



Genetic diversity of the Andean tetraploid cultivated potato (*Solanum tuberosum* L. subsp. *andigena* Hawkes) evaluated by chloroplast and nuclear DNA markers

Sukhotu, Thitaporn

Kamijima, Osamu

Hosaka, Kazuyoshi

(Citation)

Genome, 48(1):55-64

(Issue Date)

2005

(Resource Type)

journal article

(Version)

Version of Record

(URL)

<https://hdl.handle.net/20.500.14094/90000106>



Genetic diversity of the Andean tetraploid cultivated potato (*Solanum tuberosum* L. subsp. *andigena* Hawkes) evaluated by chloroplast and nuclear DNA markers¹

Thitaporn Sukhotu, Osamu Kamijima, and Kazuyoshi Hosaka

Abstract: Andigena potatoes (*Solanum tuberosum* L. subsp. *andigena* Hawkes) ($2n = 4x = 48$) are native farmer-selected important cultivars that form a primary gene pool of the common potato (*Solanum tuberosum* L. subsp. *tuberosum*). The genetic diversity of 185 Andigena accessions and 6 Chilean native potatoes (*S. tuberosum* subsp. *tuberosum*) was studied using chloroplast DNA (ctDNA) microsatellites and nuclear DNA (nDNA) restriction fragment length polymorphism (RFLP) markers. Andigena potatoes had 14 ctDNA haplotypes and showed higher variability in the central Andes, particularly in Bolivia, whereas those in the northern regions of the distribution area were remarkably uniform with A1 ctDNA and Chilean subsp. *tuberosum* with T ctDNA. Most of 123 clearly scored RFLP bands using 30 single-copy probes were randomly distributed throughout the distribution area and proved the same gene pool shared among these widely collected accessions. Nevertheless, the geographic trend of the nDNA differentiation from north to south along the Andes and the correlated differentiation between nDNA and ctDNA ($r = 0.120$) could also be revealed by canonical variates analysis. These results suggest that the genetic diversity in Andigena was brought about primarily from cultivated diploid species but considerably modified through sexual polyploidization and intervarietal and (or) introgressive hybridization and long-distance dispersal of seed tubers by humans.

Key words: Andigena, chloroplast DNA, nuclear DNA, RFLP, geographic trend, diversity.

Résumé : Les pommes de terre Andigena (*Solanum tuberosum* L. subsp. *andigena* Hawkes, $2n = 4x = 48$) sont des variétés locales, sélectionnées par les agriculteurs, et constituent le réservoir primaire de ressources génétiques pour la pomme de terre cultivée (*S. tuberosum* subsp. *tuberosum*). La diversité génétique au sein de 185 accessions Andigena et de 6 accessions de pommes de terre (*S. tuberosum* subsp. *tuberosum*) du Chili a été examinée à l'aide de microsatellites chloroplastiques et de marqueurs RFLP nucléaires. Les pommes de terre Andigena comptaient 14 haplotypes chloroplastiques et montraient plus de variabilité dans les Andes centrales, particulièrement en Bolivie. Par contre, les accessions provenant des régions plus nordiques étaient étonnement uniformes, présentant toutes l'haplotype A1, alors que les *S. tuberosum* subsp. *tuberosum* du Chili avaient un haplotype chloroplastique T. La plupart des 123 bandes RFLP identifiables sans ambiguïté avec 30 sondes à simple copie étaient réparties aléatoirement au sein de l'aire de distribution. Cela démontre qu'un même bassin de gènes est commun à l'ensemble de cette vaste collection d'accessions. Néanmoins, une tendance à la différenciation géographique de l'ADN nucléaire le long de l'axe nord-sud des Andes ainsi qu'une différenciation corrélée entre les ADN nucléaires et chloroplastiques ($r = 0,120$) a été révélée par analyse canonique des variables. Ces résultats suggèrent que la diversité génétique chez le groupe Andigena provient largement des espèces diploïdes cultivées, mais que ce bagage génétique a été modifié considérablement suite à de la polyploïdisation sexuée, des hybridations intervariétales ou introgressives ainsi que par la dispersion sur des grandes distances des tubercules de semence par les humains.

Mots clés : Andigena, ADN chloroplastique, RFLP de l'ADN nucléaire, tendance géographique, diversité.

[Traduit par la Rédaction]

Received 11 May 2004. Accepted 15 September 2004. Published on the NRC Research Press Web site at <http://genome.nrc.ca> on 9 February 2005.

Corresponding Editor: J.H. de Jong.

T. Sukhotu.² Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe 657-8501, Japan.

O. Kamijima. Faculty of Agriculture, Kobe University, Rokkodai, Nada, Kobe 657-8501, Japan.

K. Hosaka.³ Food Resources Education and Research Center, Kobe University, 1348 Uzurano, Kasai, Hyogo 675-2103, Japan.

¹The Hawkes (1990) classification system is tentatively adopted throughout the text.

²Present address: Genetics Conservation Project under the Royal Initiative of her Royal Highness Princess Maha Chakri Sirindhorn, Chitralada Palace Dusit, Bangkok 10303, Thailand.

³Corresponding author (e-mail: hosaka@kobe-u.ac.jp).

Introduction

Andigena potatoes are native farmer-selected tetraploid cultivars (*Solanum tuberosum* L. subsp. *andigena* Hawkes) ($2n = 4x = 48$) that have long been the most important staple food in the Andean highlands (2000–4000 m). Andigena potatoes are grown and tubers are formed only under short-day conditions from Mexico, Guatemala, Venezuela, and southwards along the Andes as far as northwestern Argentina. The North and Central American cultivars were most likely introduced from South America, whereas the cultivars from Colombia to Argentina were likely native (Hawkes 1990). Other tetraploid potatoes (*Solanum tuberosum* L. subsp. *tuberosum*) are grown in the coastal regions in Chile and worldwide, which are Chilean *tuberosum* and the common potato, respectively, and likely originated independently from Andigena potatoes by selection (Hawkes 1956, 1990; Brücher 1963; Hosaka and Hanneman 1988a).

Solanum stenotomum Juz. et Buk. is one of 7 cultivated potato species and believed to be the most primitive diploid species grown from central Peru to central Bolivia (Hawkes 1990). Andigena originated via polyploidization from an intervarietal or interspecies cross within cultivated diploid potatoes involving *S. stenotomum* as the main germplasm (Swaminathan and Magoon 1961; Matsubayashi 1991; Hosaka 1995) or from an interspecies cross between *S. stenotomum* and a wild diploid species, *Solanum sparsipilum* (Bitt.) Juz. et Buk. (Hawkes 1956, 1990; Cribb and Hawkes 1986).

Very wide morphological and physiological variability has been recognized in Andigena (Hawkes 1956; Simmonds 1964; Ugent 1970; Joseph and Gopal 1994). Ochoa (1990) described 13 botanical varieties and 35 forms only for Bolivian Andigena. Salaman (1946) examined leaf characters of 139 Andigena cultivars and observed that the *S. tuberosum* subsp. *tuberosum*-like forms steadily increased from 5% in Bolivia and Argentina to 27% in Peru, 65% in Ecuador, and 70% in Colombia. However, extreme overlapping of characters in Andigena makes even separation from *S. tuberosum* subsp. *tuberosum* difficult (Simmonds 1964; Hawkes 1990; Huamán and Spooner 2002).

Restriction fragment length polymorphisms (RFLPs) of chloroplast DNA (ctDNA) disclosed genetic variation within Andigena potatoes (Hosaka 1986; Hosaka and Hanneman 1988a). Five basic ctDNA types (W, T, C, S, and A) have been identified among cultivated potatoes (Hosaka 1986). Andigena had A-type ctDNA in many accessions, and 4 other ctDNA types with different frequencies varied from north to south of the Andes (Hosaka and Hanneman 1988a). Most of the ctDNA diversity in Andigena was shared with *S. stenotomum* and then overlapped partly with various presumed ancestral species (Hosaka and Hanneman 1988b; Hosaka 1995; Sukhotu et al. 2005). Hosaka (1995) proposed a hypothesis of multiple origins for *S. stenotomum* by successive domestication from the ancestral species complex (the successive domestication hypothesis). Subsequent sexual polyploidization formed wide ctDNA diversity among the Andigena potatoes (Hosaka 1995).

In this paper, we studied 185 accessions of Andigena and 6 accessions of Chilean *tuberosum* potatoes using ctDNA microsatellites and nuclear DNA (nDNA) RFLPs to understand the genetic diversity in these most important Andean

potatoes. The relevant aims that will be addressed are the geographic differentiation and coevolution of ctDNA and nDNA.

Materials and methods

Plant material and DNA isolation

One hundred eighty-five Andigena accessions, collected throughout its distribution area, were used, and were grouped as regional populations: region 1, North America (Mexico and Guatemala); region 2, northern Andes (Venezuela, Colombia, and Ecuador); region 3, Peru (except Department of Puno); region 4, Lake Titicaca (Department of Puno, Peru, and Department of La Paz, Bolivia); region 5, Bolivia (except Department of La Paz); region 6, north Argentina. Since Lake Titicaca lies between Peru and Bolivia and was often cited as the center of diversity of Andean cultivated potatoes (Simmonds 1976; Hawkes 1990), region Lake Titicaca was set as a separate region from the others. Six accessions of Chilean *tuberosum* were also included as samples from region 7, coastal Chile (Fig. 1; Table 1). Accessions with CIP numbers have all been clonally propagated at the International Potato Center, Lima, Peru, which were obtained as DNA samples (see Sukhotu et al. 2004). Accessions with PI numbers were obtained as seeds from the Potato Introduction Station (NRSP-6), Sturgeon Bay, Wis. One seedling of each accession was used, except for 4 accessions for which 2 seedlings each were used as duplicates. Total DNA was isolated from fresh leaves by the method described in Hosaka and Hanneman (1998).

DNA analysis

For determination of ctDNA types, *Bam*HI, *Hind*III, or *Pvu*II restriction fragment patterns were used as described in Hosaka and Hanneman (1988b). Seven ctDNA microsatellite markers (NTCP markers, developed by Provan et al. 1999) and the H3 marker (developed by Hosaka 2003) were used to further reveal ctDNA polymorphisms. For nDNA RFLP analysis, 30 single-copy probes were used: TG14, TG22, TG28, TG46, TG61, TG63, TG71, TG115, TG123, TG134, TG152, TG413, TG560, P41, P73, P101, P116, P140, P215, P251, P357, P368, P537, P695, P697, P769, P808, P894, P948, and P1108. The TG probes were single-copy tomato probes obtained from Dr. S.D. Tanksley, Cornell University, Ithaca, N.Y. (Tanksley et al. 1992). The P probes were random genomic clones from *Solanum phureja* clone 1.22 (Hosaka and Spooner 1992). Detection procedures have all been described previously (Sukhotu et al. 2004).

Data analysis

For the analysis of nDNA, only visibly reliable and polymorphic (presence versus absence) bands were used as RFLPs. ctDNA microsatellites and nDNA RFLPs were scored as 1/0 for each fragment. Pairwise similarities between accessions were calculated as Jaccard coefficients (*J*) (a perfect match for *J* = 1.00) separately for ctDNA microsatellites (including H3 marker data) and for nDNA RFLPs. Population means of *J* were compared by *t* test.

Since allelic relationships among RFLPs were not known, each RFLP was considered to be a single locus, and the extent of genetic variation in a population was measured by av-

Fig. 1. Geographic distribution of accessions used in this study. Only accessions that had exact locality data (180 accessions) are presented. Regional differences were analyzed between the regions shown in this map.



average gene diversity (H) (Nei and Kumar 2000). The gene diversity at a locus (h) is defined as

$$h = 1 - \sum_{i=1}^q x_i^2$$

where x_i is the frequency of the i th allele and q is the number of alleles. The H value is the average of this quantity over all loci. The H value for ctDNA microsatellites was the average of h values from all microsatellite loci and the H3 marker.

For clustering ctDNA haplotypes, the unweighted pair group method with arithmetic means (UPGMA) was used. To search for the most probable UPGMA dendrogram, 1000 bootstraps were carried out and an unrooted phylogram was obtained using PAUP 4.0b10 (Sinauer Associates, Inc., Sunderland, Mass.).

Principal components analysis (PCA) and canonical variates analysis (CVA) were performed using 1/0-type data as variables using NTSYSpc version 2.0 (Rohlf 1997).

Pearson's correlation coefficients (r) were calculated between the geographic distance, nDNA RFLPs, and ctDNA microsatellite distance matrices prepared by measurements of Euclidean distances, and the Mantel test was performed by 1000 random permutations to test whether the original correlation occurred by chance using Mantel for Windows version 1.14 (developed by Mauro J. Cavalcanti).

Results

ctDNA haplotypes

Restriction fragment patterns of ctDNA obtained in this study were similar to those reported earlier (Hosaka and Hanneman 1988b), thus disclosing the known ctDNA types (A, S, C, W, and T) that could be assigned to all accessions used. Seven ctDNA microsatellites detected single fragments

Table 1. Andigena and Chilean *tuberosum* accessions used in this study and their ctDNA haplotypes.

Haplotype	Accessions
Region 1: North America	
A1	PI 161350, PI 161683, PI 161716, PI 184903
Region 2: Northern Andes	
A1	PI 225628, PI 225633, PI 225635, PI 237208, PI 243360, PI 243400, PI 243401, PI 243405, PI 243406, PI 243407, PI 243409, PI 243411, PI 243415, PI243424, PI 243429, PI 243430, PI 243431, PI 243441, CIP 703268, CIP 704111
A2	PI 243363, PI 243434, PI 243436
T	PI 246979
Region 3: Peru	
A1	PI 186179, PI 214424, PI 214426, PI 214429, PI 214430, PI 214434, PI 214436, PI 214441, PI 214442, PI 214443, PI 232839, PI 246499, PI 246545, PI 246555, PI 281059, PI 281060, PI 281064, PI 281066, PI 281080, PI 281093, PI 292107, CIP 700017, CIP 700045, CIP 700094, CIP 700532, CIP 700616, CIP 700767, CIP 701304, CIP 703682
S	PI 232841, CIP 700387, CIP 700771, CIP 700960, CIP 701306
C1	PI 186178, PI 195162, PI 232036, PI 232840, PI 281061, PI 281105, PI 292078, PI 292096, PI 292097, CIP 700598, CIP 700652, CIP 700863, CIP 700877, CIP 700921
C2	PI 246497, PI 281088
C3	PI 292089
W1	PI 365372
W3	PI 365345
Region 4: Lake Titicaca (Department of Puno, Peru, and Department of La Paz, Bolivia)	
A1	PI 230497, PI 246516, PI 246521, PI 255505, PI 280989, PI 280990, PI 280993, PI 281008, CIP 702453
S	PI 230496, PI 498310, PI 546018
C1	PI 230499
C2	PI 258857, PI 258936
W4	PI 546017
Region 5: Bolivia	
A1	PI 197757, PI 258927, PI 281032, PI 473508, PI 498291, PI 498309, PI 546016, PI 546023, CIP 701065, CIP 701067, CIP 703474
S	PI 275119, PI 546021, CIP 701074, CIP 704082
C1	PI 258879, PI 258881, PI 258917, PI 281021, PI 281031, PI 473390, PI 473507, PI 498307, PI 545930, PI 546020, PI 546024, PI 546026
C4	PI 498294
W1	PI 258886, PI 473391, PI 546025
W2	PI 255508, PI 498076, PI 546028
W5	PI 265882
W6	PI 473393
Region 6: North Argentina	
A1	PI 255503, PI 473196, PI 473246, PI 473249, PI 473251, PI 473253, PI 473255, PI 473265, PI 473267, PI 473269, PI 473270, PI 473275, PI 473276, PI 473277, PI 473282, PI 473283, PI 473284, PI 473286, PI 473287, PI 473288, PI 473290, PI 473291, PI 473292, PI 473296, PI 473300, PI 473301, PI 500058, PI 558137, PI 558142, PI 558144, PI 558145, CIP 704152
S	PI 473259, PI 473260, PI 473261, PI 473262, PI 558139
C1	PI 473278, PI 473293, PI 473294, PI 473295, PI 473298, PI 473299, PI 473302, PI 500057
W1	PI 473254, PI 473258, PI 473268, PI 473281, PI 473285
T	PI 473257, PI 558141
Region 7: Coastal Chile (Chilean <i>tuberosum</i>)	
A1	CIP 704165
T	CIP 703254, CIP 703610, CIP 704168, CIP 704171, CIP 704172

Note: Accessions with CIP numbers were from the International Potato Center, Lima, Peru, and those with PI numbers were from the Potato Introduction Station (NRSP-6), Sturgeon Bay, Wisconsin.

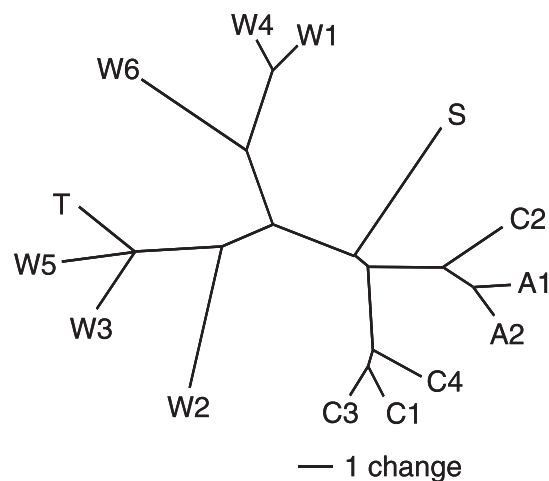
in each accession and a total of 28 different fragments were detected, while the H3 marker produced two types of restriction banding patterns (types 1 and 2; see Sukhotu et al. 2004). The combination of 30 fragments and types discriminated 14 different ctDNAs (Table 2). A-type ctDNA was

separated into 2 different ctDNAs referred to as A1 and A2 ctDNA haplotypes (suffixed numbers in order of appearance frequencies). Likewise, C-type ctDNA was separated into C1, C2, C3, and C4 and W-type ctDNA into W1, W2, W3, W4, W5, and W6 haplotypes. Both S- and T-type ctDNAs

Table 2. Characterization of ctDNA haplotypes based on ctDNA types, microsatellite fragments (base pairs), and the H3 marker.

ctDNA haplotype	ctDNA type	Microsatellite marker (NTCP) phenotype							H3 type
		6	7	8	9	12	14	18	
A1	A	174	174	250	289	237	151	186	2
A2	A	174	174	250	289	236	151	186	2
S	S	127	173	251	289	239	150	186	2
C1	C	173	174	251	289	236	150	187	2
C2	C	175	174	250	289	238	151	186	2
C3	C	173	174	251	289	236	151	187	2
C4	C	173	174	251	289	236	150	186	2
W1	W	172	172	253	247	236	151	186	1
W2	W	174	174	253	249	236	149	188	1
W3	W	173	174	253	279	235	149	188	1
W4	W	172	173	253	247	236	151	186	1
W5	W	173	174	252	279	235	150	188	1
W6	W	173	173	252	248	236	151	186	1
T	T	173	173	252	279	235	149	188	1

Note: NTCP markers reveal polymorphisms in simple sequence repeats in ctDNA (developed by Provan et al. (1999). The H3 marker produced polymorphic banding patterns by PCR amplification followed by digestion with the restriction endonuclease *DraI* (Hosaka 2003).

Fig. 2. Relationships among ctDNA haplotypes shown by an unrooted UPGMA consensus tree based on 1000 bootstraps.

were no more separated by these ctDNA microsatellites and the H3 marker.

The differences between ctDNA haplotypes are shown in Fig. 2 with a UPGMA consensus tree based on 1000 bootstrap replicates. This unrooted tree could be separated to 2 halves based on ctDNA types: a group of W- and T-type ctDNA and another of A-, S-, and C-type ctDNA, which is in good agreement with our previous studies (Sukhotu et al. 2005).

Distribution and diversity of ctDNA haplotypes

The A1 haplotype was the most frequent ctDNA in Andigena (56.8%) and the T haplotype in Chilean *tuberosum* (83.3%). The A1 ctDNA predominated in North America and northern Andes (regions 1 and 2) (Table 3). A2 ctDNA was only found in 3 Colombian accessions in region 2 (Ta-

ble 1). Toward the south of the Andes, the frequencies of A1 drastically decreased with the lowest frequency of 30.5% in Bolivia. The S ctDNA haplotype (17 accessions) was found in 18.8% of accessions from Lake Titicaca, and the within-region frequencies decreased toward the north (9.4% in Peru) and south (11.1% in Bolivia and 9.6% in north Argentina). The C1 ctDNA (35 accessions) was widely distributed in the central Andes (Regions 3–6), whereas C2 (four accessions) was limitedly distributed in Peru and Lake Titicaca and C3 and C4 (one accession each) only in Peru and Bolivia, respectively. W1 ctDNA (9 accessions) was sparsely found in the central Andes (regions 3, 5, and 6), whereas the W2–W6 haplotypes were rare in this region except in Bolivia (5 accessions). The T ctDNA haplotype, the most common ctDNA in *S. tuberosum* subsp. *tuberosum*, was found in 2 Andigena accessions from north Argentina, as similarly described in Hosaka and Hanneman (1988a) and Hosaka (2004), and in 1 Ecuadorian accession (PI 246979) (Table 1).

Regional diversity was measured by the gene diversity h for ctDNA types and by the average gene diversity H for microsatellites including the H3 marker (Table 3). Both measurements similarly indicated that levels of genetic diversity in ctDNA were very low in the northern Andes ($h = 0.00$ – 0.08 and $H = 0.00$ – 0.11), but very high in the central Andes with the highest values obtained in Bolivia ($h = 0.71$ and $H = 0.54$).

nDNA RFLPs

One hundred ninety-one accessions including 4 accessions with duplicates were analyzed using 30 single-copy RFLP probes. A total of 123 bands, or an average of 30.4 bands per sample (SD = 4.32), were clearly scored, of which nine were unique to single accessions. All samples were uniquely distinguished with an average of 26.9 bands (SD = 5.59, minimum 3 – maximum 50). Duplicates were distinguished from their counterparts by an average of 13.8 bands (SD = 4.35). There was no common band in all Chilean accessions or in all Andigena accessions.

Table 3. Geographic distribution of ctDNA haplotypes.

Region	ctDNA haplotype											ctDNA type (<i>h</i>)	Microsatellites (<i>H</i>)			
	No. of accessions	A1	A2	S	C1	C2	C3	C4	W1	W2	W3			W4	W5	W6
1. North America	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00 (0.00)
2. Northern Andes	24	20	3	0	0	0	0	0	0	0	0	0	0	0	1	0.11 (0.07)
3. Peru	53	29	0	5	14	2	1	0	1	0	1	0	0	0	0	0.38 (0.24)
4. Lake Titicaca	16	9	0	3	1	2	0	0	0	0	0	1	0	0	0	0.35 (0.22)
5. Bolivia	36	11	0	4	12	0	0	1	3	3	0	0	1	1	0	0.54 (0.17)
6. North Argentina	52	32	0	5	8	0	0	0	5	0	0	0	0	0	2	0.41 (0.14)
7. Coastal Chile	6	1	0	0	0	0	0	0	0	0	0	0	0	0	5	0.36 (0.25)
Total	191	106	3	17	35	4	1	1	9	3	1	1	1	1	8	0.43 (0.17)

Note: *h* based on ctDNA types was defined as $h = 1 - \sum_{i=1}^q x_i^2$ where x_i is the frequency of the *i*th ctDNA type and *q* is the number of ctDNA types (5) and *H* based on microsatellites was the average of *h* over 7 microsatellite loci (SD in parentheses).

An average *J* between accessions was 0.39 (SD = 0.08), which was significantly lower than that between duplicates (*J* = 0.64, SD = 0.12) (*t* = 5.750, *P* < 0.001). The average *J* between Andigena and Chilean *tuberosum* was 0.38 (SD = 0.08), which was statistically lower than that within Andigena (*J* = 0.41, SD = 0.08) (*t* = 12.105, *P* < 0.001) or within Chilean *tuberosum* (*J* = 0.44, SD = 0.13) (*t* = 2.856, *P* < 0.005). The Ecuadorian Andigena accession having T ctDNA haplotype (PI 246979) had *J* = 0.38 (SD = 0.06) to the other Andigena accessions and *J* = 0.42 (SD = 0.07) to Chilean *tuberosum*, each corresponding to the between-group similarity and within-group similarity, respectively. Thus, as hypothesized by Hosaka (2004) for this accession, it seems to be a later introduction of or a hybrid with *S. tuberosum* subsp. *tuberosum*. Consequently, this accession was omitted from further analysis.

Regional differences among 190 accessions are tabulated in Table 4. Numbers of scored bands, polymorphic bands, and endemic bands (appearing only in a certain region) were all larger in Bolivia and north Argentina, whereas their means per accession were not so large compared with those of the other regions. The average gene diversity *H* for each region ranged from 0.23 to 0.34 (Table 4). No significant difference was found between neighboring regions except for coastal Chile, which showed a significantly higher *H* than north Argentina (*t* = 3.163, *P* < 0.002). Differences within regions ranged from *J* = 0.36 for Lake Titicaca to *J* = 0.55 for North America, which were not always larger than those between regions (*J* = 0.34–0.43). In particular, *J* values within the populations from Lake Titicaca (0.36) and Bolivia (0.40) were almost at the same level as or slightly lower than those between these and the other regions. Interestingly, Chilean *tuberosum* showed the highest similarity to the accessions of the most remote location (*J* = 0.43 to North America) rather than to the closest location (*J* = 0.34 to north Argentina).

Multivariate analysis

To further reveal the nature of nDNA variation, PCA and CVA were performed using 190 accessions. The PCA plot by the first and second principal components, each explaining only 6.1% and 3.9% of the total variation, respectively, showed continuous overlapping of the accessions of Regions 2, 3, 4, 5, and 6 in that order and a rough separation between accessions with C- or S-type ctDNA and those with W-type ctDNA (plot not shown).

The CVA performed by Andigena and Chilean *tuberosum* as predefined groups significantly distinguished these 2 groups (Wilk's lambda = 0.0823, *P* (same) < 0.00001) (plot not shown). The CVA performed by regions as groups successfully disclosed a geographical cline from region 1 to region 6 by canonical variate 2 (Fig. 3) (Wilk's lambda = 0.000034, *P* < 0.00001), although regional groups were continuously overlapping. Accessions of region 7 were distinctly located on the left of canonical variate 1 axis, also indicating a good separation between Andigena and Chilean *tuberosum*.

The result of the CVA, based on nDNA RFLPs and performed by 5 ctDNA types as predefined groups, is shown in Fig. 4 (where each accession is designated by the ctDNA haplotype identity). The accessions with W- and T-type

Table 4. Regional differences revealed by nDNA RFLPs.

Region	No. of accessions	Polymorphic/ total bands	Endemic bands	H	Average J within and between regions						
					1	2	3	4	5	6	7
1. North America	4	30/44	0	0.27 (0.19)	0.55 (0.08)	0.37 (0.06)	0.40 (0.06)	0.39 (0.09)	0.40 (0.07)	0.39 (0.06)	0.43 (0.09)
2. Northern Andes	23	73/80	0	0.25 (0.18)	0.47 (0.08)	0.43 (0.08)	0.38 (0.08)	0.37 (0.07)	0.37 (0.07)	0.37 (0.07)	0.37 (0.08)
3. Peru	53	95/96	2	0.23 (0.16)		0.45 (0.08)	0.39 (0.09)	0.38 (0.08)	0.39 (0.08)	0.38 (0.07)	0.38 (0.08)
4. Lake Titicaca	16	87/91	3	0.29 (0.15)			0.36 (0.09)	0.38 (0.08)	0.36 (0.07)	0.36 (0.07)	0.37 (0.09)
5. Bolivia	36	101/104	3	0.25 (0.15)				0.40 (0.08)	0.39 (0.07)	0.42 (0.09)	0.37 (0.09)
6. North Argentina	52	99/103	5	0.26 (0.16)					0.42 (0.09)	0.34 (0.08)	0.34 (0.08)
7. Coastal Chile	6	52/57	2	0.34 (0.14)							0.42 (0.14)
Total	190	123/123	15	0.22 (0.17)							

Note: SD in parentheses. H was the average of h defined as $1 - \sum_{i=1}^q x_i^2$ where x_i was the frequency of the *i*th RFLP band and $q = 1$ because each band was assumed to be controlled by a single locus.

ctDNA haplotypes were uniquely separated, and those with C-type ctDNA were partially separated from those with A- and S-type ctDNA, which were continuously overlapped with each other (Wilk's lambda = 0.001198, $P < 0.00001$).

Correlation between distance matrices

The nDNA RFLPs and ctDNA microsatellite distance matrices using all 184 *Andigena* accessions were poorly and positively correlated ($r = 0.120$) with a statistical significance level of 1%. A geographic distance matrix was obtained from 178 *Andigena* accessions that had exact locality data. This matrix was compared with the nDNA RFLPs and ctDNA microsatellite distance matrices obtained from these limited accessions. Neither the nDNA RFLPs ($r = 0.013$) nor ctDNA microsatellites ($r = -0.101$) showed significant correlations with the geographic distance matrix, the latter of which could be expected because the A1 haplotype distributed throughout the distribution area and the diversity was not evenly distributed (Table 3).

Discussion

Genetic diversity in *Andigena* potatoes

Since *Andigena* is known as a highly polymorphic tetraploid cultigen (Hawkes 1956; Simmonds 1964; Ugent 1970; Ochoa 1990; Joseph and Gopal 1994) and can be separated from *S. stenotomum* only by the chromosome number (Huamán and Spooner 2002), it might be an assemblage of various genetically differentiated subpopulations (Ugent 1970). In our study, however, *Andigena* potatoes did not show any clearcut trends or genetically distinct subpopulations, as evaluated by nDNA RFLPs. The small number of RFLP probes used is not likely the reason because almost the same set of probes could distinguish three distinct clusters among Andean cultivated and closely related wild potato species (Sukhotu et al. 2004).

Similarity within an accession ($J = 0.64$) was certainly higher than that between accessions ($J = 0.41$), whereas those between regional populations ($J = 0.36-0.44$) were not much different from those within populations ($J = 0.38-0.56$). The high level of within-population diversity and the low level of between-population diversity at the allelic level in isozyme loci have been reported for Peruvian landraces (Zimmerer and Douches 1991; Brush et al. 1995). The average gene diversity H showed almost the same level of diversity over the distribution areas of *Andigena* from North America to north Argentina (Table 4), actually meaning that the same level of heterozygosity occurred in any sample from any region and most of alleles were likely sampled without a geographic tendency. This agrees with the Simmonds' (1964) conclusion that all characters are found in all parts of the range. The lack of apparent correlation between geographic distances and genetic differentiation strongly suggests that these accessions belonged to the same gene pool of *Andigena*, which has been maintained by frequent gene exchange by hybridization and (or) arose by common ancestry (Zimmerer and Douches 1991; Quiros et al. 1992). Whether this gene pool is a part of or an independent taxon corresponding to a species, subspecies, or other taxonomic rank not in the scope of the present study will be discussed in the future.

Fig. 3. CVA plot by first and second canonical variates based on nDNA RFLPs and performed by region as predefined groups showing regional differences. 1, North America; 2, northern Andes; 3, Peru; 4, Lake Titicaca; 5, Bolivia; 6, north Argentina; 7, coastal Chile.

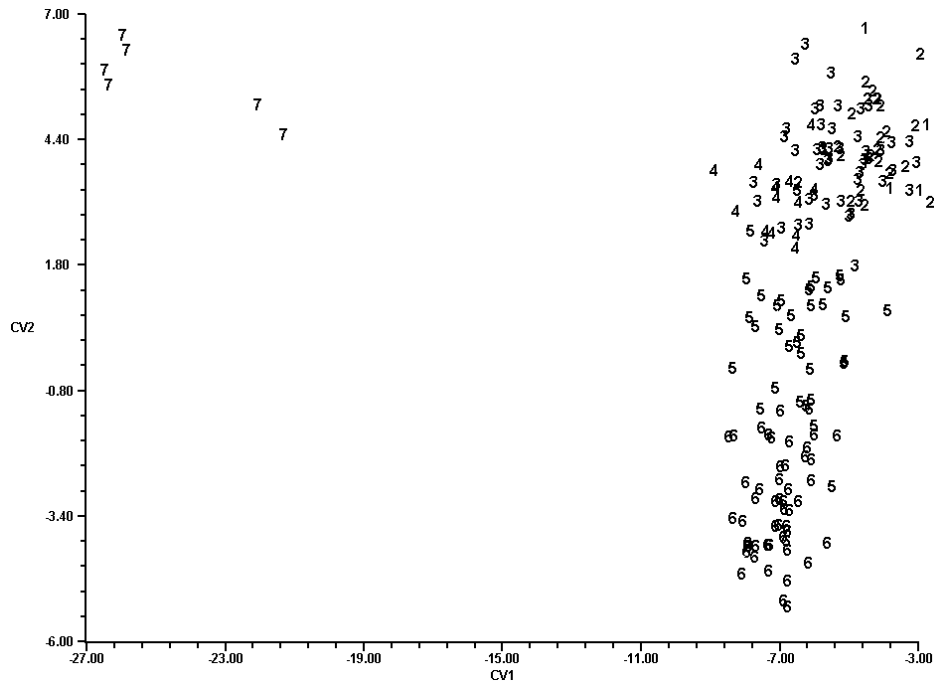
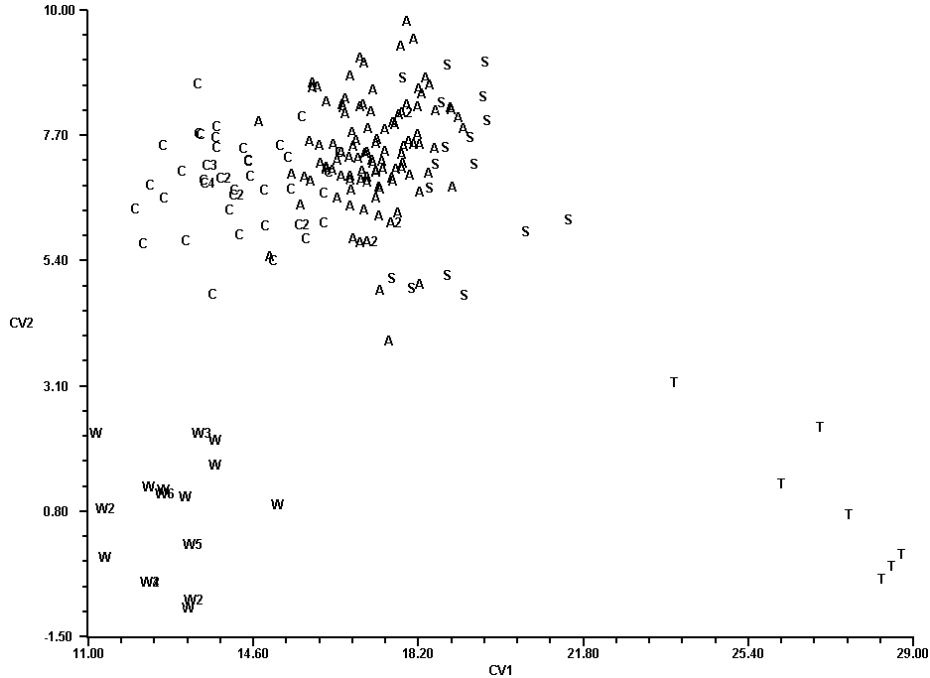


Fig. 4. CVA plot by first and second canonical variates based on nDNA RFLPs and performed by ctDNA types as predefined groups showing differences between ctDNA haplotypes. A1, C1, W1 haplotypes are represented A, C, and W, respectively.



PCA and CVA, however, could disclose a geographic trend, suggesting that some of the RFLPs did show a geographic cline. The continuous nDNA variation from the north to south of the Andes could be partly associated with morphological and physiological clines previously reported for *Andigena* populations (Salaman 1946; Simmonds 1964; Glendinning 1968; Hosaka and Hanneman 1993). A clear

geographic trend was again recognized in ctDNA variation. *Andigena* potatoes showed higher ctDNA variability in the central Andes, particularly in Bolivia, whereas those in the northern regions of the distribution area were remarkably uniform with A1 ctDNA and Chilean *tuberosum* with T ctDNA. This ctDNA difference between the 2 taxa and the geographic cline are nearly similar to those of our previous

findings (Hosaka and Hanneman 1988a). The CVA plotted by ctDNA haplotypes (Fig. 4) showed clearer separation between predefined groups than the one plotted by regions (Fig. 3). In comparisons among the geographic distance, nDNA, and ctDNA distance matrices, only the correlation between the nDNA and ctDNA matrices was significant ($r = 0.120$). This suggests that the genetic diversity in *Andigena* was more influenced by maternal parents than by geographic factors such as local climate, soil conditions, farmers agricultural practices and preferences, etc.

Origin of the genetic diversity

Solanum stenotomum, the most probable maternal ancestor of *Andigena* (Hosaka and Hanneman 1988b), likely outlined the initial genetic diversity of *Andigena*. The most predominant A1 ctDNA (or A-type ctDNA) in *Andigena* probably first arose in a putative wild ancestral species in central Peru and spread mostly in Peruvian *S. stenotomum* (Hosaka 1995). The much wider distribution of *Andigena* than *S. stenotomum* and particularly its distributional expansion towards the north seem to associate with A1 ctDNA. The S ctDNA, in turn, probably arose in wild species in southern Peru and spread throughout its distribution area of *S. stenotomum* (Sukhotu et al. 2005). The S-type ctDNA in *Andigena* was found from Peru to Bolivia, overlapping with the distribution area of *S. stenotomum*, and southward to north Argentina. This indicates that initially and locally occurred new types of ctDNA were disseminated from wild species to *S. stenotomum* and then to *Andigena*.

Hybridization occurs frequently in cultivated potato fields and natural hybrids are unintentionally incorporated into the cultivated potato gene pool (Ugent 1970; Jackson et al. 1977, 1980; Huamán et al. 1980; Johns and Keen 1986; Rabinowitz et al. 1990; Zimmerer and Douches 1991). Furthermore, Andean native farmers occasionally use true botanical seeds to renovate degenerated tubers because of viral infection (Quiros et al. 1992). Therefore, it is suggested that although the genetic diversity of *Andigena* was outlined initially by *S. stenotomum*, the present genetic diversity in *Andigena* was considerably modified and attenuated through sexual polyploidization and intervarietal and (or) introgressive hybridization after *Andigena* arose and more likely by long-distance dispersal of seed tubers by humans.

Further differentiation in *Andigena*

Chilean *tuberosum*, irrespective of its ctDNA (T or A1), was clearly distinguished from *Andigena* potatoes by nDNA RFLPs, although a very small number of Chilean *tuberosum* accessions were used, since this study was not designed to evaluate the relationship between the 2 taxa. The present results agree with those of Raker and Spooner (2002), who successfully separated Chilean *tuberosum* from *Andigena* potatoes using microsatellite markers of nDNA. Recently, Hosaka (2003) traced T ctDNA, the most common ctDNA of Chilean *tuberosum*, to some populations of a wild diploid species *Solanum tarijense* Hawkes and suggested that Chilean *tuberosum* originated by selection from natural hybrids between *S. tarijense* as females and *Andigena* as males (Hosaka 2004). Thus, the distinctiveness of Chilean *tuberosum* might be due to incorporated germplasm from *S. tarijense*.

Compared with diploid cultigens, *Andigena* potatoes extended the growing area much farther south, which accompanied various haplotypes of W-type ctDNA in Bolivia and T ctDNA in north Argentina (Table 3). Likewise with T ctDNA incorporated from *S. tarijense*, this W-type ctDNA might also be incorporated by introgressive hybridization from Bolivian and Argentine wild species, most of which have W-type ctDNA (Nakagawa and Hosaka 2002; Sukhotu et al. 2004). The importance of introgressive hybridization and (or) polyploidization from any wild and cultivated species via $2x \times 4x$ or $4x \times 2x$ crosses has been emphasized to broaden the genetic diversity and adaptability of *Andigena* (Ugent 1970; Bukasov 1978; Hosaka and Hanneman 1988b; Grun 1990). Alternatively, according to the successive domestication hypothesis (Hosaka 1995), *Andigena* potatoes having W-type ctDNA are thought to be the most primitive varieties. To solve the controversy on the conflicting hypotheses as to the W-type ctDNA and the species origin of *Andigena*, evolutionary relationships of *Andigena* with closely related species including various hypothesized ancestral species should be investigated in a future study.

Acknowledgements

We thank the US gene bank at Sturgeon Bay, Wisconsin, and the CIP gene bank at Lima, Peru, for providing seed and DNA samples. We also thank our technical staff M. Miyake, S. Kakehi, H. Hashizume, M. Fujimatsu, and S. Kitanishi for plant maintenance.

References

- Brücher, H. 1963. Untersuchungen über die *Solanum (Tuberosum)* — Cultivare der Insel Chiloé. *Z. Pflanzenzücht.* **49**: 7–54.
- Brush, S., Kesseli, R., Ortega, R., Cisneros, P., Zimmerer, K., and Quiros, C. 1995. Potato diversity in the Andean center of crop domestication. *Conserv. Biol.* **9**: 1189–1198.
- Bukasov, S.M. 1978. Systematics of the potato. *Bull. Appl. Bot. Genet. Breed.* **62**: 1–42.
- Cribb, P.J., and Hawkes, J.G. 1986. Experimental evidence for the origin of *Solanum tuberosum* subspecies *andigena*. In *Solanaceae: biology and systematics*, Columbia University Press, New York. pp. 383–404.
- Glendinning, D.R. 1968. Regional variation in leaf form and other characters of *Solanum tuberosum* Group *Andigena*. *Eur. Potato J.* **11**: 277–280.
- Grun, P. 1990. The evolution of cultivated potatoes. *Econ. Bot.* **44**(Suppl. 3): 39–55.
- Hawkes, J.G. 1956. Taxonomic studies on the tuber-bearing *Solanums*. 1. *Solanum tuberosum* and the tetraploid complex. *Proc. Linn. Soc. Lond.* **166**: 97–144.
- Hawkes, J.G. 1990. The potato — evolution, biodiversity and genetic resources. Belhaven Press, London, UK.
- Hosaka, K. 1986. Who is the mother of the potato? — restriction endonuclease analysis of chloroplast DNA of cultivated potatoes. *Theor. Appl. Genet.* **72**: 606–618.
- Hosaka, K. 1995. Successive domestication and evolution of the Andean potatoes as revealed by chloroplast DNA restriction endonuclease analysis. *Theor. Appl. Genet.* **90**: 356–363.
- Hosaka, K. 2003. T-type chloroplast DNA in *Solanum tuberosum* L. ssp. *tuberosum* was conferred from some populations of *S. tarijense* Hawkes. *Am. J. Potato Res.* **80**: 21–32.

- Hosaka, K. 2004. Evolutionary pathway of T-type chloroplast DNA in potato. *Am. J. Potato Res.* **81**: 155–160.
- Hosaka, K., and Hanneman, R.E., Jr. 1988a. The origin of the cultivated tetraploid potato based on chloroplast DNA. *Theor. Appl. Genet.* **76**: 172–176.
- Hosaka, K., and Hanneman, R.E., Jr. 1988b. Origin of chloroplast DNA diversity in the Andean potatoes. *Theor. Appl. Genet.* **76**: 333–340.
- Hosaka, K., and Hanneman, R.E., Jr. 1993. Effects of different chloroplast DNA types on tuber yield in the Andean and Chilean tetraploid potatoes — preliminary report. *Sci. Rep. Fac. Agric. Kobe Univ.* **20**: 177–181.
- Hosaka, K., and Hanneman, R.E., Jr. 1998. Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 2. Localization of an S locus inhibitor (Sli) gene on the potato genome using DNA markers. *Euphytica*, **103**: 265–271.
- Hosaka, K., and Spooner, D.M. 1992. RFLP analysis of the wild potato species, *Solanum acaule* Bitter (*Solanum* sect. *Petota*). *Theor. Appl. Genet.* **84**: 851–858.
- Huamán, Z., and Spooner, D.M. 2002. Reclassification of landrace populations of cultivated potatoes (*Solanum* sect. *Petota*). *Am. J. Bot.* **89**: 947–965.
- Huamán, Z., Hawkes, J.G., and Rowe, P.R. 1980. *Solanum ajanhuiri*: an important diploid potato cultivated in the Andean Altiplano. *Econ. Bot.* **34**: 335–343.
- Jackson, M.T., Hawkes, J.G., and Rowe, P.R. 1977. The nature of *Solanum* × *chaucha* Juz. et Buk., a triploid cultivated potato of the South American Andes. *Euphytica*, **26**: 775–783.
- Jackson, M.T., Hawkes, J.G., and Rowe, P.R. 1980. An ethnobotanical field study of primitive potato varieties in Peru. *Euphytica*, **29**: 107–113.
- Johns, T., and Keen, S.L. 1986. Ongoing evolution of the potato on the Altiplano of western Bolivia. *Econ. Bot.* **40**: 409–424.
- Joseph, T., and Gopal, J. 1994. Variation in tuber dormancy of andigena potato (*Solanum tuberosum* ssp. *andigena*). *Indian J. Agric. Sci.* **64**: 247–248.
- Matsubayashi, M. 1991. Phylogenetic relationships in the potato and its related species. In *Chromosome engineering in plants: genetics, breeding, evolution*. Part B. Edited by T. Tsuchiya and P.K. Gupta. Elsevier, Amsterdam, Oxford, New York, and Tokyo. pp. 93–118.
- Nakagawa, K., and Hosaka, K. 2002. Species relationships between a wild tetraploid potato species, *Solanum acaule* Bitter, and its related species as revealed by RFLPs of chloroplast and nuclear DNA. *Am. J. Potato Res.* **79**: 85–98.
- Nei, M., and Kumar, S. 2000. *Molecular evolution and phylogenetics*. Oxford University Press, New York.
- Ochoa, C.M. 1990. *The potatoes of South America: Bolivia*. Cambridge University Press, Cambridge, UK.
- Provan, J., Powell, W., Dewar, H., Bryan, G., Machray, G.C., and Waugh, R. 1999. An extreme cytoplasmic bottleneck in the modern European cultivated potato (*Solanum tuberosum*) is not reflected in decreased levels of nuclear diversity. *Proc. R. Soc. Lond. B Biol. Sci.* **266**: 633–639.
- Quiros, C.F., Ortega, R., van Raamsdonk, L., Herrera-Montoya, M., Cisneros, P., Schmidt, E., and Brush, S.B. 1992. Increase of potato genetic resources in their center of diversity: the role of natural outcrossing and selection by the Andean farmer. *Genet. Resour. Crop Evol.* **39**: 107–113.
- Rabinowitz, D., Linder, C.R., Ortega, R., Begazo, D., Murguía, H., Douches, D.S., and Quiros, C.F. 1990. High levels of interspecific hybridization between *Solanum sparsipilum* and *S. stenotomum* in experimental plots in the Andes. *Am. Potato J.* **67**: 73–81.
- Raker, C.M., and Spooner, D.M. 2002. Chilean tetraploid cultivated potato, *Solanum tuberosum*, is distinct from the Andean populations: microsatellite data. *Crop Sci.* **42**: 1451–1458.
- Rohlf, F.J. 1997. NTSYSpc: numerical taxonomy system. Verion 2.0. Exeter Publishing Ltd., Setauket, N.Y.
- Salaman, R.N. 1946. The early European potato: its character and place of origin. *J. Linn. Soc. Lond. Bot.* **53**: 1–27.
- Simmonds, N.W. 1964. Studies of the tetraploid potatoes. 1. The variability of the Andigena Group. *J. Linn. Soc. Lond. Bot.* **58**: 461–474.
- Simmonds, N.W. 1976. *Evolution of crop plants*. Longman Group Limited, New York.
- Sukhotu, T., Kamijima, O., and Hosaka, K. 2004. Nuclear and chloroplast DNA differentiation in Andean potatoes. *Genome*, **47**: 46–56.
- Sukhotu, T., Kamijima, O., and Hosaka, K. 2005. Chloroplast DNA variation in the most primitive cultivated diploid potato species *Solanum stenotomum* Juz. et Buk., and its putative wild ancestral species using high-resolution markers. *Genet. Resour. Crop Evol.* In press.
- Swaminathan, M.S., and Magoon, M.L. 1961. Origin and cytogenetics of the commercial potato. In *Advances in genetics*. Vol. 10. Academic Press, London, UK. pp. 217–256.
- Tanksley, S.D., Ganai, M.W., Prince, J.P., de Vicente, M.C., Bonierbale, M.W., Broun, P., Fulton, T.M., Giovannoni, J.J., Grandillo, S., Martin, G.B., Messeguer, R., Miller, J.C., Miller, L., Paterson, A.H., Pineda, O., Roder, M.S., Wing, R.A., Wu, W., and Young, N.D. 1992. High density molecular linkage maps of the tomato and potato genomes. *Genetics*, **132**: 1141–1160.
- Ugent, D. 1970. The potato. *Science (Wash., DC)*, **170**: 1161–1166.
- Zimmerer, K.S., and Douches, D.S. 1991. Geographical approaches to crop conservation: the partitioning of genetic diversity in Andean potatoes. *Econ. Bot.* **45**: 176–189.