the 20 are still present in the 2× Phureja Group. A new selection was made using 91 nSSR alleles (Table 5). As shown previously with RAPD data, the selection of most genetically diverse samples based on marker data can effectively reduce the sample size down to less than 20 individuals. Here, with only 18 of them less than 8% of the nSSR alleles are lost. The addition of 52 RAPD marker data did not maximize the ratio sample size versus allele content. On the contrary, the addition of RAPD marker data reduces the percent of marker allele loss for the same sample size (Table 5).

**Table 5** Comparison of core collection selection using 91 nSSR alleles alone and with the addition of 52 RAPD marker data of both on 70 genotypes of the Phureja Group

| Size | Similarity index | Molecular variance | Allele loss (%) |
|---|---|---|---|
| 91 nSSR alleles | | | |
| 70 | 0.98 | 9.33 | 0 |
| 54 | 0.95 | 7.56 | 2.20 |
| 48 | 0.90 | 6.85 | 3.30 |
| 37 | 0.85 | 5.60 | 5.49 |
| 24 | 0.80 | 3.88 | 6.59 |
| 18 | 0.75 | 3.09 | 7.69 |
| 12 | 0.70 | 2.08 | 16.48 |
| 91 nSSR alleles + 52 RAPD marker data | | | |
| 70 | 0.98 | 17.38 | 0 |
| 66 | 0.95 | 16.49 | 0 |
| 57 | 0.90 | 14.42 | 1.40 |
| 51 | 0.85 | 13.06 | 2.10 |
| 37 | 0.80 | 9.85 | 2.10 |
| 20 | 0.75 | 5.64 | 6.99 |
| 11 | 0.70 | 3.09 | 19.58 |

Similarity index is calculated using the UPGMA method and the DICE similarity coefficient

## Discussion

### Nuclear SSR markers as support for a cultivar-group classification of potato

Huamán and Spooner (2002) recognized cultivar groups as appropriate to serve the practical purposes to users, but not as natural taxa deserving species status. They defined the Phureja Group by short-day adaptation, diploid ploidy ($2n = 2x = 24$), and tubers lacking dormancy. Chromosome counts were not available to Hawkes and Ochoa who originally identified these accessions, and their assumption of diploidy for these accessions was apparently incorrect in some cases. Our new chromosome count data show that cultivars with no tuber dormancy span diploid and tetraploid ploidy levels. They provide further support to the lack of coherent traits maintaining the classification of cultivated potatoes as separate species, and further support a cultivar-group classification of cultivated potatoes. We propose a refinement of the definition of the Phureja Group here be redefined to include cultivars lacking tuber dormancy, irrespective of their ploidy.

Until recently, potato was hypothesized to have multiple and independent origins (reviewed in Huamán and Spooner 2002), but Spooner et al. (2005a) found support for a single origin of potato from a member of the wild potato *Solanum brevicaule* complex from southern Peru. Many studies have shown that potato fields in the Andes contain mixtures of cultivated species at different ploidy levels (e.g., Ochoa 1958; Brus et al. 1981; Johns et al. 1987; Zimmerer 1991). Watanabe and Peloquin (1989, 1991) showed both diploid and unreduced gametes to be common in the South American wild and cultivated species, potentially allowing gene transfer among different ploidy levels. Under such a scenario it is therefore, not surprising that there has not been a divergence of *S. tuberosum* into well-defined species.

### Concordance between RAPD and SSR trees

Our results revealed extremely low concordance between the RAPD and the nSSR trees. Similar very low concordance values between RAPD and nuclear RFLP also were found by Powell et al. (1996, $r = 0.150$; for infraspecific comparisons as in our present study), Milbourne et al. (1997, $r$ not calculated but other correlations measures relatively low), and Russell et al. (1997, $r = 0.235$). Sun et al. (1999, $r = 0.267$) similarly found very low concordance of RAPDs and nSSRs. However, Sun et al. (1997) found a good correlation value of 0.80 when comparing wheat nSSR markers and RAPD markers in several species of *Elymus*. A possible explanation for such difference may rely in the type of information provided by each molecular marker system. RAPD markers are randomly distributed in the genome and hence largely from the intergenic regions. In contrast half of the nSSR markers used in the present study were developed from the untranslated regions and introns of genes (8 in 5′or 3′ UTR and 5 in introns). It is possible that the selection on gene bias the nSSR allele distribution as reported in a review on microsatellites within genes by Li et al. (2004). In addition, nSSR markers do not undergo the same evolutionary process as AFLPs, RAPDs, or nRFLPs, but evolve mostly by a slippage mechanism more frequently that base pair changes and insertions and deletions (Tautz and Rentz 1984; Tautz et al. 1986; Delsney et al. 1988). Another factor is possible non-homology of bands (Chalmers et al. 1992; Tinker et al. 1993). Homoplastic bands are logically more frequently observed on low-resolving gels such as on agarose gels (used in RAPD) than on polyacrylamide gels (used for nSSR). Finally, the number of markers might be too low to provide good estimate of the genetic variability (Messmer et al. 1991; Smith et al. 1992), as precision improves with greater number of markers (Moser and Lee 1994; Spooner et al. 1995).

Gene flow among populations is thought to diminish with geographic distance and thus, germplasm collectors are advised to collect from as many geographically distant sites as possible to maximize genetic variation

(Marshall and Brown 1975; Chapman 1989). Actual measurements of genetic diversity and geographic distance, however, have shown that this expectation is not always met. The lack of such association we show here also was found in crop wild relatives by Lamboy et al. (1996) in *Malus*, Gallois et al. (1998) in *Fagus*; Fahima et al. (1999) in wheat; and del Rio et al. (2001), del Rio and Bamberg (2002) and McGregor et al. (2002) in wild potatoes. As pointed out by del Rio et al. (2001), the low size of potato populations (generally < 100 plants per population) may make other stochastic events more important in partitioning of genetic diversity, such as environmental changes, demographic factors (i.e., chance differences among individuals in survivorship or fecundity) and genetic drift. In a cultivated species, where genotypes are actively dispersed through trade, such an association is even less likely to occur.

### Nuclear SSR and/or RAPD markers for constructing a core collection

Our results reveal the superior polymorphism detection power of nSSR markers compared to RAPD markers for similar number of markers. The utility of this nSSR core collection is suspect, however, for actual use by breeders for two reasons. First, the failure of RAPD and nSSR data to give similar results suggests that these diversity estimates do not measure the same genetic diversity. Second, studies in common bean show no relationship between neutral marker diversity and functional diversity (Skroch et al. 1998). Perhaps the assumption by Brown (1989) and others that neutral marker diversity is of superior use to construct core collections should be challenged and tested with assays of functional diversity.

**Acknowledgments** We thank Rene Gomez for providing the plant material and associated information used in the present study and Jorge Núñez for his valuable help on SSR data management.

## References

Basigalup DH, Barnes DK, Stucker RE (1995) Development of a core collection for perennial *Medicago* plant introductions. Crop Sci 35:1163–1168

Brown AHD (1989) Core collections: a practical approach to genetic resources management. Genome 31:818–824

Brown AHD, Clegg MT (1983) Isozyme assessment of plant genetic resources. In: Rattazzi MC, Scandalios JG, Whitt GS (eds) Isozymes: current topics in biological and medical research, vol 11. Alan R. Liss, New York, NY, pp 285–295

Brus SB, Carney HJ, Huamán Z (1981) Dynamics of Andean potato agriculture. Econ Bot 35:70–88

Chalmers KJ, Waugh R, Sprent JI, Simmons AJ, Powell W (1992) Detection of genetic variation between and within populations of *Gliricidia sepium* and *G. maculata* using RAPD markers. Heredity 69:465–472

Chapman CGD (1989) Collection strategies for the wild relatives of field crops. In: Brown ADH, Frankel OH, Marshall DR, Williams JT (eds) The use of plant genetic resources. Cambridge University Press, Cambridge, UK, pp 263–279

Crawford DJ (1990) Plant molecular systematics: macromolecular approaches. Wiley, New York

del Rio AH, Bamberg JB (2002) Lack of association between genetic and geographical origin characteristics for the wild potato *Solanum sucrense*. Am J Potato Res 79:335–338

del Rio AH, Bamberg JB, Huamán Z, Salas A, Vega SE (2001) Association of ecogeographical variables and RAPD marker variation. Wild potato populations of the USA. Crop Sci 41:870–878

Delsney M, Grellet F, Tremousaygue D, Raynal M, Panabieres F (1988) Structure, evolution et expression de l'DNA nucleaire. Bull Soc Bot Fr 135:23–38

Diwan N, Bauchan GR, McIntosh MS (1994) A core collection for the United States annual *Medicago* germplasm collection. Crop Sci 34:279–285

Fahima T, Sun GL, Beharav A, Krugman T, Beiles A, Nevo E (1999) RAPD polymorphism of wild emmer wheat populations, *Triticum dicoccoides*, in Israel. Theor Appl Genet 98:434–447

Frankel OH (1984) Genetic perspectives of germplasm conservation. In: Arber WK, Llimensee K, Peacock WJ, Starlinger P (eds) Genetic manipulation: impact on man and society. Cambridge University Press, Cambridge, UK, pp 161–170

Frankel OH, Brown AHD (1984) Current plant genetic resources—a critical appraisal. In: Chopra VL, Joshi BC, Sharma RP, Basnal HC (eds) Genetics: new frontiers. In: Proceeding of the fifth International Congress General, vol 4. Oxford and IBH Publishing Co., New Delhi, India, pp3–13

Gallois A, Audran JC, Burus M (1998) Assessment of genetic relationships and population discrimination among *Fagus sylvatica* L. by RAPD. Theor Appl Genet 97:211–219

Gepts P (1993) The use of molecular and biochemical markers in crop evolution studies. Evol Biol 27:51–94

Gepts P (1995) Genetic markers and core collections. In: Hodgkin T, Brown AHD, van Hintum TJL, Morales EAV (eds) Core collections of plant genetic resources. Wiley, Chichester, UK, pp127–146

Ghislain M, Zhang D, Fajardo D, Huamán Z, Hijmans R (1999) Marker-assisted sampling of the cultivated Andean potato *Solanum phureja* collection using RAPD markers. Genet Res Crop Evol 46:547–555

Ghislain M, Zhang D, Herrera-Montoya M (eds) (1997) Molecular biology laboratory protocols: Plant genotyping. Genetic resources department training manual. CIP, Lima, Peru

Ghislain M, Spooner DM, Rodríguez F, Villamón F, Núñez J, Vásquez C, Waugh R, Bonierbale M (2004) Selection of highly informative and user-friendly microsatellites (SSRs) for genotyping of cultivated potato. Theor Appl Genet 108:881–890

Grauke LJ, Thompson TE (1995) Evaluation of pecan [*C. illinoinensis* (Wangenh.) K. Koch] germplasm collection designation of a core subset. HortScience 30:950–954

Grenier C, Bramel-Cox PJ, Hamon P (2001) Core collection of sorghum. I. Stratification based on geographical data. Crop Sci 41:234–240

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

Hijmans RJ, Guarino L, Cruz M, Rojas E (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. Pl Genet Res Newsl 127:15–19

Holbrook CC, Anderson WF, Pittman RN (1993) Selection of a core collection from the US germplasm collection of pea. Crop Sci 33:859–861

Huamán Z, Ortiz R, Gómez R (2000) Selecting a *Solanum tuberosum* subsp. *andigena* core collection using morphological, geographical, disease and pest descriptors. Am J Potato Res 77:183–190

Huamán Z, Spooner DM (2002) Reclassification of landrace populations of cultivated potatoes (*Solanum* sect. *Petota*). Am J Bot 89:947–965

Johns T, Huamán Z, Ochoa CM, Schmiediche PE (1987) Relationships among wild, weed, and cultivated potatoes in the *Solanum ajanhuiri* complex. Syst Bot 12:541–552

Lamboy WF, Yu J, Forsline PL, Weeden NF (1996) Partitioning of allozyme diversity in wild populations of *Malus sieversii* L. and implications for germplasm collection. J Am Soc Hort Sci 121:982–987

Li Y-C, Korol AB, Fahima T, Nevo E (2004) Microsatellites with genes: structure, function, and evolution. Mol Biol Evol 21:991–1007

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

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

McGregor CE, van Treuren R, Hoekstra R, van Hintum TJL (2002) Analysis of the wild potato germplasm of the series *Acaulia* with AFLPs: implications for ex situ conservation. Theor Appl Genet 104:146–156

Messmer MM, Melchinger AE, Woodman WL, Lee EA, Lamkey KR (1991) Genetic diversity among progenitors and elite lines from the Iowa Stiff Stalk Synthetic (BSSS) maize populations: comparison of allozyme and RFLP data. Theor Appl Genet 83:97–107

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

Milbourne D, Meyer RC, Collins AJ, Ramsay LD, Gebhardt C, Waugh R (1998) Isolation, characterization and mapping of simple sequence repeat loci in potato. Mol Gen Genet 259:233–245

Moser H, Lee M (1994) RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variation in oats. Theor Appl Genet 87:947–956

Ochoa CM (1958) Expedición colectora de papas cultivadas a la cuenca del Lago Titicaca. I. Determinación sistemática y número cromosómico del material colectado. Programa Cooperativo de Experimentación Agropecuaria (PCEA), Ministerio de Agricultura, Lima, Perú

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

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

Provan J, Powell W, Waugh R (1996) Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*). Theor Appl Genet 92:1078–1084

Rohlf FJ (1993) NTSYS-pc, Numerical taxonomy and multivariate system. Exeter Publishing, Ltd., New York, NY, USA

Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W, Waugh R (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor Appl Genet 95:714–722

Skroch P, Nienhuis J, Beebe S, Tohme J, Pedraza F (1998) Comparison of Mexican common bean (*Phaseolus vulgaris* L.) core and reserve collections. Crop Sci 38:488–496

Smith BW (1974) Cytological evidence. In: Radford AE, Dickison WC, Massey JR, Bell CR (eds) Vascular plant systematics. Harper and Row, New York, pp237–258

Smith JSC, Chin ECL, Shu L, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997) An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor Appl Genet 95:163–173

Smith OS, Smith JSC, Bowen SL, Tenborg RA (1992) Numbers of RFLP probes necessary to show associations between lines. Maize Genet Coop News Lett 66:66

Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005a) A single domestication for potato based on multilocus AFLP genotyping. Proc Natl Acad Sci USA 120:14694–14699

Spooner DM, Tivang J, Nienhuis J, Miller JT, Douches DS, Contreras-MA (1995) Comparison of four molecular markers in measuring relationships among the wild potato relatives *Solanum* section *Etuberosum* (subgenus *Potatoe*). Theor Appl Genet 92:532–540

Spooner DM, van Treuren RR, de Vicente MC (2005b) Molecular markers for germplasm and genebank management. Tech Bull 10. International Plant Genetic Resources Institute, Rome, Italy, pp1–136

Sun GL, Diaz O, Salomon B, von Bothmer R (1999) Genetic diversity in *Elymus caninus* as revealed by isozyme, RAPD, and microsatellite markers. Genome 42:420–431

Sun GL, Salomon B, von Bothmer R (1997) Analysis of tetraploid *Elymus* species using wheat microsatellites markers and RAPD markers. Genome 40:806–814

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. Biotechnology 7:257–260

Tautz D, Rentz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12:4127–4138

Tautz D, Trick M, Dover GA (1986) Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652–656

Tinker NA, Fortin MG, Mather DE (1993) Random amplified polymorphic DNA and pedigree relationships in spring barley. Theor Appl Genet 85:976–984

Van Hintum TJL, Haalman D (1994) Pedigree analysis for composing a core collection of modern cultivars, with examples from barley (*Hordeum vulgare* s. lat.). Theor Appl Genet 88:70–74

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, 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

Watanabe K, Peloquin SJ (1989) Occurrence of 2*n* pollen and *ps* gene frequencies in cultivated groups and their related wild species in the tuber-bearing Solanums. Theor Appl Genet 78:329–336

Watanabe K, Peloquin SJ (1991) The occurrence and frequency of 2n pollen in 2x, 4x, and 6x wild, tuber-bearing *Solanum* species from Mexico, and Central and South America. Theor Appl Genet 82:621–626

Wendel JF, Doyle JJ (1998) Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis DE, Soltis PS, Doyle JJ (eds) Molecular systematics of plants II: DNA sequencing. Kluwer Academic Publishers, Boston, pp 265–296

Wieczorek J, Guo Q, Hijmans RJ (2004) The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. Inter J Geogr Inform Sci 18:745–767

Williams J, Kubelik A, Livak K, Rafalski JA, Tingey S (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Yap I, Nelson RJ (1996) WinBoot: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. IRRI discussion paper series No. 14, International Rice Research Institute, Manila, Philippines

Zimmerer K (1991) The regional biogeography of native potato cultivars in highland Peru. J Biogeogr 18:165–178