The Americas Journal of Plant Science and Biotechnology ©2010 Global Science Books

Genetic Diversity in Wild Apple (*Malus* sp.) Populations in Argentina

Paula Calvo[1*] • Alicia Carrera[2] • Enrique Sánchez[1] • Mónica Poverene[2]

[1] INTA, EEA-Alto Valle, General Roca, Río Negro, Argentina
[2] Universidad Nacional del Sur, Departamento de Agronomía y CERZOS (CONICET) Bahía Blanca, Argentina

Corresponding author: * **pcalvo@correo.inta.gov.ar**

ABSTRACT

In Argentina, wild apple trees are distributed in the western areas of Neuquén province, along the Andean low hillsides. A collection of 23 wild populations was investigated with RAPD primers to evaluate variability and genetic relationships. The data were analyzed by multivariate analysis. A low germplasm variability of $H = 0.115$ was estimated with Nei's genetic diversity index. Cluster and principal coordinate analyses did not reveal distinct geographic patterns or strong grouping among individuals of the same origin, but there was a clear partition between wild and cultivated materials. AMOVA showed the largest level of diversity within populations (57%). A strong human influence seems to have shaped wild apple distribution.

Keywords: germplasm, genetic variability, plant conservation, RAPD

INTRODUCTION

Apple (*Malus x domestica* Borkh) is the economically most important fruit tree crop of temperate zones. It is also widely cultivated in most of temperate latitudes and/or on heights (Luby 2003). Despite this wide distribution, the genetic basis of domesticated apple has been eroded and most of the world production relies on few varieties only (Hokanson *et al*. 2001). Moreover, apple breeding programs normally use a restricted number of founding clones (Noiton and Alspach 1996). The available diversity pool for fruit improvement could be expanded by including wild *Malus* in breeding programs.

Wild and cultivated apples can be maintained in gene banks. In order to increase genetic variability wild sexual individuals, propagated by seeds, represent valuable materials where recombination could occur as a consequence of self-incompatibility and outbreeding (Kitahara *et al*. 2005).

The center of *Malus* genetic diversity lies in East Asia (Juniper *et al*. 1998). From there, the fruits dispersed through Persia and Greece to the Roman Empire. The Romans introduced apple fruits in Europe and then the domesticated apple has diversified and settled worldwide (Harris *et al.* 2002; Büttner *et al.* 2004). Populations of wild apple trees (*Malus* sp.) introduced in the 17th century, became established in the Andean zone of Neuquén, Argentina. The trees propagated by indians survived and reproduced naturally and remained in the region since then (Bandieri 2005). The area has very extreme weather conditions, with low winter temperatures and strong winds. With the aim to characterize and collect these materials, two collection trips were accomplished for this study. Wild germplasm could provide valuable traits for breeding programs. The populations are endangered by overgrazing since the trees are very old and little natural regeneration is observed. The collected material was grafted on apple cloned rootstocks and incorporated to INTA Alto Valle Experimental Station apple genebank for its conservation and evaluation (Calvo *et al*. 2005).

The knowledge of the levels and patterns of genetic variation in populations of wild plants constitute a paramount aspect for the development of appropriate strategies for conservation, further collection, and preservation of genetics resources (Frankel *et al.* 1995; Zhao *et al*. 2008).

Traditionally, morphological characters have been used for apple germplasm characterization but many of these characters are strongly influenced by the environment. The use of molecular techniques became popular during the past decade and has contributed to a better understanding of the genetic composition and distribution patterns of numerous species (Engelman and Engels 2002). Molecular markers have a main advantage because they are not influenced by the environment and, particularly in species with a long juvenile period, the time for characterization can be greatly reduced (Goulão and Oliveira 2001). RAPD markers have been used in apple species for the identification of cultivars (Koller *et al.* 1993; Mulcahy *et al*. 1993; Royo and Itoiz 2004; Garkava-Gustavson and Nybom 2007), for construction of a saturated genetic map (Hemmat *et al.* 1994), for paternity determination (Gardiner *et al*. 1996), and to study genetic diversity (Dunemann *et al*. 1994; Zhou and Li 2000; Coart *et al*. 2003). Possible reproducibility problems can be avoided if the same protocol is carefully followed and only clear, strong and consistent bands are scored (Koller *et al.* 1993; Mulcahy *et al*. 1993).

The goal of this study was to assess the levels of genetic diversity in Argentinean populations of wild apple, as well as the distribution of variability within and among populations. RAPD markers were chosen for a first approach of diversity.

MATERIALS AND METHODS

Plant material

A total of 105 individuals representing 23 populations, uniformly distributed in the collecting area, were sampled. The number of trees in each population ranged from 3 to 40. For each population, every tree in enough good conditions to obtain grafting material was sampled. The collecting area is located between 38º 50' to 40º 11' S latitude and 70º 55' to 71º 40' W longitude, in the Andean Region of Argentina (**Fig. 1**).

Received: 12 May, 2009. Accepted: 10 December, 2010.

Original Research Paper

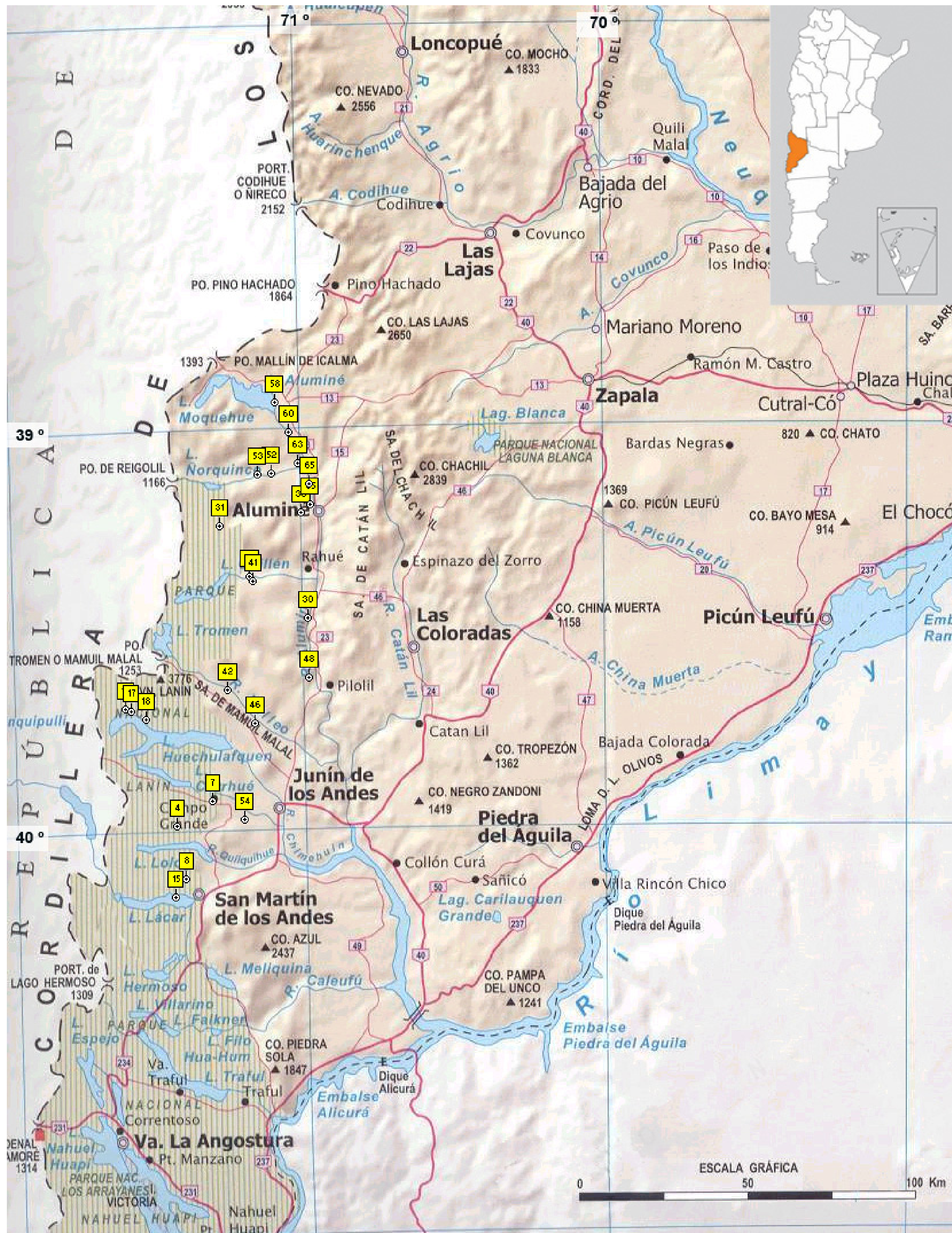

Fig. 1 Locations of the population included in this study. The numbers indicate the collection sites for each population. The map is based on IGN (National Geographic Institute of Argentina), © Instituto Geográfico Militar, with kind permission.

Sample consisted of shoots grown within the year of 0.4 to 1 cm diameter, having at least three winter buds. They were conserved at 4°C until the time for grafting on clonal apple rootstocks (M7, East Malling). Fresh, fully expanded leaves from each sampled tree were collected from the gene bank in the growing season, and were shipped chilled to the laboratory. The leaves were frozen at -20°C and then lyophilized for 48 hrs. The lyophilized material was stored at low humidity and room temperature until DNA extraction. Samples of the major commercial apple varieties, 'Royal Gala' (RG), 'Granny Smith' (GS), 'Red Delicious' (RD) and 'Golden Delicious' (GD) were also included for comparison.

RAPD analysis

The DNA was obtained from 200 mg of lyophilized leaf tissue, using a modified CTAB extraction protocol (Hoisington *et al.* 1994). The extraction buffer was prepared with 100 mM Tris pH 8.0, 20 mM EDTA pH 8.0, 1.5 M NaCl, 2% CTAB and 2% β-mercapthoetanol, with the addition of 2% polyvinylpyrrolidone (PVP-40). PCR reactions were carried out in 25 µl of final mix containing 20 ng of DNA template, 0.2 mM of each dNTPs (Invitrogen Life Science Technologies), 30 ng primer (Alpha DNA), 2.5 µl buffer 10X, 2.5 mM $MgCl_2$ and 2 U *Taq* DNA polymerase (Invitrogen Life Science Technologies). The steps for amplification reaction were: 4 min at 94°C, 36 cycles of 30 sec at

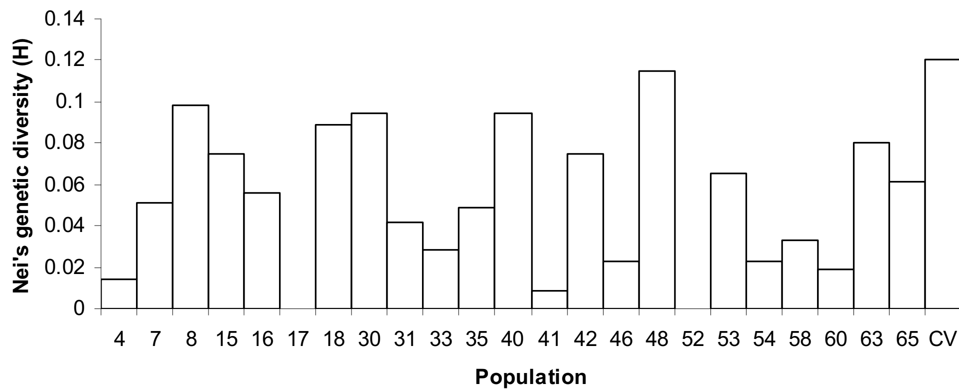

Fig. 2 Genetic diversity measure (H) in wild populations (numbers) and the commercial cultivars group (CV).

94°C, 1 min at 36°C and 1 min at 72°C and a final exten-sion of 6 min at 72°C. DNA amplifications were carried out in a PCT-100 MJ Research Inc. thermocycler.

Twelve Operon primers were evaluated and five of them were selected for the final study based on quality (clarity and signal strength) and reproducibility. DNA of the same four trees was included in all the runs as controls, to verify reproducibility.

The RAPD markers were resolved by electrophoresis in 1.5% agarose gels using 0.1 × TBE buffer and stained with ethidium bromide. A 100 bp ladder (PB-L Productos Bio-Lógicos, Arg.) was used as a molecular weight marker to estimate the size of the amplification products.

Data analysis

Amplification products were scored manually. Each band was considered as an independent locus. The data were entered into a binary matrix as discrete variables (1 for presence and 0 for absence).

Genetic diversity was calculated using the unbiased estimate of Nei (Nei 1978) for expected heterozygosis (H) as:

$$\hat{H} = \frac{N}{N-1}\left(1-\sum_{i=1}^{l}\hat{p}_i^{\,2}\right)$$

where p is the frequency of the ith allele and N the number of samples.

These measures were calculated for each loci and each population. At the population level statistical tests were performed to compare the average level of genetic variability among populations using Friedman's nonparametric test.

The Info-Gen software for statistical analysis of genetics data was used (Balzarini and Di Rienzo 2003).

Two approaches were used to analyze relationships between individuals: Cluster analysis and Principal Coordinate Analysis (PCO). A distance matrix, based on Jaccard's coefficient was calculated for all pairwise comparison between individuals. Matrix distance was used to perform a cluster analysis using UPGMA (Average linkage algorithm) to generate the corresponding dendrogram. The goodness of fit was measured by the cophenetic correlation test. Cluster analysis was performed with Info-Gen software.

A distance matrix from a Euclidean binary distance was used for PCO analysis. The calculation of genetics distances followed the method of Huff *et al*. (1993).

The relationships among populations were studied using a genetic distance matrix (Nei 1972).The relationships among populations were represented by PCO analysis performed with GenAlEx 6 (Peakall and Smouse 2006).

The individual pairwise Euclidean distance matrix obtained with RAPD data set was analyzed using analysis of molecular variance (AMOVA) (Excoffier *et al*. 1992) and Φ-statistics were computed. Total genetic variation was partitioned into two levels: within and among populations. AMOVA analysis was performed using GenAlEx 6 (Genetic Analysis in Excel; Peakall and Smouse 2006: http://www.anu.edu.au/BoZo/GenAlEx (freely available).

The correlation between geographic and genetic distance matrices was analyzed by the Mantel test of matrix correspondence. Statistical significance was determined by random permutation, with the number of permutations set to 1000 (GenAlEx6).

Table 1 RAPD primer, sequences, number of polymorphic bands and size of bands.

Primer	Sequence: 5′ to 3′	№ of scored bands	№ of polymorphic bands	Size range (bp)
OPB04	GGA CTG GAG T	5	5	370-900
OPC05	GAT GAC CGC C	9	8	170-1150
OPA01	CAG GCC CTT C	11	10	380-1250
OPA03	AGT CAG CCA C	8	5	530-1000
OPD04	TCT GGT GAG G	5	3	650-1050

Fig. 3 Populations of wild apple trees in the area of collection.

RESULTS AND DISCUSSION

Five selected primers produced 38 clear and consistent bands, 31 of these (81.6%) were polymorphic (**Table 1**). A single primer detected from 3 to 10 bands, with size range from 170 to 1250 bp. The number of polymorphic bands obtained with each primer was quite variable, ranging from three bands (OPD04) to 10 (OPA01). Our results were similar to others obtained in apple (Royo and Itoiz 2004; Garkava-Gustavsson and Nybom 2007).

Genetic diversity value in the entire data set (H_T) was 0.115. Evaluation of heterozygosity in each wild population showed population 48 as the most diverse (H = 0.115), while populations 17 and 52 revealed as the most homogeneous (H = 0) (**Fig. 2**). The H_T value was lower than the mean value obtained by Coart *et al*. (2003) in *Malus sylvestris* using AFLP, and also lower than the values obtained by Garkava-Gustavsson and Nybom (2007) and Goulão and Oliveira (2001) using RAPD in an old commercial apple variety collection. In our study the commercial variety group resulted the most diverse (H = 0.12).

Normally, we can expect that genetic diversity is main-

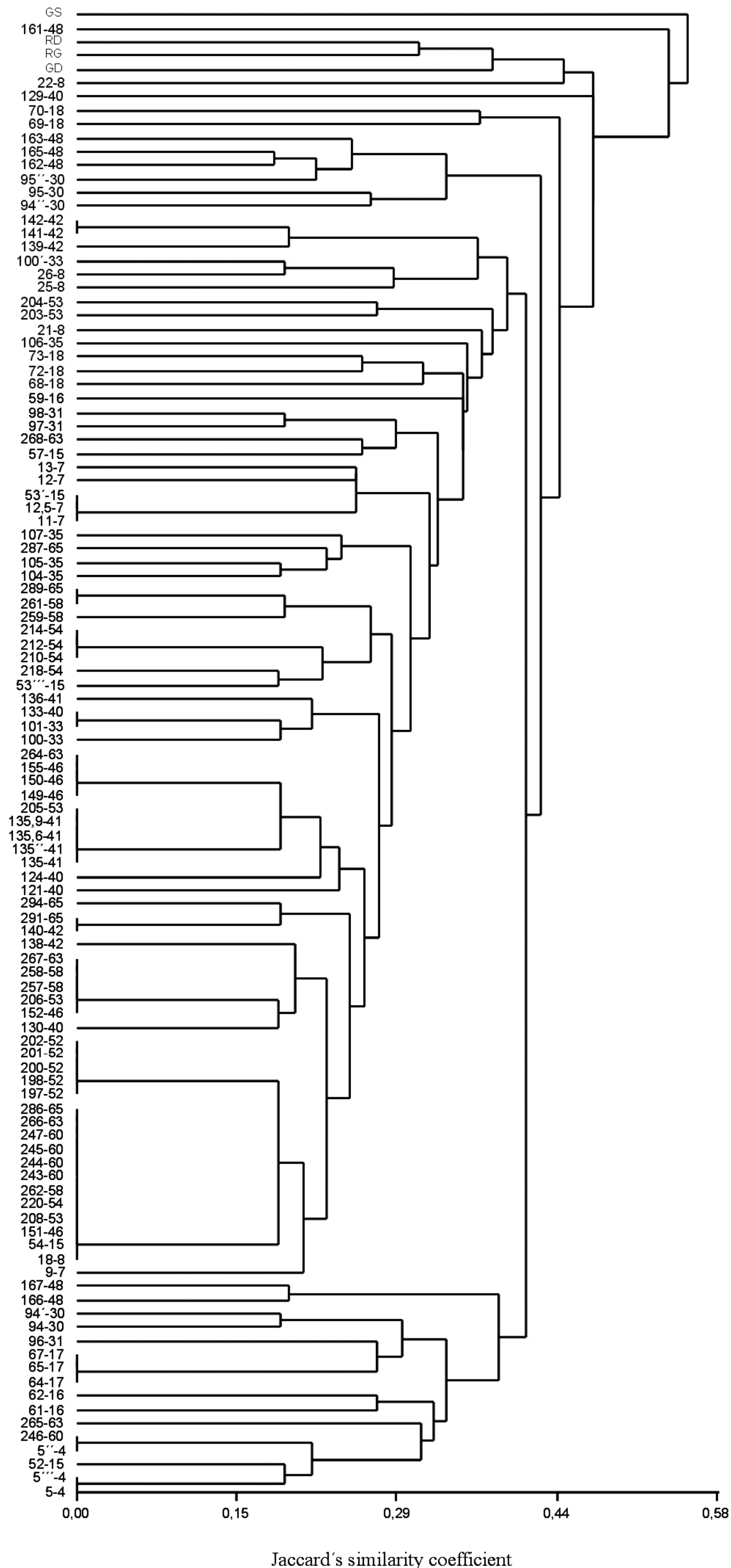

Fig. 4 UPGMA dendrogram representing genetic relationships among individuals. Numbers following hyphenate indicate population.

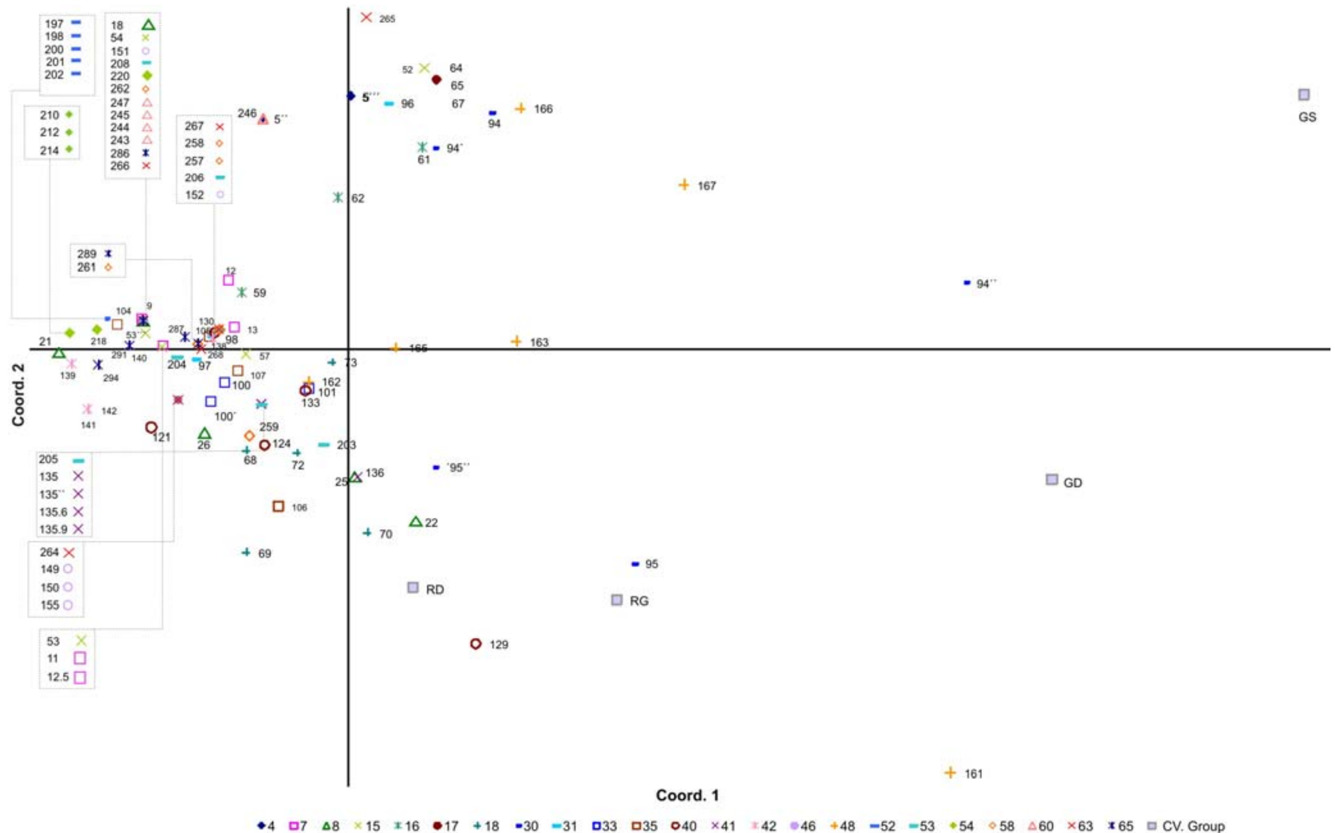

Fig. 5 PCO plot of the first two axes based on the individual distance binary. Percentage of variation explained by the first 2 axes is 54.98%. Different symbols mean different populations. Individuals in the same box were plotted in the same points.

tained at higher levels in larger populations (Hartl and Clark 1997). Here there was no obvious relationship between genetic diversity and population size, in fact the largest population (15) presented intermediate values and the largest values corresponded to a relatively small population (48). A similar situation was also reported by Jordano and Godoy (2000) in *Prunus mahaleb* populations under parallel landscapes. As *Malus* has a long life cycle, most environmental constraints (grazing, erosion) and scarce regeneration do not affect adult trees to a large extent; nevertheless they would affect juvenile individuals and recruitment in the future (**Fig. 3**).

The Friedman non-parametric test showed highly significant statistic differences in the amount of genetic diversity (*H*) in different populations. A cluster analysis was performed to study the relationships among all individuals. The dendrogram showed several clusters. Some individuals were grouped in accordance with their population of origin but this correlation cannot be generalized. A number of trees were very similar or identical in their patterns (**Fig. 4**).

Goodness-of-fit analysis suggests that average linkage is useful for this kind of representation (cophenetic correlation: 0.887).

PCO analysis showed that many individuals were very similar and it was not possible to recognize clear groups or cluster the individuals by population, whereas commercial varieties clearly separated from wild materials (**Fig. 5**).

Some wild individuals (161, 94´´and 95) segregated from the rest of the wild individuals (**Fig. 5**) and interspersed with the commercial varieties. They belong to populations where individuals showed similarity to other wild samples. This could be explained through the occurrence of an escape of edible apples attributable to local residents or tourists. Alternatively, given that at the collecting time these trees displayed a wild morphological phenotype, they could simply represent a component of the local wild apple genetic variability. These individuals are currently growing and are been evaluated at an experimental station.

As in our study, Coart *et al*. (2003), using AFLP, were able to differentiate edible apple cultivars from wild genotypes of *Malus sylvestris*, but within this last group the individuals were not distributed according to their origin. Garkava-Gustavsson and Nybom (2007) observed also little grouping associated with different origin in cultivated apple. The distribution pattern obtained in our analysis can be attributed to strong human factors in the dispersion of the populations in the area under study.

The PCO analysis in **Fig. 6** represented the relatedness among the 23 populations and commercial varieties group based on Nei's distance matrix. Commercial variety group was clearly distinct from the wild accessions showing the effect of domestication and selection. However, wild populations were widespread enough to reflect a large amount of genetic variation.

A matrix of geographical distances was compared with Nei's genetic distance matrix to investigate the possible association between genetic and geographic distance. The Mantel test showed a low correlation between genetic and geographic distance (R^2 0.0031). Additional insight on this weak geographical-genetic distance relationship is offered by examining **Figs. 1** and **6**. An example is that populations 17 and 18 are separated by 6 km with a genetic distance of 0.162, while populations 46 and 33 are separated by 75 km and genetic distance is only 0.041. The absence of correlation between genetic and geographic distance suggests that population dispersion could be caused by forces such as human migrations in the area under study (Bandieri 2005).

Analysis of molecular variance (AMOVA) for the total of wild RAPD data set (commercial varieties were excluded) showed that 43% of the total variability was attributable to differences among populations and 57% to within population differences ($P < 0.001$), as expected for an outcrossing tree species (Hamrick and Godt 1996). This pattern of diversity suggests gene flow and low genetic isolation among populations. Nevertheless, there was substantial variation among populations. To satisfactorily capture most of the variation, the collector should sample a considerable number of individuals in each population. Similar strategy

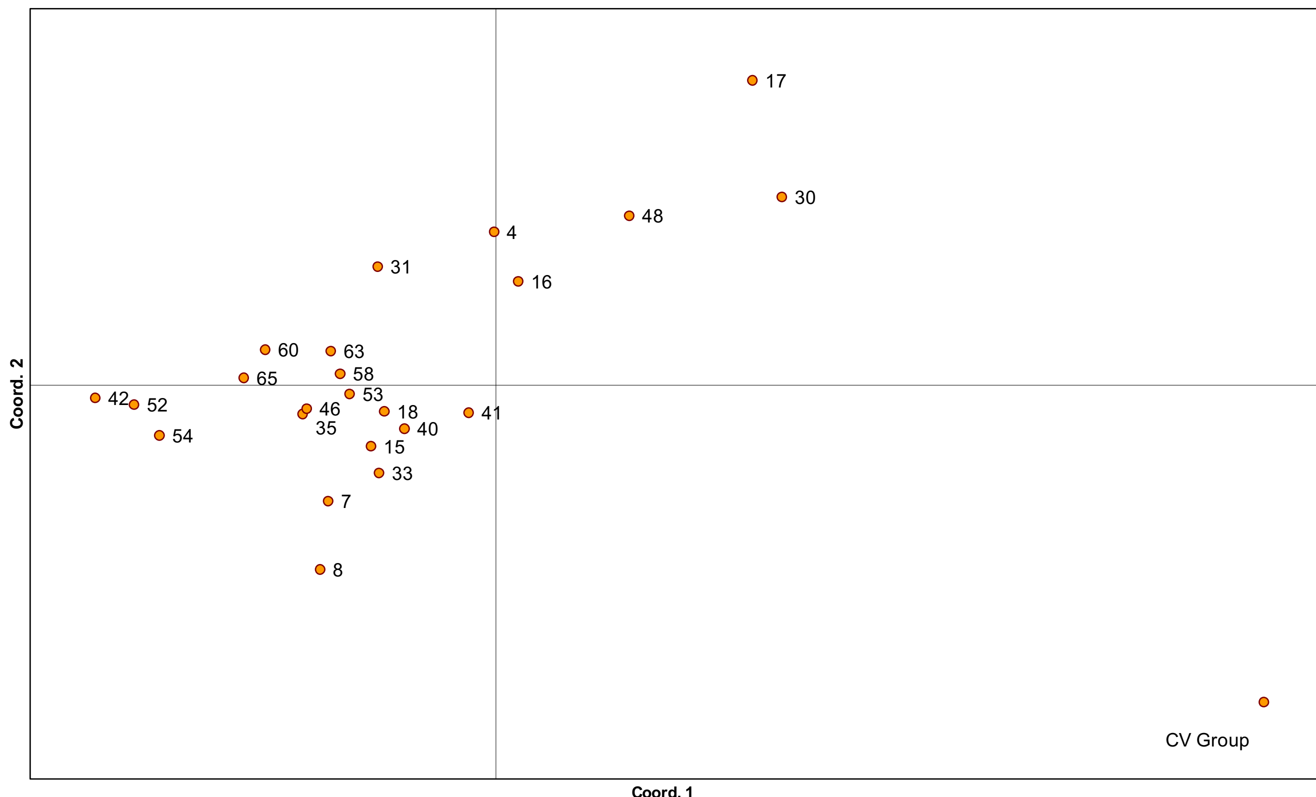

Fig. 6 PCO plot of the first two axes based on Nei's distance matrix. Percentage of variation explained by the first 2 axes is 66%.

was proposed by Volk *el al.* (2005) for *Malus sieversii* collection. Furthermore, the location of the populations in PCO analysis may be useful to exclude from the sampling job those populations with high genetic similarity.

The low diversity levels in wild apples, including two sites without any variation, could be addressed to a founder effect in those populations. The molecular pattern shows a strong human impact in the distribution of wild apple variability in Argentina.

The differences observed among wild apple germplasm and cultivated material, the existence of small populations and almost null regeneration observed, justifies *ex-situ* conservation of the collected germplasm.

ACKNOWLEDGEMENTS

To Administración de Parques Nacionales de la República Argentina who authorized our collecting mission in protected areas.

REFERENCES

Balzarini M, Di Rienzo J (2003) *Info-Gen*: Software para análisis estadístico de datos genéticos. Facultad de Ciencia Agropecuarias. Universidad Nacional de Córdoba. Argentina

Bandieri S (2005) *Historia de la Patagonia*, Sudamericana, Buenos Aires, 448 pp

Büttner R, Manfred F, Forsline P, Geibel M, Ponomarenko V (2004) Gene banks for the preservation of wild apple genetic resources. *Journal of Fruit and Ornamental Plant Research* **12 (Special Issue)**, 99-105

Calvo P, Carrera A, Poverene M (2005) Colecta y conservación de manzanos silvestres (*Malus* sp). V Sirgealc Symposium of Genetic Resources for Latin America and the Caribbean, 23-25 November 2005, Montevideo, Uruguay. Abstract N 71

Coart E, Vekemans X, Smulders MJM, Wagner I, van Huylenbroeck J, van Bockstaele E, Roldán-Ruiz I (2003) Genetic variation in the endangered wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as revealed by amplified fragment length polymorphism and microsatellite markers. *Molecular Ecology* **12**, 845-857

Dunemann F, Kahnau R, Shmidt H (1994) Genetic relationships in *Malus* evaluated by RAPD 'fingerprinting' of cultivars and wild species. *Plant Breeding* **113**, 150-159

Engelmann F, Engels JMM (2002) Technologies and Strategies for *ex situ* conservation In: Johannes MM, Engels V, Ramanatha R, Brown AHD, Jackson MT (Eds) *Managing Plant Genetic Diversity*, CABI Publishers, 487 pp

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

Frankel O, Brown A, Burdon J (1995) *The Conservation of Plant Biodiversity*, Cambridge University Press, United Kingdom, 299 pp

Gardiner SE, Bassett HCM, Madie C, Noiton DAM (1996) Isozyme, RAPD and RFLP markers used to deduce a putative parent for the "Braeburn" apple. *Journal of the American Society for Horticultural Science* **121**, 996-1001

Garkava-Gustavsson L, Nybom H (2007) Genetic diversity in a collection of apple (*Malus x domestica* Borkh.) cultivars as revealed by RAPD markers. *International Journal of Horticultural Science* **13**, 1-11

Goulão L, Oliveira CM (2001) Molecular characterization of cultivars of apple (*Malus* x *domestica* Borkh.) using microsatellite (SSR and ISSR) markers. *Euphytica* **122**, 81-89

Hamrick JL, Godt MJW (1996) Effect of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London* B **351**, 1291-1298

Hartl DL, Clark AG (1997) *Principles of Population Genetics* (3rd Edn), Sinauer Associates Inc., Sunderland, MA, 542 pp

Harris SA, Robinson JP, Juniper BE (2002) Genetic clues to the origin of the apple. *Trends in Genetics* **18**, 426-430

Hemmat M, Weeden NF, Manganaris AG, Lawson DM (1994) Molecular marker linkage map in *Malus*. *Journal of Heredity* **85**, 4-11

Hoisington D, Khairallah M, Gonzalez de Leon (1994) *Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory* (2nd Edn), CIMMYT, Mexico. DF, 50 pp

Hokanson SC, Lamboy WF, Szewc-McFadden AK, McFerson JR (2001) Microsatellite (SSR) variation in a collection of *Malus* (apple) species and hybrids. *Euphytica* **118**, 281-294

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalograss [*Buchloë dactyloides* (Nutt.) Engelm]. *Theoretical and Applied Genetics* **86**, 927-934

Jordano P, Godoy JA (2000) RAPD variation and population genetic structure in *Prunus mahaleb* (Rosaceae), an animal dispersed tree. *Molecular Ecology* **9**, 1293-1305

Juniper BE, Watking R, Harris SA (1998) The origin of the apple. *Acta Horticulturae* **484**, 27-33

Kitahara K, Matsumoto S, Yamamoto T, Soejima J, Kimura T, Komatsu H, Abe K (2005) Molecular characterization of apple cultivars in Japan by S-RNase and SSR markers. *Journal of the American Society for Horticultural Science* **130**, 885-892

Koller B, Lehmann A, McDermott JM, Gessler C (1993) Identification of apple cultivars using RAPD markers. *Theoretical and Applied Genetics* **85**, 901-904

Luby J (2003) Taxonomic classification and brief history. In: Ferree DC (Eds) *Apples: Botany, Production and Uses*, Ohio State University, USA and HortResearch, Warrington, New Zealand, pp 1-14

Mulcahy DL, Cresti M, Sansavini S, Douglas GC, Linskens HF, Mulcahy GB, Vignani R, Pancaldi M (1993) The use of random amplified polymorphic DNAs to fingerprint apple genomes. *Scientia Horticulturae* **54**, 89-96

Nei M (1972) Genetic distance between populations. *American Naturalist* **106**, 283-392

Nei M (1978) Estimation of average heterozygocity and genetic distance from small number of individuals. *Genetics* **89**, 583-590

Noiton LAM, Alspach P (1996) Founding clones, inbreeding, coancestry, and status number of modern apple cultivars. *Journal of the American Society for Horticultural Science* **121**, 773-782

Peakall R, Smouse PE (2006) GENALEX: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295

Royo JB, Itoiz R (2004) Evaluation of the discriminance capacity of RAPD, isoenzymes and morphologic markers in apple (*Malus* x *domestica* Borkh.) and the congruence among classifications. *Genetic Resources and Crop Evolution* **51**, 153-160

Volk GM, Richards CM, Reilley AA, Henk AD, Forsline PL, Aldwinckle HS (2005) *Ex situ* conservation of vegetatively propagated species: Development of a seed-based core collection for *Malus sieversii*. *Journal of the American Society for Horticultural Science* **130**, 203-210

Zhao NX, Gao YB, Wang JL, Ren AZ (2008) Population structure and genetic diversity of *Stipa grandis* P. Smirn, a dominant species in the typical steppe of northern China. *Biochemical Systematics and Ecology* **36**, 1-10

Zhou ZQ, Li YN (2000) The RAPD evidence for the phylogenetic relationship of the closely related species of cultivated apple. *Genetic Resources and Crop Evolution* **47**, 353-357