AFLP Assessment of Sweetpotato Genetic Diversity in Four Tropical American Regions

D.P. Zhang, M. Ghislain, Z. Huamán, J.C. Cervantes[1], and E.E. Carey[2]

Sweetpotato (*Ipomoea batatas* (L.) Lam.) was originally domesticated in the New World (Austin, 1988). The exact site of origin and domestication of the sweetpotato has not been well defined; neither has the wild ancestor of this species been found. Based on the numerical analysis of key morphological characters of sweetpotato and the wild *Ipomoea* species, Austin (1988) postulated that the center of origin of *I. batatas* was somewhere between the Yucatán Peninsula of Mexico and the mouth of the Orinoco River in Venezuela.

Linguistic evidence suggests that sweetpotato was dispersed to the rest of the world in three lines. The *kumara* line is prehistoric and based on lexical parallels between the Quechua name and Polynesian word *kumara*. That could explain the transfer of sweetpotato by Peruvian or Polynesian voyagers from northern South America to eastern Polynesia around AD 400. The *batata* line dates from the first voyage of Columbus in 1492, which resulted in the introduction of West Indian sweetpotatoes to Western Europe. Portuguese explorers transferred sweetpotatoes grown in western Mediterranean Europe to Africa, India, and the East Indies in the 16th century. The plant was recorded in South China by 1594 and in southern Japan by 1698. The *camote* (name derived from the word *camotli* in the Mayan language Nahuatl) line represents the direct transfer of Mexican sweetpotatoes by Spanish trading galleons between Acapulco, Mexico, and Manila, Philippines, in the 16th century (O'Brien, 1972; Yen, 1982).

Today sweetpotato, with an annual production of 124 million t, is the world's seventh most important crop after wheat, rice, maize, potato, barley, and cassava (CIP, 1996). It is widely grown in tropical, subtropical, and warm temperate regions. Developing countries account for 98% of the world's sweetpotato production. Its wide adaptability on marginal land and rich nutritional content provide an enormous potential for preventing malnutrition and enhancing food security in the developing world.

CIP has carried out numerous sweetpotato-collecting expeditions in Latin America and the Caribbean since 1985. Donations from other countries and the transfer of sweetpotato collections previously maintained in other international research centers further expanded the collection. The sweetpotato gene bank maintained at CIP now contains 5,526 cultivated accessions from 57 countries, 2,589 of them from Latin America. Peru alone contributed 1,099 accessions (Huamán and Zhang, 1997).

The success of any genetic conservation and breeding program is dependent on understanding the amount and distribution of genetic diversity present in the gene pool. For example, information on diversity distribution and domestic history is crucial for constructing core collections. They are a limited subset of accessions, derived from a larger germplasm collection, chosen to represent the genetic spectrum in the whole collection. Maximum genetic diversity is

1 CIP, Lima, Peru.
2 CIP, Nairobi, Kenya.

the key to establishing a good core subset (Brown, 1989).

In the present study, we applied amplified fragment length polymorphism (AFLP) fingerprinting to a group of Latin American sweetpotato cultivars maintained at CIP to answer two basic questions related to genetic structure in this gene pool:

1. Is there any regional differentiation in the sweetpotato diversity from tropical America? For purposes of this study, we defined four regions: (1) Central America, (2) Colombia-Venezuela, (3) Peru-Ecuador, and (4) Caribbean.
2. What is the geographical pattern of genetic diversity in these regions?

This survey is part of CIP's molecular assessment of genetic diversity in the sweetpotato collection. The information obtained will enhance our understanding of the genetic structure in the American sweetpotato gene pool and help to rationalize its management.

Materials and Methods

Plant materials

Four to ten sweetpotato cultivars were randomly selected from each country of the four regions, which resulted in a total of eighty cultivars with a geographical coverage ranging from Mexico to southern Peru (Table 1). The plant materials were obtained from the sweetpotato gene bank at CIP. Healthy young leaves were collected from accessions maintained in a screenhouse and in vitro. The leaf tissue was immediately immersed in liquid N and then transferred to –80°C, freeze-dried, and stored at room temperature until use.

AFLP protocol

A modified DNA miniprep procedure of Doyle and Doyle (1990) was used to extract DNA. The AFLP technique is described by Vos et al. (1995). DNA restriction uses the enzyme combination *Eco*RI/*Mse*I. After adapter ligation, DNA fragments are amplified using polymerase chain reaction (PCR). Primer annealing is targeted at the adapter and restriction-site sequence. Three-nucleotide extensions on both *Eco*RI and *Mse*I primers cause selective amplification of fragments. Primer combinations were chosen to produce a high number of unambiguous polymorphisms in a set of 10 sweetpotato genotypes tested. The eight primer combinations that were used are presented in Table 2.

Scoring and data analysis. Different fragments produced with each primer were treated as a unit character and numbered sequentially. Genotypes were scored for the presence (1) or absence (0) of each fragment. Only those fragments with medium or high intensity were taken into account. Fragments with the same mobility on the gel, but with different intensities, were not distinguished from each other when cultivars were compared. Monomorphic fragments were not scored.

From these data, a matrix of pairwise Euclidean distances defined by Excoffier et al., (1992) and a matrix of similarity based on simple matching coefficients (Sneath and Sokal, 1973) were calculated using the package RAPDistance v1.03 (Armstrong et al., 1995).

The similarity-based relationships between the 80 cultivars were then presented in a dendrogram following the unweighted pair group method with arithmetic mean algorithm (UPGMA) using SAHN clustering analysis of NTSYSpc (Rohlf, 1992).

Based on the Euclidean distances, the analysis of molecular variance (AMOVA) procedure (Excoffier et al., 1992) was applied to estimate variance components for AFLP phenotypes. Individual variation was partitioned within and between regions. The variance components of interest were extracted and tested using nonparametric permutational procedures. Variation between regions was then partitioned into pairwise distance between

Table 1. Name, collection number, and origin of 80 sweetpotato cultivars from the collection held at CIP.

No.	Name	Collection No.	Country of origin	No.	Name	Collection No.	Country origin
1	Cuba 9	CTX 34	Mexico	43	Catemaco	DLP 868	Venezuela
2	Linea 96	CTX 9	Mexico	44	Unknown	DLP 869	Venezuela
3	Cuba 3	CTX 5	Mexico	45	Chaco Morado	DLP 2869	Venezuela
4	Linea 475	CTX 16	Mexico	46	Unknown	DLP 2902	Venezuela
5	Unknown	DLP 3874	Panama	47	Unknown	DLP 2884	Venezuela
6	Columbiaorado	DLP 3892	Panama	48	Unknown	DLP 842	Venezuela
7	Unknown	DLP 3837	Panama	49	Unknown	DLP 824	Venezuela
8	Unknown	DLP 3834	Panama	50	Patatilla	DLP 806	Venezuela
9	Unknown	DLP 4617	Nicaragua	51	Chaco Columbiaorado	DLP 2868	Venezuela
10	Unknown	DLP 4678	Nicaragua	52	Morado	DLP 2896	Venezuela
11	Unknown	DLP 4686	Nicaragua	53	Unknown	DLP 1870	Colombia
12	Unknown	DLP 4675	Nicaragua	54	Unknown	DLP 1685	Colombia
13	Camote Blanco	GUA 494	Guatemala	55	Blanca	DLP 1731	Colombia
14	Santa Rosa (A)	GUA STRA	Guatemala	56	Morada	DLP 1793	Colombia
15	Camote Morado	GUA 940	Guatemala	57	Unknown	DLP 1858	Colombia
16	Camote Morado	GUA 948	Guatemala	58	Exquisita	DLP 1736	Colombia
17	Unknown	DLP 4545	Honduras	59	Camota	DLP 1771	Colombia
18	Unknown	DLP 4494	Honduras	60	Morado	DLP 1011	Colombia
19	Unknown	DLP 4521	Honduras	61	Blanca	DLP 1792	Colombia
20	Unknown	DLP 4558	Honduras	62	Unknown	DLP 2104	Colombia
21	Gumbs	SVG 12	El Salvador	63	Horse Money	DLP 3205	Jamaica
22	Lover's Nome	SVG 24	El Salvador	64	Tis-2544	DLP 3247	Jamaica
23	Six Weeks White	SVG 27	El Salvador	65	Clipperu	DLP 3211	Jamaica
24	Rasta	SVG 8	El Salvador	66	Tis-5093	DLP 3232	Jamaica
25	Papa Blanca	INVT 72	Cuba	67	Papota	SPV 44	Puerto Rico
26	Jiquima	INVT 51	Cuba	68	Sunny	SPV 43	Puerto Rico
27	Nigua	INVT 63	Cuba	69	Mojave	SPV 65	Puerto Rico
28	Pata de Gallina	INVT 74	Cuba	70	Toquecita	SPV 55	Puerto Rico
29	Unknown	CSDA 22	Dominican Rep	71	Unknown	DLP 3433	Peru
30	3-c-n	3-C-N	Dominican Rep	72	Unknown	DLP 253	Peru
31	Tunita	DLP 1567	Dominican Rep	73	Unknown	RCB IN-199	Peru
32	Yema de Huevo	SOSA 33	Dominican Rep	74	Unknown	DLP 909	Peru
33	Unknown	DLP 1257	Ecuador	75	Unknown	DLP 1921	Peru
34	De Dulce	DLP 1149	Ecuador	76	Unknown	RCB IN- 90	Peru
35	Blanco Papa	DLP 1435	Ecuador	77	Unknown	ARB 355	Peru
36	Camote de Dulce	DLP 1231	Ecuador	78	Unknown	ARB 455	Peru
37	Bunuelo	DLP 1484	Ecuador	79	Unknown	DLP 2298	Peru
38	Irizo Blanco	DLP 1405	Ecuador	80	Unknown	ARB 97	Peru
39	Unknown	DLP 1186	Ecuador				
40	Niguilla	DLP 1397	Ecuador				
41	Unknown	DLP 1157	Ecuador				
42	Camote Dulce	DLP 1498	Ecuador				

Table 2. Primer combinations used in AFLP analysis of sweetpotato diversity.

Number	E+3/M+3 nucleotide extensions
E34/M33	5'-GACTGCGTACCAATTC AAT-3'
	5'-GATGAGTCCTGAGTAA AAG-3'
E38/M38	5'-GACTGCGTACCAATTC ACT-3'
	5'-GATGAGTCCTGAGTAA ACT-3'
E33/M38	5'-GACTGCGTACCAATTC AAG-3'
	5'-GATGAGTCCTGAGTAA ACT-3'
E38/M32	5'-GACTGCGTACCAATTC ACT-3'
	5'-GATGAGTCCTGAGTAA AAC-3'
E36/M37	5'-GACTGCGTACCAATTC ACC-3'
	5'-GATGAGTCCTGAGTAA ACG-3'
E38/M34	5'-GACTGCGTACCAATTC ACT-3'
	5'-GATGAGTCCTGAGTAA AAT-3'
E36/M33	5'-GACTGCGTACCAATTC ACC-3'
	5'-GATGAGTCCTGAGTAA AAG-3'
E36/M34	5'-GACTGCGTACCAATTC ACC-3'
	5'-GATGAGTCCTGAGTAA AAT-3'

regions to examine the regional contribution to total molecular diversity (Excoffier et al., 1994).

Results

AFLP profile, similarity and cluster analysis

The eight primer combinations generated 210 polymorphic and clearly scorable fragments across the 80 cultivars. There were no region-specific markers (present in one region but absent in the other). However, the mean genetic similarities vary considerably between regions, ranging from 0.592 for Central America to 0.660 for Peru-Ecuador (Table 3).

The overall geographic proximity is low as reflected by the UPGMA dendrogram (Figure 1). No clusters clearly distinguished cultivars from Central America, Caribbean, and Colombia-Venezuela. Within each cluster, however, cultivars from the same geographical origins tend to group together. In sharp contrast with the closer relationship between cultivars from Central America, the Caribbean, and Colombia-Venezuela, the Peruvian-Ecuadorian cultivars were clearly distinguishable from the others. All of the 40 cultivars from Peru-Ecuador were grouped into one single cluster at a higher similarity level.

The mean similarity and the UPGMA dendrogram both suggest that Central American cultivars apparently have the highest genetic diversity, whereas Peru-Ecuador cultivars have the lowest. The Caribbean and Colombia-Venezuela cultivars have an intermediate level. The cluster analysis demonstrated a substantial regional differentiation in the Latin American sweetpotato cultivars. Sweetpotatoes from Peru-Ecuador contributed most to the regional differentiation.

Analysis of molecular variance

The results of AMOVA (Table 4) revealed large within-region variations, which accounted for 91.3% of the total molecular variance. This large within-region variation could be the contributions from two sources. The first is the expected large

Table 3. Number of AFLP-generated polymorphic markers and mean similarity in sweetpotato cultivars from four Latin American regions.

Region	Accessions (no.)	Polymorphic fragments within regions	Mean similarities within regions (no.)
Central America	24	213	0.592
Colombia-Venezuela	20	210	0.611
Caribbean	16	202	0.639
Peru-Ecuador	20	195	0.660

variation between individual cultivars, because sweetpotato is an outcrossing hexaploid, therefore the variation due to sexual reproduction is large. The second source is the between-country variation within a region. In this study, each region included more than one country, and the between-country variation was confounded by the between-individual variation. The between-country component was not separated in the analysis because the sample sizes in some countries, particularly those from Central America and the Caribbean, were too small to represent one country.

The between-region variation, although accounting for only 8.7% of total molecular variance, was statistically significant using a nonparametric test (Excoffier, 1994). That means there is measurable divergence between the four regions. Individuals from different regions were more divergent, on average, than individuals from the same regions.

Partitioning between-region variability into pairwise distance gave a clearer picture of the extent to which each region contributes to the total molecular diversity (Table 5). Pairwise distances between regions differ greatly. The shortest distance was between Central America and the Caribbean (0.0234), whereas the longest distance was between Central America and Peru-Ecuador (0.1269). Peru-Ecuador has the greatest mean distance to the rest of the regions (0.1007), indicating that sweetpotato from Peru-Ecuador has the most differentiation thus contributed most to the between-region variation. This result fully agreed with the observed grouping pattern of Peru-Ecuador cultivars in the cluster analysis.

There is a significant geographic pattern of diversity distribution (Table 4). The Central America region contains the most internal diversity (0.1990), whereas the Peru-Ecuador region contains the least (0.1677). The intermediate within-region diversities were found in Colombia-Venezuela (0.1837) and the Caribbean (0.1795). This pattern is the same as that based on mean similarity and cluster analysis. It thus confirms the higher level of genetic diversity in Central America. The diversity level decreased as sweetpotato was dispersed from the center of origin toward the south and the east.

Discussion

This study is the first case of applying AFLP fingerprints for diversity assessment in sweetpotato. The high polymorphism of AFLP makes it a powerful tool for genotyping a large number of accessions. It is suitable for genetic diversity assessment in large sweetpotato gene banks.

Austin (1988) postulated that the center of origin of *I. batatas* was somewhere between the Yucatán Peninsula of Mexico

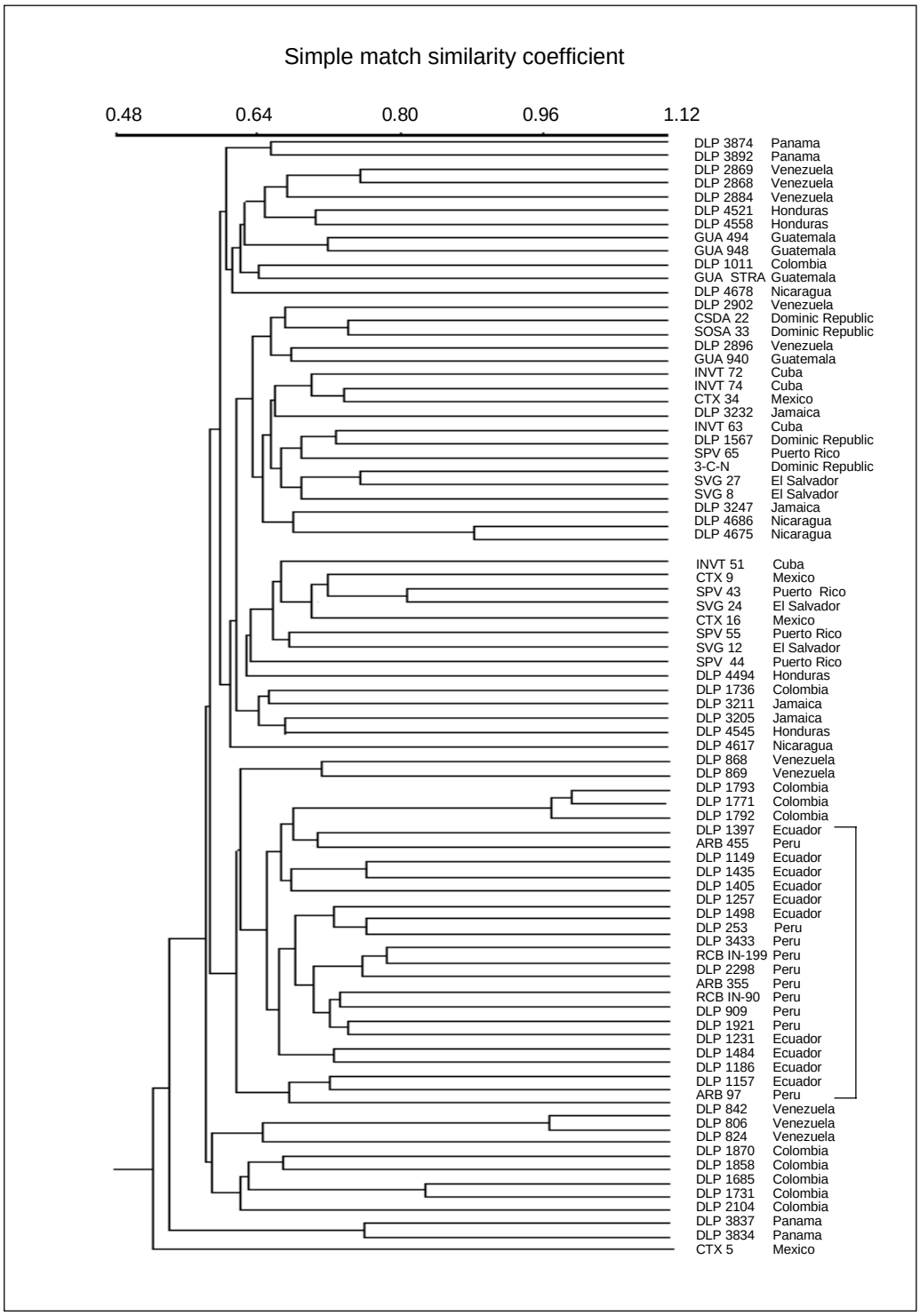

Figure 1. UPGMA dendrogram of 80 sweetpotato varieties based on simple matching coefficients from AFLP markers.] = Regional differentiation of Peru-Ecuador varieties and the lower internal diversities in the region.

Table 4. AMOVA format for the extraction of components of AFLP variation of sweetpotato clones between regions, and between individuals within regions.

Variation source	Df	SSD	MSD[a]	Variance component	Total[b] (%)	P value[c]
Between regions	3	1.593	0.531	0.0174	8.55	<0.001
Individuals within regions	76	14.129	0.1859	0.1859	91.45	
Central America	23	4.5758	0.1990			
Colombia-Venezuela	19	3.6745	0.1837			
Peru-Ecuador	19	3.1868	0.1677			
Caribbean	15	2.6920	0.1795			
Total	**79**	**15.722**				

a. Mean squared deviations.
b. Percent of total molecular variance.
c. Probability of obtaining a larger component estimate. Number of permutations = 1000

Table 5. Genetic distances between sweetpotato cultivars from four Latin American regions (Distances = PhiST between pairs of regions).

Regions	Central America	Colombia- Venezuela	Peru-Ecuador	Caribbean
Central America	0.0000	0.0391	0.1269	0.0234
Colombia-Venezuela	0.0391	0.0000	0.1130	0.0622
Peru-Ecuador	0.1269	0.1130	0.0000	0.0622
Caribbean	0.0234	0.0622	0.0622	0.0000
Mean distance to other regions	0.0631	0.0714	0.1007	0.0493

and the mouth of the Orinoco River in Venezuela. That is where *I. trifida* might have been crossed with another putative ancestor, *I. triloba*, and might have produced the wild ancestor of *I. batatas*. By at least 2,500 BC, the cultigen had most likely been spread by the local people to almost the limits for cultivation in Central and South America that existed when the Europeans arrived nearly 4,000 years later. The greater molecular diversity in Central America provides strong evidence that Central America is the primary center of diversity. It should also be the center of origin as well, considering the richness of wild *Ipomoea* species that are closely related to sweetpotato.

The much lower molecular diversity found in Peru-Ecuador suggests that cultivars in this region constitute a subset from the center of primary diversity. This result reflects the traditional pattern of crop distribution from the primary center of origin. As the crop dispersed along migration and trade routes, and through exchanges between native communities, the diversity level decreased due to the limited donor populations. A similar pattern was demonstrated in our previous work on a group of Papua New Guinea sweetpotatoes (Zhang et al., 1998).

Sweetpotato, however, is an ancient crop in Peru (Engel, 1970; Yen, 1974). The

unique ancient agriculture and other ethnographic factors may have led to a greater morphological diversity based on relatively narrow donor diversity. That could explain why Peru-Ecuador hosts a large number of sweetpotato cultivars but has a lower level of diversity. The natural and human selection forces also caused a significant regional differentiation, which was reflected as the greater genetic distance between Peru-Ecuador and the other regions. Therefore, Peru-Ecuador could be considered as a secondary center of sweetpotato diversity.

References Cited

Armstrong, J., A. Gibbs, R. Peakall, and G. Weiller. 1995. RAPDistance, Package Manual. Version 1.03. Australian National University, Canberra.

Austin, D.F. 1988. The taxonomy, evolution and genetic diversity of sweetpotatoes and related wild species, In: Gregory, P. (ed.). Exploration, maintenance, and utilization of sweetpotato genetic resources. CIP, Lima, Peru. p. 27-60.

Brown, A.H.D. 1989. Core collection: A practical approach to genetic resources management. Genome 31:818-824.

CIP, 1996. Sweetpotato facts. International Potato Center, Lima, Peru. (Brochure.)

Doyle, J.J. and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Excoffier, L. and P.E. Smouse. 1994. Using allel frequencies and geographic subdivision to reconstruct gene trees within a species: Molecular variance parsimony. Genetics 136:343-359.

Excoffier, L., P.E. Smouse, and J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Applications to human mitochondrial DNA restriction data. Genetics 131:479-491.

Engel, E. 1970. Exploration of the Chilca Canyon. Cur. Anthropol. 11:55-58.

Huamán, Z. and D.P. Zhang. 1997. Sweetpotato. In: Fuccillo, D., et al. (eds). Biodiversity in trust - Conservation and use of plant genetic resources in CGIAR Centers. Cambridge University Press, Cambridge, UK. p. 29-38.

O'Brien, P.J. 1972. The sweet potato: Its origin and dispersal. Am. Anthropologist 74:343-365.

Rohlf, F.J. 1992. NTSYS-PC numerical taxonomy and multivariate analysis system, Version 1.70. Owner's manual. Exter Publications, Setauket, NY, USA.

Sneath, P.H.A. and R.R. Sokal. 1973. Numerical taxonomy. The principles and practice of numerical classification. W.H. Freeman, San Francisco, CA, USA.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van der Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: A new technique for DNA fingerprinting. Nuc. Ac. Res. 23:4407-4414.

Yen, D.E. 1974. The sweetpotato and Oceania. Bishop Museum Bull. (Honolulu) 236:1-389.

Yen, D.E. 1982. Sweet potato in historical perspective. In: Villareal, R.L. and T.D. Griggs (eds.). Sweet potato, Proceedings of First International Symposium. AVRDC Publ. No. 82-172. p. 17-30.

Zhang, D.P., M. Ghislain, Z. Huamán, A. Golmirzaie, and R.J. Hijmans. 1998. RAPD variation in sweetpotato [*Ipomoea batatas* (L.) Lam] cultivars from South America and Papua New Guinea. Gen. Res. Crop Evolution 45:271-277.