

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Publications from USDA-ARS / UNL Faculty

U.S. Department of Agriculture: Agricultural  
Research Service, Lincoln, Nebraska

---

12-2007

## Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification

David M. Spooner

*University of Wisconsin, david.spooner@ars.usda.gov*

Jorge Nunez

*International Potato Center*

Guillermo Trujillo

*International Potato Center*

Maria del Rosario Herrera

*International Potato Center*

Frank Guzman

*International Potato Center*

*See next page for additional authors*

Follow this and additional works at: <https://digitalcommons.unl.edu/usdaarsfacpub>

---

Spooner, David M.; Nunez, Jorge; Trujillo, Guillermo; del Rosario Herrera, Maria; Guzman, Frank; and Ghislain, Marc, "Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification" (2007). *Publications from USDA-ARS / UNL Faculty*. 1586.

<https://digitalcommons.unl.edu/usdaarsfacpub/1586>

This Article is brought to you for free and open access by the U.S. Department of Agriculture: Agricultural Research Service, Lincoln, Nebraska at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Publications from USDA-ARS / UNL Faculty by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

---

## Authors

David M. Spooner, Jorge Nunez, Guillermo Trujillo, Maria del Rosario Herrera, Frank Guzman, and Marc Ghislain

# Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification

David M. Spooner<sup>\*†</sup>, Jorge Núñez<sup>‡</sup>, Guillermo Trujillo<sup>‡</sup>, María del Rosario Herrera<sup>‡</sup>, Frank Guzmán<sup>‡</sup>, and Marc Ghislain<sup>‡</sup>

<sup>\*</sup>Vegetable Crops Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706-1590; and <sup>‡</sup>Applied Biotechnology Laboratory, International Potato Center, Apartado 1558, Lima 12, Peru

Communicated by S. J. Peloquin, University of Wisconsin, Madison, WI, October 16, 2007 (received for review August 15, 2007)

Contrasting taxonomic treatments of potato landraces have continued over the last century, with the recognition of anywhere from 1 to 21 distinct Linnean species, or of Cultivar Groups within the single species *Solanum tuberosum*. We provide one of the largest molecular marker studies of any crop landraces to date, to include an extensive study of 742 landraces of all cultivated species (or Cultivar Groups) and 8 closely related wild species progenitors, with 50 nuclear simple sequence repeat (SSR) (also known as microsatellite) primer pairs and a plastid DNA deletion marker that distinguishes most lowland Chilean from upland Andean landraces. Neighbor-joining results highlight a tendency to separate three groups: (i) putative diploids, (ii) putative tetraploids, and (iii) the hybrid cultivated species *S. ajanhuiri* (diploid), *S. juzepczukii* (triploid), and *S. curtilobum* (pentaploid). However, there are many exceptions to grouping by ploidy. Strong statistical support occurs only for *S. ajanhuiri*, *S. juzepczukii*, and *S. curtilobum*. In combination with recent morphological analyses and an examination of the identification history of these collections, we support the reclassification of the cultivated potatoes into four species: (i) *S. tuberosum*, with two Cultivar Groups (Andigenum Group of upland Andean genotypes containing diploids, triploids, and tetraploids, and the Chilotanum Group of lowland tetraploid Chilean landraces); (ii) *S. ajanhuiri* (diploid); (iii) *S. juzepczukii* (triploid); and (iv) *S. curtilobum* (pentaploid). For other classifications, consistent and stable identifications are impossible, and their classification as species is artificial and only maintains the confusion of users of the gene banks and literature.

cultivated | microsatellites | sect. *Petota* | *Solanum tuberosum* | taxonomy

The cultivated potato represents one of the most important food plants worldwide, yet interpretation of its gene pool structure remains controversial. Contrasting taxonomic treatments of the landraces have continued over last century, with the recognition of anywhere from 1 to 21 distinct Linnean species, or of various Cultivar Groups within the single species *S. tuberosum* (1). For consistency in usage in our article and to maintain the names most familiar to scientists, we use the seven species terminology of Hawkes (2). Indigenous cultivated (landrace) potatoes are widely distributed in the Andes from western Venezuela, south to northern Argentina, and with another set of landraces in south-central Chile in Chiloé Island and the adjacent Chonos Archipelago. The Chilean landraces, although once proposed to have arisen independently from central Chile (3), are secondarily derived from the Andean ones (2), likely after hybridization with the Bolivian and Argentinean species *Solanum berthaultii* (4), a species recently combined with the formerly recognized wild species *S. tarijense* (5). Three of the Andean-cultivated species are hypothesized to be of hybrid origins with cultivated potatoes and wild species: *S. ajanhuiri* [*S. stenotomum* cultivated  $\times$  *S. megistacrolobum* wild (6)], *S. juzepczukii* [*S. stenotomum*  $\times$  *S. acaule* wild (7, 8)], and *S. curtilobum* [*S. andigenum* cultivated  $\times$  *S. juzepczukii* (7, 8)]. The latter three “bitter potatoes” are grown in upland habitats and are not

grown nearly as extensively as *S. tuberosum*, as outlined by Huamán and Spooner (1).

The relationships and extent of genetic differentiation between the Andean and Chilean landraces has long been controversial. Based on cytoplasmic sterility factors, geographical isolation, and ecological differences, Grun (9) suggested that Chilean landraces were distinct from Andean landraces. Hawkes (2) distinguished the tetraploid Chilean from Andean landraces by characters of the leaf and flower pedicel. Plastid restriction site data documented five genotypes (A, C, S, T, and W types) in the diploid and tetraploid Andean landraces, and the Chilean landraces had three types, A, T, and W (10, 11). The most frequently observed type in the Chilean landraces (21 of 24 or 87.5% of the accessions examined) is type T, which is characterized by a 241-bp deletion (12). Conversely, 5 of the 113 (4.4%) accessions of *S. tuberosum* subsp. *andigenum* had the T type (10–12).

Potato landraces have been classified into 21 species (13, 14), 7 species with seven subspecies (2) and 9 species with two subspecies (15, 16), or as the single species *S. tuberosum* with 8 user-defined Cultivar Groups (1). Cultivar Groups are taxonomic categories used by the International Code of Nomenclature of Cultivated Plants to associate cultivated plants with traits that are of use to agriculturists and are not meant to represent natural groups or species in any classification philosophy. Ploidy levels in cultivated potatoes range from diploid ( $2n = 2x = 24$ ), to triploid ( $2n = 3x = 36$ ), to tetraploid ( $2n = 4x = 48$ ), to pentaploid ( $2n = 5x = 60$ ). Huamán and Spooner (1) examined the morphological support for the various classifications of potato landraces using representatives of all seven species from the classification of Hawkes (2). The results showed some morphological support for *S. ajanhuiri*, *S. chaucha*, *S. curtilobum*, and *S. juzepczukii*, lesser support for *S. tuberosum* subsp. *tuberosum*, and no support for *S. phureja* and *S. stenotomum*. Whatever morphological support for these entities was present was only by using a suite of characters, all of which are shared with other taxa (polythetic support). These results, combined with their likely hybrid origins, multiple origins, and evolutionary dynamics of continuing hybridization, led Huamán and Spooner (1) to recognize all landrace populations of cultivated potatoes as a single species, *S. tuberosum*, with the eight Cultivar Groups: Ajanhuiri Group, Andigenum Group, Chaucha Group, Chilotanum Group, Curtilobum Group, Juzepczukii Group, Phureja Group, and Stenotomum Group (the latter containing all landraces of the Goniocalyx Group).

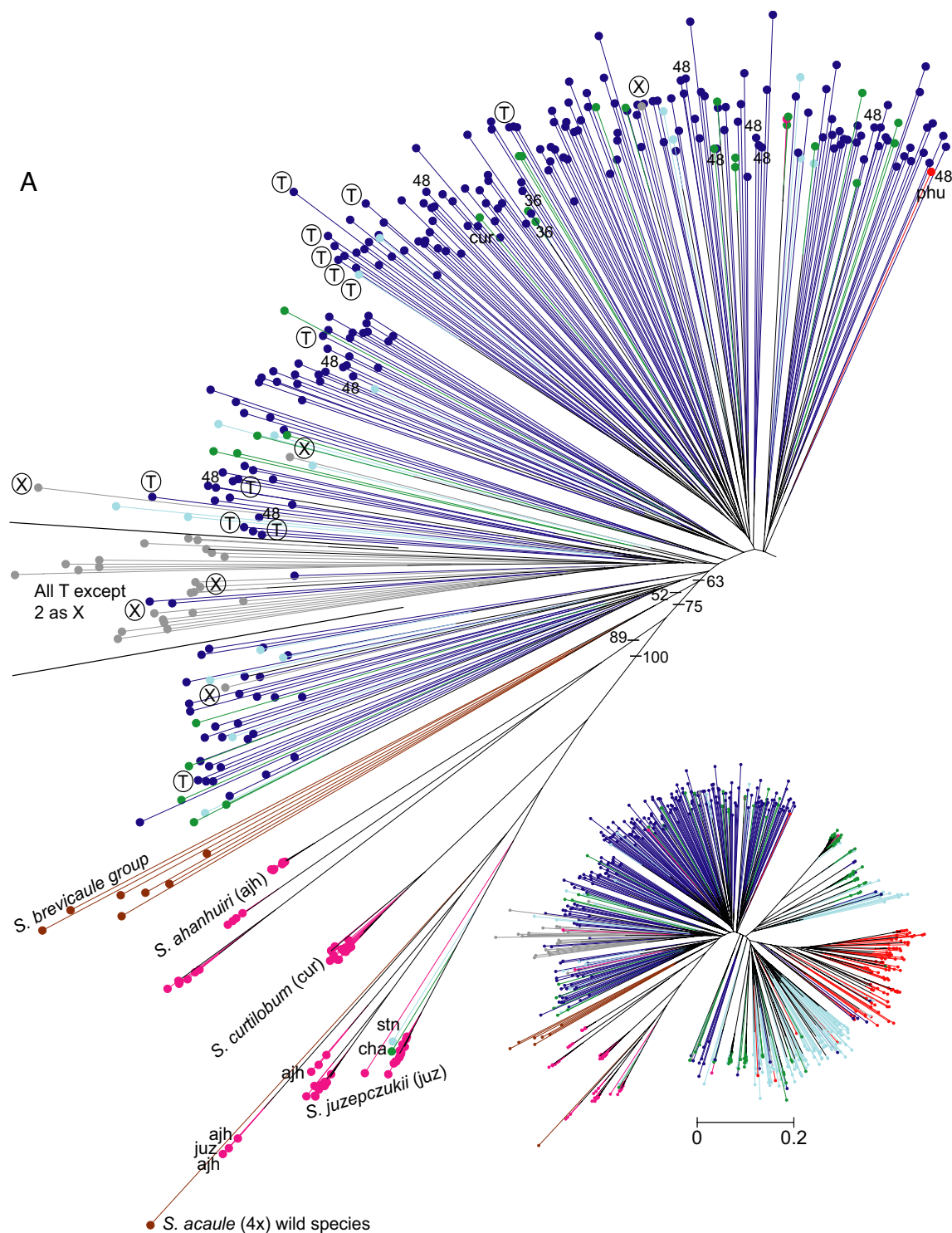
Author contributions: J.N., G.T., M.d.R.H., F.G., and M.G. performed research; and D.M.S. wrote the paper.

The authors declare no conflict of interest.

<sup>†</sup>To whom correspondence should be addressed. E-mail: david.spooner@ars.usda.gov.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0709796104/DC1](http://www.pnas.org/cgi/content/full/0709796104/DC1).

© 2007 by The National Academy of Sciences of the USA



**Fig. 1.** (Figure continues on the opposite page.)

The wild relatives of these landraces (*Solanum* section *Petota*) are all tuber-bearing and include  $\approx 190$  wild species that are widely distributed in the Americas from the southwestern United States to southern Chile (17); they possess all ploidy levels of the cultivars, as well as hexaploids ( $2n = 6x = 72$ ). Spooner *et al.* (18) studied the origin of the landraces *S. tuberosum* subsp. *andigenum*, *S. tuberosum* subsp. *tuberosum*, *S. phureja*, and *S. stenotomum* with amplified fragment length polymorphisms (AFLPs). They discovered that (i) in contrast to all prior hypotheses, these species were shown to have

a monophyletic origin; (ii) the wild species progenitors were from a group of very similar wild potato species classified in the *Solanum brevicaulle* complex; and (iii) the landraces had their origin in the highlands of southern Peru.

Although ploidy has been a major feature to define the cultivated species, many cultivated potato germplasm collections lack chromosome numbers, and many assumptions of ploidy are likely in error. For example, Ghislain *et al.* (19) used simple sequence repeats (SSRs) (also known as microsatellites) to

**Pink** - *Solanum ajanhuiri* (ajh, 2x), *S. juzepczukii* (juz, 3x), *S. curtilobum* (cur, 5x)  
**Dark blue** - *S. tuberosum* subsp. *andigenum* (adg, 4x)  
**Green** - *S. chaucha* (cha, 3x)  
**Gray** - *S. tuberosum* subsp. *tuberosum* (tub, 4x)  
**Red** - *S. phureja* (phu, 2x)  
**Light blue** - *S. stenotomum* (stn, 2x)

**Fig. 1.** Jaccard's tree based on a dissimilarity matrix of 742 potato landraces and 8 wild species examined with 50 microsatellite primer pairs. (*Inset*) Overall view of the entire tree. (A) Detail of the left-hand side of this tree, which contains mostly putative polyploid accessions (the "polyploid cluster") and accessions of *S. ajanhuiri*, *S. curtilobum*, *S. juzepczukii*, and wild species. (B) Mostly putative diploid accessions and some triploids (the "diploid cluster"). Chromosome numbers (36 and 48) are labeled for accessions not 24 within the main red *S. phureja* cluster (in red) on the left right side of the tree or previously identified as *S. phureja* but falling outside of the red cluster. The circled "T" (possessing a 241-bp deletion) and "X" (no deletion) designate accessions unexpected to occur in these clusters because they either lack the T cytoplasm thought to be largely confined to lowland Chile (area in the wedge) or lack this deletion (area outside of the wedge containing mostly accessions from Venezuela to Argentina).





Groups except for the bitter potato species *S. ajanhuiri*, *S. curtilobum*, and *S. juzepczukii*. As shown by Huamán and Spooner (1), most traditionally recognized cultivated potato species have little morphological support, and then only by using a suite of characters, all of which are shared with other taxa (polythetic support).

The International Potato Center has collected cultivated potatoes for 30 years and has invested tremendous effort in their identification. An examination of identification records at the International Potato Center shows many changes over the years, further showing the lack of stability of any character set to reliably define most cultivated species.

Potato gene banks are in great need of an integrated and comprehensive program of ploidy determinations; controlled and replicated studies of tuber dormancy (which we suspect will highlight grades of dormancy, not the present/absent determinations that exist today); photographically documented determinations of tuber and flesh colors and tuber shapes; and determinations of tuber pigments, glycoalkaloid contents, carbohydrates, proteins, amino acids, minerals, and secondary metabolites, using functional genomics approaches, with all data publicly integrated into a readily searchable web-based bioinformatics database. Such a multicomponent system will serve the breeding community much better than the outdated, unstable, and phylogenetically indefensible traditional classifications that exist today.

## Materials and Methods

**Plant Materials.** A total of 742 potato landraces of all cultivated potato species were examined: *S. tuberosum* subsp. *andigenum*, putatively tetraploid (251 accessions); *S. ajanhuiri*, diploid (22); *S. chaucha*, triploid (151 accessions); *S. tuberosum* subsp. *tuberosum*, tetraploid (27 accessions); *S. curtilobum*, pentaploid (21 accessions); *S. juzepczukii*, triploid (35 accessions); *S. phureja*, diploid (104 accessions); *S. stenotomum*, diploid (131 accessions); 7 diploid wild species accessions in the northern *S. brevicaulle* complex *S. ambosinum* Ochoa (1 accession), *S. bukasovii* Juz. (4 accessions), and *S. multiinterruptum* Bitter (2 accessions); and the wild tetraploid species *S. acaule* Bitter (1 accession) (750 accessions in total with the 8 wild species). Selection of these wild species is based on recent amplified fragment length polymorphism (AFLP) studies that documented the northern *S. brevicaulle* complex wild species to be the progenitors of the cultivated potatoes and *S. acaule* believed to be a wild species parent in the hybrid species *S. juzepczukii* and *S. curtilobum*. We qualify landrace collection ploidy as “putative” because only *S. phureja* has been counted in detail (19), that showed extensive examples of incorrect assumptions of ploidy as discussed above. Data of these accessions that includes International Potato Center accession number, taxonomic identification, ploidy when known, locality of collection, and average number of SSR alleles per accession are available as a [supporting information \(SI\) Dataset](#).

**DNA Extraction, SSR Primers, PCR Conditions, and Electrophoresis.** Genomic DNA was obtained by using standard protocols at the International Potato Center (35). DNA concentration was calculated by using PicoGreen dsDNA quantitation reagent (Molecular Probes) and a TBS-380 Fluorometer (Turner BioSystems). DNA dilutions were performed to achieve a final concentration of 3 ng/μl, using 96-well plates. We used 50 nuclear SSRs (see [SI Dataset](#)) screened from 88 that included the 22 from the Potato Genetic Identity (PGI) kit (20), 13 from ESTs developed at the Scottish Crop Research Institute (36), 30 identified by using the

potato EST database at TIGR, and 23 from the University of Idaho (37). PCR reactions were performed in a 10-μl volume containing 100 mM Tris-HCl (Sigma), 20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (Merck), 2.5 mM MgCl<sub>2</sub> (Merck), 0.2 mM each dNTP (Amersham Biosciences), 0.3 μM labeled M13 forward primer (LI-COR IRDye 700 or 800), 0.3 μM M13-tailed SSR forward primer (Invitrogen), 0.2 μM SSR reverse primer (Invitrogen), 1 unit of Taq polymerase (GIBCO/BRL), and 15 ng of genomic DNA. PCR was carried out in a PTC-200 thermocycler (MJ Research). The program used was the following: 4 min at 94°C, followed by 33 cycles of 50 sec at 94°C, 50 sec at annealing temperature (*T*<sub>a</sub>), and 1 min at 72°C, then 4 min at 72°C as a final extension step. PCR products were separated by electrophoresis on a LI-COR 4300 DNA analyzer system. The molecular weight ladder was the LI-COR IRDye 50–350 bp size standard and was loaded into gel each eight samples.

**SSR Allele Scoring.** SSR alleles were detected and scored by using SAGA Generation 2 software (LI-COR). Size calibration and an SSR “smiling line” were performed by using the molecular weight ladder (LI-COR IRDye 50–350). The SSR alleles were determined for size in bp of the upper band of the allele and scored as present (1) or absent (0). Missing data were scored as “9.”

**Data Analysis.** Genetic analysis was performed by using the program DARwin (38). A dissimilarity matrix was calculated by using Jaccard’s coefficient, 60% of minimal proportion of valid data required for each unit pair, and 500 replicate bootstrapping. The dendrogram was built by using the NJ method, using the seven wild species accessions in the northern *S. brevicaulle* complex as outgroup. The NJ method developed by Saitou and Nei (39) estimates phylogenetic trees. Although based on the idea of parsimony (it does yield relatively short estimated evolutionary trees), the NJ method does not attempt to obtain the shortest possible tree for a set of data. Rather, it attempts to find a tree that is usually close to the true phylogenetic tree (40). This method allows the rooting of trees on outgroups (in this case, the seven accessions of the *S. brevicaulle* complex). The polymorphic information content (PIC) was calculated as  $PIC = 1 - \sum(p_i^2)$ , where  $p_i$  is the frequency of the *i*th allele detected in all accessions (41). Data of somatic chromosome counts for accessions of *S. phureja* were obtained from Ghislain *et al.* (19).

**Plastid DNA Polymorphism Detection.** The 241-bp deletion was analyzed for all 742 cultivated accessions by using the primers from ref. 23. PCR amplification was performed in a volume of 10 μl consisting of 18 ng of genomic DNA, 0.4 μM each of primers (Invitrogen), 1× PCR buffer (PerkinElmer), 2.5 mM MgCl<sub>2</sub> (PerkinElmer), 200 μM each dNTP (Amersham Biosciences), and 0.25 unit of Taq DNA polymerase (GIBCO/BRL). Thermal cycling was carried out in a PTC-200 thermocycler (MJ Research) (one cycle of 4 min at 94°C, followed by 40 cycles of 45 sec at 94°C, 45 sec at 59°C, and 45 sec at 72°C, then terminated with one cycle of 4 min at 72°C). PCR products were separated by electrophoresis in a 1% agarose gel, and lambda phage digested by PstI was used as a molecular weight marker. The 241-bp plastid polymorphism was determined for size in bp and scored as “T” (≈200 bp for deleted type) and “X” (≈440 bp for undeleted type).

We thank David Douches and Lynn Bohs for comments on an earlier draft of the manuscript. This work was supported by the International Potato Center, U.S. Department of Agriculture, Generation Challenge Program Grant SP1C2-2004-5, and National Science Foundation Grant DEB 0316614 entitled “A world-wide treatment of *Solanum*” (to D.M.S.).

1. Huamán Z, Spooner DM (2002) *Am J Bot* 89:947–965.
2. Hawkes JG (1990) *The Potato: Evolution, Biodiversity, and Genetic Resources* (Belhaven, London).
3. Ugent D, Dillehay T, Ramirez C (1987) *Econ Bot* 41:17–27.

4. Hosaka K (2003) *Am J Potato Res* 80:21–32.
5. Spooner DM, Fajardo D, Bryan GJ (2007) *Taxon* 56:987–999.
6. Johns T, Huamán Z, Ochoa CM, Schmiediche PE (1987) *Syst Bot* 12:541–552.
7. Hawkes JG (1962) *Z Pflanzenzuecht* 47:1–14.

8. Schmiediche PE, Hawkes JG, Ochoa CM (1982) *Euphytica* 31:395–707.
9. Grun P (1990) *Econ Bot* 44(Suppl 3):39–55.
10. Hosaka K, Hanneman RE, Jr (1988) *Theor Appl Genet* 76:333–340.
11. Hosaka K, de Zoeten GA, Hanneman RE, Jr (1988) *Theor Appl Genet* 75:741–745.
12. Kawagoe Y, Kikuta Y (1991) *Theor Appl Genet* 81:13–20.
13. Bukasov SM (1971) in *Flora of Cultivated Plants*, eds Bukasov SM (Kolos, Leningrad, Russia), Vol 9, pp 5–40.
14. Lechnovich VS (1971) in *Flora of Cultivated Plants*, eds Bukasov SM (Kolos, Leningrad, Russia), Vol 9, pp 41–302.
15. Ochoa CM (1990) *The Potatoes of South America: Bolivia* (Cambridge Univ Press, Cambridge, UK).
16. Ochoa CM (1999) *Las Papas de Sudamérica: Perú, Part 1* (International Potato Center, Lima, Peru).
17. Spooner DM, Salas A (2006) in *Handbook of Potato Production, Improvement, and Post-Harvest Management* (Haworth, Binghamton, NY), pp 1–39.
18. Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005) *Proc Natl Acad Sci USA* 120:14694–14699.
19. Ghislain M, Andrade D, Rodríguez F, Hijmans R, Spooner DM (2006) *Theor Appl Genet* 113:1515–1527.
20. Ghislain M, Spooner DM, Rodríguez F, Villamón F, Núñez J, Vásquez C, Waugh R, Bonierbale M (2004) *Theor Appl Genet* 108:881–890.
21. Raker C, Spooner DM (2002) *Crop Sci* 42:1451–1458.
22. Ríos D, Ghislain M, Rodríguez F, Spooner DM (2007) *Crop Sci* 47:1271–1280.
23. Hosaka K (2002) *Am J Potato Res* 79:119–123.
24. Huamán Z, Golmirzaie A, Amoros W (1997) in *Biodiversity in Trust: Conservation and Use of Plant Genetic Resources in CGIAR Centres*, eds Fuccillo D, Sears L, Stapleton P (Cambridge Univ Press, Cambridge, UK), pp 21–28.
25. Hosaka K (2004) *Am J Potato Res* 81:153–158.
26. Ochoa CM (1958) *Expedición Colectora de Papas Cultivadas a la Cuenca del Lago Titicaca* (Ministerio de Agricultura, Lima, Perú), Vol 1.
27. Jackson MT, Hawkes GJ, Rowe PR (1980) *Euphytica* 29:107–113.
28. Brush SB, Carney HJ, Huamán Z (1981) *Econ Bot* 35:70–88.
29. Johns T, Keen SL (1986) *Econ Bot* 40:409–424.
30. Quiros CF, Brush SB, Douches DS, Zimmerer KS, Huestis G (1990) *Econ Bot* 44:254–266.
31. Quiros CF, Ortega R, Van Raamsdonk L, Herrera-Montoya M, Cisneros P, Schmidt E, Brush SB (1992) *Genet Res Crop Evol* 39:107–113.
32. Zimmerer K (1991) *J Biogeogr* 18:165–178.
33. Bukasov SM (1939) *Physis (Buenos Aires)* 18:41–46.
34. Hawkes JG, Hjerting JP (1989) *The Potatoes of Bolivia: Their Breeding Value and Evolutionary Relationships* (Oxford Univ Press, Oxford).
35. Ghislain M, Zhang D, Herrera M, eds (1997) *Molecular Biology Laboratory Protocols: Plant Genotyping*, Genetic Resources Department Training Manual (International Potato Center, Lima, Peru).
36. Milbourne D, Meyer R, Collins A, Ramsay L, Gebhardt C, Waugh R (1998) *Mol Gen Genet* 259:233–245.
37. Feingold S, Lloyd J, Norero N, Bonierbale M, Lorenzen J (2005) *Theor Appl Genet* 111:456–466.
38. Perrier X, Jacquemoud-Collet JP (2006) DARwin (<http://darwin.cirad.fr/darwin>), Version 5.0.
39. Saitou N, Nei M (1987) *Mol Biol Evol* 4:406–425.
40. Rohlf FJ (1997) *NTSYSpc: Numerical Taxonomy and Multivariate Analysis System* (Exeter Software, Setauket, NY), Version 2.0.
41. Nei M (1973) *Proc Natl Acad Sci USA* 70:3321–3323.