Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/sjar

Spanish Journal of Agricultural Research 2009 7(3), 572-584
ISSN: 1695-971-X

Assessment of the genetic and phenotypic diversity maintained in apple core collections constructed by using either agro-morphologic or molecular marker data

L. G. Santesteban*, C. Miranda and J. B. Royo

Dpto. Producción Agraria, Universidad Pública de Navarra (UPNa). Campus de Arrosadia, 31006 Pamplona, Spain.

Abstract

Various types of data have been used for sampling plant core collections, including morphological, agronomic and eco-geographical traits, and molecular and biochemical markers. However, little is known about the ability of woody perennial core collections to retain the diversity and structure of the whole collection for characters that were not considered in the selection, especially when molecular markers are used. In this study, three core subsets were established for the apple germplasm bank curated at the Public University of Navarre (UPNa, Spain): based upon the diversity found with 10 SSR markers, another based upon the diversity assessed with 12 isozyme loci; and a third based upon morpho-agronomic diversity evaluated by 23 morpho-agronomic traits. Comparisons between these three subsets and to the whole collection were assessed to determine the impact of the data used in the selection on phenotypic and genetic diversity and on their population structure. The three subsets had a similar diversity and they did not differ from the original collection, according to Nei and Shannon-Weaver indices. The allelic/class frequencies were also always maintained in the three subsets. Overall, the kind of data used to constitute a core collection had little influence on the phenotypic and genetic diversity retained, so in the case of apple collections the use of molecular markers is preferable for this task because they allow a rapid and reliable characterization.

Additional key words: characterization, germplasm, isozyme, *Malus* × *domestica*, microsatellite, pomology, stepwise clustering.

Resumen

Evaluación de la diversidad genética y fenotípica retenida en colecciones nucleares de manzano formadas a partir de caracterizaciones agro-morfológicas o moleculares

Las colecciones nucleares de plantas pueden formarse a partir de la información obtenida en caracterizaciones morfológicas, agronómicas o eco-geográficas, así como a partir de caracterizaciones bioquímicas o moleculares. Un aspecto poco estudiado de la formación y validación de colecciones nucleares es la capacidad de éstas para conservar la estructura y diversidad de la colección global en aquellos caracteres que no han sido empleados para formar la colección nuclear. En el presente estudio se determinaron tres colecciones nucleares para el Banco de Germoplasma de Manzano de la Universidad Pública de Navarra: para el primero se emplearon los datos de caracterización de 10 marcadores SSR, la segunda fue formada a partir de la caracterización hecha con 12 loci de isoenzimas, y la tercera se construyó a partir de la información proporcionada por 23 caracteres morfo-agronómicos. Se comparó la estructura y diversidad genotípica y fenotípica de las tres colecciones, entre sí y respecto de la colección original, para determinar el impacto que sobre dicha estructura y diversidad tenía el tipo de dato con que se formó la colección nuclear. Las tres colecciones conservaron una diversidad similar y no difirieron de la original en sus índices de Nei y Shannon-Weaver, y siempre se mantuvieron las frecuencias de las clases alélicas o fenotípicas. En conjunto, el tipo de caracterización empleada en la formación de la colección nuclear tuvo escasa influencia sobre la diversidad retenida, de lo que se concluye que en el manzano los microsatélites son especialmente adecuados para formar dichas colecciones.

Palabras clave adicionales: agrupación por pasos, caracterización, germoplasma, isoenzimas, *Malus* × *domestica*, microsatélites, pomología.

* Corresponding author: gonzaga.santesteban@unavarra.es
Received: 05-11-08. Accepted: 16-06-09.

All authors are members of the SECH.

Introduction

The conservation of biodiversity can result in large collections that are difficult to characterize, evaluate, utilize and maintain. To facilitate the characterization process, core subsets have been developed to represent the genetic diversity found within the entire collection (Brown, 1989). One of the objectives of these subsets is to reduce the collection to a manageable size that facilitates the systematic and rigorous characterization of a more tractable set of genotypes within current budget and time constraints. Various types of data have been used to analyze the genetic diversity in the collections, including morphological, agronomic and eco-geographical traits or molecular and biochemical markers (van Hintum *et al.*, 2000), and each of these has its advantages and disadvantages. In the case of crops, core collections have been habitually obtained from passport, eco-geographical and morphological data (Balakrishnan *et al.*, 2000; Grenier *et al.,* 2000; Hu *et al.,* 2000; Li *et al.*, 2004), because these collections are usually very large (thousands of accessions) and the use of molecular markers in the entire collection is not feasible because of the cost (Grenier *et al.*, 2000; Hu *et al.*, 2000). Woody perennial collections constitute a different case, because they usually hold fewer accessions, maintained as clones in field genebanks. For instance, the world's largest *Malus* germplasm collections maintain over no more than 2500 accessions, and most of national or regional collections hold less than 500 accessions (Hokanson *et al.*, 1998; Maggioni *et al.*, 2004). Additionally, management costs in field banks are usually higher than those to maintain seed banks, they are highly vulnerable to environmental conditions, and field analyses are usually a lengthy and costly process, as it may be necessary waiting for several years from planting the trees to obtain their first crop. Molecular markers provide here a cost-feasible alternative that can be used on young, non-bearing trees, but nowadays the examples of the use of molecular markers to develop core collections in woody perennial species are yet restricted to few cases in apple (Laquidain *et al.,* 2005; Volk *et al.,* 2005), pear (Santesteban *et al.,* 2008), table grapevine (Jiménez-Cantizano *et al*., 2008), cashew (Dhanaraj *et al.,* 2002), sandalwood (Shashidhara *et al.,* 2003) or cherimoya (Escribano *et al.*, 2008). Two important issues in the context of using molecular markers for constructing core collections could explain this.

The first issue relates to the number of markers needed to provide a subset which represents the genetic structure of the collection. In apple, isoenzymatic characterization can reveal high levels of diversity within collections, but their relatively low level of polymorphism does not provide enough information to unambiguously differentiate all the accessions in a collection (Itoiz and Royo, 2003), so complementary molecular characterization using DNA-based markers is required (Royo and Itoiz, 2004). In the last years, microsatellite markers (SSR) have become favored for establishing unique genetic identities or fingerprints and assessing genetic diversity within a collection, because of their high polymorphism level, reproducibility and relative ease of analysis (Schlötterer, 2004). Several studies have found that a reduced set of SSR markers (around 6-12) can unambiguously differentiate accessions (Hokanson *et al.*, 1998; Laurens *et al.,* 2004; Galli *et al*., 2005; Kitahara *et al*., 2005; Guarino *et al*., 2006).

The second issue relates to the ability of core collections constituted from molecular markers to retain the agro-morphological diversity from the entire collection. This has particular relevance for collections not yet evaluated and for which the cost of field analyses is much higher than for molecular analyses, such as many woody perennial collections. This issue was studied in a limited manner by Grenier *et al.* (2000) in sorghum [*Sorghum bicolor* (L.) Moench]; they considered a small part or their collection (521 accessions out of >36,000) as the entire collection and used five SSR as sample criteria to establish a 10% size subset. They found that molecular sampling on that set did not induce changes in the distribution for most of morpho-agronomic characters. The authors are not aware of similar studies performed on woody perennial collections.

The main goal of this work was to determine the ability of morphological and molecular markers (isozymes or SSR) to constitute core collections representative of the genetic and phenotypic diversity conserved in an apple collection maintained *ex situ*.

Abbreviations used: AFLP (amplified fragment length polymorphism), CI (core subset obtained from isozyme data), CM (core subset obtained from morphological data), CN (number of classes defined for each quantitative trait), CS (core subset obtained from SSR data), DI (Nei [1987] gene diversity index), H' (Shannon-Weaver diversity index; Poole, 1974), PCA (principal component analysis), SSR (short sequence repeat, microsatellite), UPGMA (unweighted pair-group average method), UPNa (Public University of Navarre), WC (whole collection).

Material and methods

Plant material

The data set obtained in this study was obtained from the phenotypic and genotypic analysis of 254 apple (*Malus* × *domestica* Borkh) accessions maintained at the Public University of Navarre (UPNa, Spain) germplasm bank. This collection comprises mostly ancient cultivars native to Northern Navarre (42-43ºN, 1-3ºE).

Microsatellite analysis

Genomic DNA was extracted from young leaves by the Fulton *et al.* (1995) protocol. Ten SSR primer pairs were selected from the list of 140 described by Liebhard *et al.* (2002), namely: NZ02b01, CH01b12, CH01e12, CH01f02, CH01h10, CH02c09, CH02d08, CH04e05, CH01h02 and COL. The first five primer pairs were selected due to their demonstrated high level of polymorphism (Guilford *et al.,* 1997; Gianfranceschi *et al*., 1998) and the other five primer pairs belong to the list of nine 'standard' primers proposed by EGP/GR working group in *Pyrus/Malus* (Maggioni *et al*., 2004). PCR amplifications were performed as described in Gianfranceschi *et al.* (1998). PCR products were separated on acrylamide gel using a Sequi-Gen GT (Bio-Rad) sequencing unit, and visualized by silver staining (Bassam *et al*., 1991). A 30-330 bp AFLP DNA ladder (Invitrogen Life Technologies, Barcelona, Spain) was used as molecular size standard.

Isozyme analysis

The isoenzymatic profiles obtained in a previous work (Itoiz and Royo, 2003) for each accession at 12 loci using the enzyme systems ACP, ENP, EST, GOT, LAP, PER and SOD were used as source data.

Morpho-agronomic traits analysis

For observations on fruit, 10 typical fruits were selected out of a minimum of 20 from each accession, as defined by UPOV (1995). Similar criteria were used for leaf and flower observations. A total of 68 morpho-agronomic traits indicated by UPOV (1981, 1995), IPGBR (1982) and Bore and Fleckinger (1997) were evaluated in 1999-2005. Principal component analysis (PCA) was used to select the most discriminant traits (Pereira *et al*., 2003) and 23 traits were selected from the 68 initial traits. These traits had the highest eigenvectors in the first 21 principal components, which accounted for 70% of the variance. Eleven traits were quantitative (Table 1), and 12 were qualitative and scored as class variables (Table 2). Quantitative and qualitative traits were combined to perform a single cluster analysis after the transformation of the quantitative traits into qualitative ones. The number of classes (CN) for each quantitative trait was defined as follows:

$$CN = \frac{R_B}{\left(\frac{\sum_{i=1}^{n} R_{Ai}}{n}\right) + 1.15\left(\frac{\sum_{i=1}^{n} SD_{Ai}}{n}\right)}$$

Table 1. List of the morphological qualitative traits used in this study

Trait	Nº of possible classes	Described in
Flower: Color at ballon stage	6	UPOV (1995)
Flower: Petal color at bloom stage	3	This study[a]
Flower: Point of styles welding	2	UPOV (1981)
Fruit: Aperture of eye	3	UPOV (1995)
Leaf: Attitude in relation to shoot	3	UPOV (1995)
Leaf: General contour	6	Bore and Fleckinger (1997)
Leaf: Pubescence on lower side	3	UPOV (2005)
Shoot: Pubescence	5	UPOV (1995)
Susceptibility: Green aphid (*Aphis pomi* De Geer)	3	IPGBR (1982)
Susceptibility: Red mite (*Panonychus ulmi* Koch)	3	IPGBR (1982)
Susceptibility: Scab (*Venturia inequalis* (Cke.) Wint.)	3	IPGBR (1982)
Tree: Bearing habit	5	IPGBR (1982)

[a] Levels 1:white; 2: pinkish white, 3: pink.

Table 2. List of the morphological quantitative traits used in this study. For each trait, the highest (Max) and lowest (Min) average trait values observed in the bank, their difference (R_B), the mean of the differences between the highest and lowest trait values observed within each accession (mean R_A), the mean of the standard deviation of the variable values observed within each accession SD_A, and the number of qualitative classes defined (CN) are shown

Quantitative trait (UPOV, 1995)	Min	Max	R_B	Mean R_A	Mean SD_A	CN
Shoot: Length of internode (cm)	0.81	3.6	2.78	1.33	0.46	2
Leaf blade: ratio length/width	1.10	2.0	0.90	0.27	0.14	2
Petiole: length (cm)	1.30	3.9	2.60	0.70	0.29	3
Fruit: size (g)	61.80	503.2	441.40	46.06	35.62	5
Fruit: thickness of stalk (cm)	0.20	0.6	0.43	0.19	0.08	2
Fruit: lenght of stalk (cm)	0.56	2.9	2.34	1.18	0.48	2
Fruit: ground colour (CIE a/b)	-0.49	0.3	0.78	0.20	0.10	3
Fruit: amount russet on cheeks (%)	0.00	100.0	100.00	41.13	22.92	2
Fruit: firmness (kg)	3.90	12.0	8.10	2.72	1.32	2
Time of beginning of flowering (d)	-3.30	23.4	26.70	6.84	3.45	3
Time of maturity (d)	95.00	190.5	95.50	7.86	7.64	6

where n is the number of accessions in the germplasm bank, R_B the difference between the highest and lowest average trait values observed between accessions in the bank, R_A the difference between the highest and lowest trait values observed within each accession, and SD_A the standard deviation of the variable values observed within each accession. CN value was rounded to the closest integer. The purpose of this procedure was to define class intervals broader than the average differences that can be found for each character within an accession, because the procedures available in the literature (Kaufman and Rosseeuw, 1990; Pecetti *et al*., 1992) were considered as unsatisfactory. Class intervals for each redefined variable are shown on Table 2. Binomial variables were then created for each of the classes, where 1 or 0 indicated presence or absence of the observed state in the accession (Kaufman and Rosseeuw, 1990). Overall, the 23 traits were split into 75 binomial variables.

Construction of core collections

The procedure of stepwise clustering with random sampling proposed by Hu *et al*. (2000) was employed in developing three core subsets, one from each source of data: the subset named CS was obtained from microsatellite data, CI subset was obtained from isozyme data and CM from morpho-agronomic traits data. Dice distance and unweighted pair-group average method (UPGMA) of hierarchical cluster analysis (Sneath and Sokal, 1973) were used for grouping accessions. The stepwise cluster analysis was conducted until selected accessions were reduced to ≈20% of the initial collection (Crossa *et al*., 1995; Yonezawa *et al*., 1995). The number of accessions eventually retained on each core collection was 55 (21.7%) for the CI subset, 57 (22.4%) for the CS subset and 50 (19.7%) for the CM subset. The analyses and the phenograms were computed with the NTSYS-pc, ver. 2.11w (Rohlf, 1993).

Evaluation of the diversity in the collections

For each core subset and for the whole collection, the gene diversity index per locus and trait (corrected for sample size) was calculated as

$$DI = \frac{n\left(1 - \sum_{i=1}^{n} p_i\right)}{n-1}$$

where n is the number of phenotypic classes or alleles, and p_i the proportion of the total number of accessions in the i^{th} class (Nei, 1987). The diversity was also estimated for each locus and trait by the Shannon-Weaver Diversity Index defined by Poole (1974) as

$$H'_C = \sum_{i=1}^{n} p_i \log_e p_i$$

Due to its additive property (Poole, 1974), Shannon-Weaver diversity indices obtained for each character were pooled. The significance of the Shannon diversity indices among the subsets and the whole collection was performed by the *t*-test described by Hutcheson (1970). Morpho-agronomic traits and biochemical marker fre-

quency distributions were compared between the three subsets and to the whole collection with a chi-square goodness of fit test. The tests were performed only on the frequent alleles/classes (those with frequency higher than 5%).

Validation of core collections

The representativeness of the three core subsets was validated according to the following criteria [modified from Brown (1989) and Grenier *et al.* (2000)]: a) recovery of all the alleles/classes present in more than five accessions in the whole collection, b) no significant differences in frequency distribution of alleles/classes in at least 95% of loci/traits between the core and the whole collection, and c) no significant differences in variability parameters between the core and the whole collection.

Results

Comparison of the whole collection and the core subsets for SSR loci

A total of 159 alleles were recorded for the 10 analyzed SSR loci, out of which 59 were present in more than 5% of the accessions (frequent alleles). In the whole collection (WC), the mean allelic richness was 15.9 alleles per locus, and ranged from 8 alleles at the locus CH01e12 to 26 alleles at the locus CH02d08. The gene diversity (DI) was relatively high and ranged (Table 3) between 0.72 and 0.94 (mean of 0.83). Comparison between the whole collection and the core subsets revealed similar tendencies, no matter which data were originally used to create the subset: The mean allelic richness was of the same magnitude (14.5, 13.9 and 13.9 alleles per locus for the CS, CI and CM subsets), and all the frequent alleles were present in the three subsets. The genetic diversity (DI) retained was high, with similar ranges and mean values to that of the whole collection not only for the CS subset created from SSR data, but also for CI and CM subsets. The Shannon-Weaver diversity indices (H') for the whole collection and the three subsets were also statistically similar (H'=2.17±0.47 for the WC and 2.22±0.47, 2.12±0.42 and 2.11±0.48 for the CS, CI and CM subsets, respectively).

When compared to the whole collection, the chi-square tests showed that the three core subsets had a similar allelic distribution for all of the 10 loci (Table 4). Comparisons between subsets showed significant differences between CS and CI at one loci (COL), and between CM and CI at two loci (CH02d08 and COL), whereas CS and CM had similar allelic distributions.

Table 3. Microsatellite diversity assesed with the number of frequent alleles (A) in the whole collection (WC), the number of rare alleles (B) in the WC, the number of alleles that were rare in the WC and are present in each subset (C), and the genetic diversity (DI) as defined by Nei (1987)

Locus	WC			CS[1]		CI		CM	
	A	B	DI	C	DI	C	DI	C	DI
NZ02b01	8	13	0.88	11	0.91	8	0.88	9	0.89
CH01b12	4	10	0.82	9	0.84	10	0.86	8	0.79
CH02c09	11	4	0.92	4	0.93	3	0.92	4	0.93
COL	6	12	0.88	11	0.91	11	0.85	9	0.88
CH02d08	8	18	0.94	15	0.95	12	0.94	15	0.94
CH04e05	3	12	0.75	10	0.73	11	0.77	11	0.77
CH01e12	5	3	0.79	3	0.81	3	0.78	3	0.77
CH01f02	6	16	0.90	12	0.92	10	0.90	12	0.90
CH01h02	5	5	0.73	4	0.76	5	0.70	3	0.72
CH01h10	3	7	0.72	7	0.75	7	0.72	6	0.74
Total	59	100		86		80		80	
Mean	5.9	10.0	0.83	8.6	0.85	8.0	0.83	8.0	0.83

[1] CS: Core subset created from microsatellite data; CI: core subset created from isozyme data; CM: Core subset created from morphological data.

Table 4. Comparison for the frequency distribution for microsatellite markers through a *Chi*[2] test (A) between the three subsets and the whole collection, and (B) between the three subsets

Locus	(A) Comparison to the whole collection[1]						(B) Comparison between subsets					
	CS *vs* WC[a]		CI *vs* WC		CM *vs* WC		CS *vs* CI		CS *vs* CM		CI *vs* CM	
NZ02b01	4.84	NS[2]	8.69	NS	1.36	NS	11.35	NS	7.43	NS	9.66	NS
CH01b12	0.47	NS	3.20	NS	2.16	NS	5.61	NS	1.46	NS	5.26	NS
CH02c09	4.94	NS	4.36	NS	6.61	NS	6.34	NS	9.29	NS	9.22	NS
COL	5.50	NS	4.52	NS	3.40	NS	23.34	*	7.87	NS	19.20	*
CH02d08	5.08	NS	9.68	NS	12.31	NS	9.93	NS	16.27	NS	30.67	*
CH04e05	0.22	NS	0.74	NS	2.15	NS	0.24	NS	3.72	NS	3.33	NS
CH01e12	2.92	NS	3.06	NS	3.98	NS	2.99	NS	3.55	NS	2.21	NS
CH01f02	5.66	NS	0.97	NS	7.92	NS	8.54	NS	2.18	NS	4.68	NS
CH01h02	6.14	NS	8.29	NS	0.75	NS	4.29	NS	2.99	NS	5.64	NS
CH01h10	2.21	NS	2.79	NS	0.72	NS	2.08	NS	2.47	NS	5.35	NS

[1] CS: Core subset created from microsatellite data; WC: Whole collection; CI: core subset created from isozyme data; CM: Core subset created from morphological data. [2] NS: non-significant differences at $P<0.05$ level; *: significant differences at $P<0.05$ level.

In the whole collection, 100 alleles have had a frequency lower than 5%. From these 100 rare alleles, 86 were maintained in the CS subset, and three of them were present only in this subset. The CI and CM subsets retained a lower number (80) of rare alleles, out of which five were found only in the CI subset, and two in the CM subset. The rare alleles lost in the three subsets were always present in less than five accessions of the whole collection.

Comparison of the whole collection and the core subsets for isozyme loci

The level of polymorphism detected in the whole collection by the 12 isozyme loci was much lower than the observed with SSRs (Table 5), as only 33 alleles were recorded (just two of them were rare). The mean allelic richness (2.75 alleles per locus) and diversity indices (mean DI=0.48) were accordingly much lower. Comparisons between whole collection and CI subset revealed that all the frequent alleles (31) were maintained and their allelic richness and gene diversity had similar values. The latter was also true for the CM and CS subsets. Similar results were obtained when diversity was evaluated through H' diversity indices (H'=0.78±0.30 for the WC and 0.82±0.28, 0.78±0.29 and 0.75±0.29 for the CI, CS and CM subsets, respectively).

When allelic distribution was compared (Table 6), CI and CM subsets had one locus (*SOD-4*) significantly different to the WC, whereas the CS subset had similar allelic frequencies at all the 12 loci. Comparisons between the three subsets showed similar allelic distributions for nearly all loci: CI and CS subsets differed only at *Est-1* loci, CI and CM subsets differed at two loci (*Sod-3* and *Sod-4)*, while CS and CM showed different distributions for three loci (*Got*, *Sod-3* and *Sod-4)*.

There were only two rare alleles (at *Acp-2* and *Per-2* loci) in the whole collection. Though the rare allele in *Acp-2* was maintained in all the subsets, only the CI subset could retain the rare allele in *Per-2,* which was present in only three accessions of the whole collection.

Comparison of the core subsets and the whole collection for morpho-agronomic traits

A total of 75 phenotypic classes were recorded in the collection out of the 77 possible, the absent classes were *'Wijcik habit'* level for *Tree Bearing Habit* and 'Strong' level for *Leaf pubescence*. The mean phenotypic richness in the collection was 3.26 classes per trait (Table 7) and the phenotypic diversity level ranged from 0.08 for the trait *Leaf pubescence* to 0.79 for *Time of maturity* (mean of 0.49). As with molecular markers, comparisons between whole collection and CM subset showed that all the frequent classes (64) were retained

Table 5. Isozyme diversity assessed with the number of frequent alleles (A) in the whole collection (WC), the number of rare alleles (B) in the WC, the number of alleles that were rare in the WC and are present in each subset (C), and the genetic diversity (DI) as defined by Nei (1987)

Locus	WC			CI[1]		CS		CM	
	A	B	DI	C	DI	C	DI	C	DI
Acp-2	2	1	0.52	1	0.54	1	0.52	1	0.52
Enp-1	2	0	0.40	0	0.40	0	0.37	0	0.38
Est-1	2	0	0.41	0	0.46	0	0.34	0	0.41
Est-4	2	0	0.40	0	0.44	0	0.41	0	0.44
Got-2	3	0	0.49	0	0.54	0	0.56	0	0.46
Lap-3	4	0	0.63	0	0.63	0	0.63	0	0.62
Per-1	3	0	0.58	0	0.57	0	0.56	0	0.61
Per-2	4	1	0.75	1	0.76	0	0.74	0	0.75
Per-3	2	0	0.24	0	0.30	0	0.25	0	0.25
Sod-2	3	0	0.63	0	0.65	0	0.65	0	0.63
Sod-3	2	0	0.43	0	0.49	0	0.45	0	0.34
Sod-4	2	0	0.23	0	0.35	0	0.30	0	0.13
Total	31	2		2		1		1	
Mean	2.6	0.2	0.48	0.2	0.51	0.1	0.48	0.1	0.46

[1] See Table 4.

and diversity had similar values. Once again, similar results were obtained for the CS and CI subsets, which did not included phenotypic traits as source data. Finally, statistically similar results between WC and the three subsets were obtained for H' diversity indices (H'=0.85±0.39 for the WC and 0.92±0.38, 0.86±0.41 and 0.85±0.40 for the CM, CS and CI subsets, respectively).

Table 6. Comparison for the frequency distribution for isozyme markers through a *Chi*[2] test (A) between the three subsets and the whole collection, and (B) between the three subsets

Locus	(A) Comparison to the whole collection[1]						(B) Comparison between subsets					
	CS *vs* WC		CI *vs* WC		CM *vs* WC		CS *vs* CI		CS *vs* CM		CI *vs* CM	
Acp-2	1.70	NS[2]	0.08	NS	0.07	NS	2.49	NS	1.09	NS	0.29	NS
Enp-1	0.45	NS	0.00	NS	0.21	NS	0.44	NS	0.05	NS	0.20	NS
Est-1	1.64	NS	0.91	NS	0.01	NS	6.02	*	1.39	NS	1.14	NS
Est-4	0.03	NS	1.10	NS	1.02	NS	0.77	NS	0.65	NS	0.00	NS
Got-2	3.44	NS	3.25	NS	1.24	NS	0.40	NS	6.41	*	4.33	NS
Lap-3	0.06	NS	3.45	NS	0.36	NS	3.26	NS	0.17	NS	3.53	NS
Per-1	1.64	NS	0.36	NS	1.60	NS	0.69	NS	4.81	NS	2.52	NS
Per-2	0.77	NS	0.08	NS	0.24	NS	0.72	NS	1.24	NS	0.09	NS
Per-3	0.03	NS	1.70	NS	0.01	NS	1.23	NS	0.01	NS	1.42	NS
Sod-2	1.17	NS	0.99	NS	0.10	NS	0.00	NS	2.04	NS	1.76	NS
Sod-3	0.13	NS	2.03	NS	2.69	NS	1.10	NS	5.17	*	11.88	*
Sod-4	2.49	NS	8.28	*	3.99	*	1.36	NS	22.72	*	41.60	*

[1,2]: See Table 4.

Table 7. Morphological diversity assessed with the number of frequent classes (A) in the whole collection (WC), the number of rare classes in the WC (B), the number of classes that were rare in the WC and present in each subset (C), and with the genetic diversity DI as defined by Nei (1987)

Trait	WC			CM[1]		CS		CI		CR	
	A	B	DI	C	DI	C	DI	C	DI	C	DI
Flower: Color at ballon stage	4	2	0.70	2	0.77	2	0.73	1	0.68	1	0.71
Flower: Petal color at bloom stage	2	1	0.51	1	0.53	0	0.51	1	0.51	0	0.51
Flower: Point of styles welding	2	0	0.35	0	0.45	0	0.42	0	0.39	0	0.40
Fruit: amount russet on cheeks (%)	2	0	0.11	0	0.22	0	0.07	0	0.00	0	0.19
Fruit: Aperture of eye	3	0	0.63	0	0.68	0	0.66	0	0.62	0	0.63
Fruit: firmness (kg)	2	0	0.49	0	0.51	0	0.48	0	0.46	0	0.50
Fruit: ground colour (CIE a/b)	2	1	0.52	1	0.48	1	0.56	1	0.52	1	0.57
Fruit: lenght of stalk (cm)	2	0	0.21	0	0.36	0	0.22	0	0.31	0	0.19
Fruit: size (g)	3	2	0.66	1	0.69	2	0.66	1	0.64	1	0.60
Fruit: thickness of stalk (cm)	2	1	0.35	0	0.41	0	0.29	0	0.26	1	0.39
Leaf blade: ratio length/width	2	0	0.47	0	0.48	0	0.49	0	0.51	0	0.37
Leaf: Attitude in relation to shoot	3	0	0.66	0	0.66	0	0.64	0	0.64	0	0.67
Leaf: General contour	4	2	0.60	2	0.73	1	0.67	2	0.66	1	0.60
Leaf: Pubescence on lower side	2	0	0.08	0	0.14	0	0.07	0	0.14	0	0.08
Petiole: length (cm)	3	0	0.36	0	0.54	0	0.38	0	0.44	0	0.40
Shoot: Length of internode (cm)	2	0	0.49	0	0.49	0	0.51	0	0.51	0	0.50
Shoot: Pubescence	4	1	0.74	1	0.72	1	0.77	1	0.77	1	0.75
Susceptibility: Green aphid (*Aphis pomi*)	3	0	0.64	0	0.67	0	0.66	0	0.65	0	0.62
Susceptibility: Red mite (*Panonychus ulmi*)	3	0	0.56	0	0.61	0	0.58	0	0.63	0	0.60
Susceptibility: Scab (*Venturia inequalis*)	2	0	0.24	0	0.34	0	0.25	0	0.25	0	0.18
Time of beginning of flowering (d)	3	0	0.54	0	0.57	0	0.52	0	0.60	0	0.53
Time of maturity (d)	5	1	0.79	1	0.80	1	0.78	1	0.82	1	0.80
Tree: Bearing habit	4	0	0.65	0	0.72	0	0.70	0	0.65	0	0.61
Total	64	11		9		8		8		7	
Mean	2.78	0.48	0.49	0.39	0.55	0.35	0.50	0.35	0.51	0.30	0.50

[1] See Table 4.

When compared to the collection (Table 8), CM had two traits (*petiole length* and *stalk length*) with a significantly different class distribution, whereas CS and CI subsets had similar frequencies for all the 23 traits. Comparisons between subsets showed significant differences between CM and CS for one trait (*russet on cheeks*). The class distribution in the CM and CI subsets was significantly different for three traits (s*talk thickness, russet on cheeks* and *color of petals*). Also, the CS and CI had two traits (*fruit size* and *color of petals*) with significantly different class distributions.

From the 11 rare classes observed in the collection, nine were maintained in the CM subset. An equivalent number of rare classes (eight) were maintained in the CI and CS subsets, respectively. Once again the rare classes lost in the core subsets were present in less than five accessions of the collection.

Discussion

The number of apple accessions (254) maintained by the UPNa in their collection averages the usual size for European apple regional collections, which ranges from ≈100 to 500 accessions (Maggioni *et al.,* 2004). Despite its narrow geographical origin, mostly ancient cultivars native to Northern Navarre, our collection has proven to be rather diverse: the allelic richness and phenotypic/genetic diversity levels found here are similar than those reported for other local European collections. Laurens *et al.* (2004) assessed the genetic diversity among 142 French apple local cultivars belonging to various morphological groups from several geographic regions with a nine SSR set and found from 11 to 21 alleles per locus (mean of 15.3) and an average DI close to 0.8, values nearly identical to ours. Moreover, Wagn-

Table 8. Comparison for the frequency distribution for the morphological traits through a *Chi*[2] test (A) between the three subsets and the whole collection, and (B) between the three subsets

Trait	(A) Comparison to the whole collection [1]						(B) Comparison between subsets					
	CS *vs* WC		CI *vs* WC		CM *vs* WC		CS *vs* CI		CS *vs* CM		CI *vs* CM	
Flower: color at ballon stage	7.41	NS[2]	2.96	NS	2.57	NS	19.83	*	5.07	NS	7.89	*
Flower: petal color at bloom stage	0.03	NS	2.08	NS	0.09	NS	1.65	NS	0.02	NS	1.33	NS
Flower: point of styles welding	1.28	NS	0.38	NS	3.57	NS	0.17	NS	0.47	NS	1.09	NS
Fruit: amount russet on cheeks (%)	0.64	NS	3.40	NS	2.32	NS	0.00	NS	4.21	*	7.57	*
Fruit: aperture of eye	0.80	NS	0.58	NS	4.02	NS	2.16	NS	1.25	NS	5.58	NS
Fruit: firmness (kg)	0.56	NS	1.31	NS	0.52	NS	0.18	NS	2.15	NS	3.34	NS
Fruit: ground colour (CIE a/b)	0.00	NS	1.29	NS	2.54	NS	1.62	NS	2.74	NS	0.17	NS
Fruit: lenght of stalk (cm)	0.01	NS	2.41	NS	6.38	*	2.13	NS	3.59	NS	0.47	NS
Fruit: size (g)	1.62	NS	1.78	NS	0.49	NS	5.83	*	4.08	NS	3.09	NS
Fruit: thickness of stalk (cm)	0.66	NS	1.47	NS	1.22	NS	0.22	NS	3.13	NS	4.42	*
Leaf blade: ratio length/width	0.16	NS	1.37	NS	0.02	NS	0.60	NS	0.07	NS	1.09	NS
Leaf: attitude in relation to shoot	1.41	NS	1.56	NS	0.32	NS	0.15	NS	0.68	NS	1.07	NS
Leaf: general contour	0.09	NS	1.27	NS	1.02	NS	2.43	NS	1.08	NS	0.01	NS
Leaf: pubescence on lower side	2.58	NS	4.84	NS	4.16	NS	1.56	NS	5.45	NS	6.69	NS
Petiole: length (cm)	2.58	NS	1.19	NS	10.31	*	4.14	NS	5.10	NS	2.85	NS
Shoot: length of internode (cm)	1.06	NS	1.04	NS	0.32	NS	0.00	NS	2.62	NS	2.55	NS
Shoot: pubescence	1.10	NS	1.20	NS	3.17	NS	1.15	NS	4.80	NS	5.23	NS
Susceptibility: green aphid (*Aphis pomi*)	0.20	NS	1.49	NS	3.38	NS	2.06	NS	1.89	NS	4.95	NS
Susceptibility: red mite (*Panonychus ulmi*)	0.49	NS	2.69	NS	1.20	NS	1.84	NS	0.76	NS	0.30	NS
Susceptibility: scab (*Venturia inequalis*)	0.00	NS	0.03	NS	2.54	NS	0.01	NS	1.69	NS	1.40	NS
Time of beginning of flowering (d)	0.78	NS	0.99	NS	0.12	NS	2.42	NS	1.19	NS	0.29	NS
Time of maturity (d)	2.59	NS	1.36	NS	2.34	NS	3.49	NS	4.03	NS	4.93	NS
Tree: bearing habit	0.93	NS	2.25	NS	4.25	NS	2.65	NS	0.91	NS	5.26	NS

[1,2] See Table 4.

er *et al.* (2004) used nine isozyme loci, to assess genetic diversity on 321 old German cultivars, and their reported mean Shannon-Weaver index (H'=0.776) nearly identical to ours (H'=0.780).

The three core subsets developed in this study have revealed as representative of the phenotypic and genetic diversity of the collection, because they retained all the frequent allele/classes, and all of the rare ones present in more than five accessions in the collection, without altering their frequency distributions or losing diversity. Since the concept of the core collection was proposed (Brown, 1989), various types of data, such as morphological, agronomical or ecogeographical traits, molecular and biochemical markers etc., have been used to analyze genetic diversity (van Hintum *et al*., 2000). To our knowledge, the ability of those different types of data to constitute core collections adequately retaining the overall diversity within a collection has been analyzed only by Grenier *et al.* (2000) in sorghum and in a limited manner. In our study, comparative analyses for each character and locus between the three subsets and the whole collection illustrate some minor differences only at the conservation of very rare alleles/classes: each strategy tends to preserve slightly better the rare alleles of those characters used for the constitution of the subset. Comparisons between the three subsets highlighted the similar genetic diversity retained by each sampling. Allelic and class frequencies were found significantly similar for nearly all loci between the subsets. Overall, our results confirm the preliminary work of Grenier *et al*. (2000) and suggest that the phenotypic and genetic diversity is not affected as a result of the kind of characters used for the constitution of the specific core collection. Therefore, this particular issue

seems as of little relevance when deciding what kind of data should be used, so that other criteria such as time, cost or ease for data acquisition should be taken into account. In large collections with thousands of accessions, such as many crop collections, assessing the entire collection with molecular markers (even with small sets) is generally accepted as not feasible because of the cost (Grenier *et al*., 2000). Moreover, a small number of morphological characters (≈20-30) are generally described for annual crop species, where less than a half of them suffice and are required for assessing distinctness (UPOV, 1989, 1996, 1998). Those are some of the reasons why core collections for crop species have been traditionally obtained using a reduced set of passport, eco-geographical and morphological data (Balakrishnan *et al*., 2000; Hu *et al*., 2000; Li *et al*., 2004) and only at a later stage the genetic diversity of the constituted core collection has been assessed with molecular markers (Hokanson *et al*., 1998, Grenier *et al*., 2000, Liu *et al*., 2001). However, there are reasons to take a different approach to constitute core collections for many woody perennial species. Although these collections are usually constituted by less than a thousand accessions (Maggioni *et al*., 2004), many fruit tree species have a much higher number of morphological characters described (50-70 characters), and more than a half of them are required for assessing distinctness (UPOV, 2000, 2005). Moreover, the population structure of conserved germplasm in most cultivated woody perennials differs from that of annual species because, in vegetatively propagated fruit tree species, the domestication process has involved usually few recombination cycles and, consequently, domesticated genotypes are only a few generations apart from their wild ancestors (Escribano *et al*., 2008). Moreover, field analyses are a lengthy process, as it may be necessary waiting for several years from the implantation of the trees to obtain their first crop. For instance, in our collection 10 accessions, grafted in 1998, have not been yet fully characterized as they have not produced either flowers or fruit. The cost and effort of conducting the morphological characterization of a woody perennial collection can be consequently much higher than for assessing with molecular markers (Karp *et al*., 1997). Additionally, genetic sorting based on phenotypic data could not correctly reflect the genetic diversity of the initial germplasm resources due to environmental errors in the field or genotype-ambient interaction (Hu *et al*., 2000). Therefore, in regional and national apple collections not yet characterized, molecular markers could constitute a rapid and economic tool for the estimation of genetic variability and the constitution of core subsets, so in a later stage it will be feasible to perform an exhaustive and efficient characterization through agro-morphological descriptors at the core collection.

An important issue that arises in the context of constructing core collections relates to the number of markers or traits needed to provide a subset which represents not only the genetic structure of the collection but also the phenotypic structure. Pereira *et al.* (2003) assessed the diversity within the Galician apple germplasm collection with 24 morpho-agronomic traits which were selected (by means of PCA) as the most discriminant from an exhaustive characterization with 89 traits. In our study a similar number of discriminant traits (23) were selected out of a total of 68 evaluated traits and the core collection we created using them adequately retained the genetic diversity and structure assessed by isozymes and SSRs. In the case of molecular markers, it is noteworthy that a reduced set of isozyme loci (12) resulted as valid as a similar magnitude (10 loci) but much more polymorphic SSR set to obtain a representative core subset. Such number of markers is generally accepted as sufficient to establish genetic identities (Hokanson *et al.,* 1998; Laurens *et al.*, 2004; Galli *et al*., 2005; Kitahara *et al.*, 2005; Guarino *et al*., 2006). However, a different allelic distribution and lower diversity indices would have been expected for the CI subset at any of the SSR loci or morpho-agronomic traits evaluated, due to the kind of information provided by each type of marker. Isozyme variation only reflects differences in protein-coded sequences and, hence, only in a small fraction of all mutational events (Clegg, 1989). Moreover, isozyme markers have a small number of available marker loci compared to the DNA markers, and do not provide enough information to unambiguously differentiate all the accessions in a collection (Itoiz and Royo, 2003), whereas the high polymorphism, reproducibility and relative ease of SSR analysis favor the establishment of unique genetic identities or fingerprints and the assessment of genetic diversity within a collection (Schlötterer, 2004). For all those reasons, though the results presented here show isozymes and SSR as equally effective at sampling for core collections, SSR should be preferred over isozymes.

In conclusion, the results obtained in this study suggest that the kind of data used to constitute a core collection has little influence on the phenotypic and genotypic diversity retained. However, molecular markers are more adequate because a lower number of them are

enough to retain the phenotypic diversity and structure of apple collections. Moreover, molecular markers allow a more rapid characterization and they can be applied at earlier stages of tree development.

Aknowledgments

This study was funded by the Spanish Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) grants Nº RF200400008-C03-00 and Nº RFP2004-00003-00-00. The authors wish to acknowledge the Instituto Técnico y de Gestión Agrícola (ITGA) of Navarre role in the initial assembly of this collection.

References

BALAKRISHNAN R., NAIR N.V., SREENIVASAN N.V., 2000. A method for establishing a core collection of *Saccharum officinarum* L. germplasm based on quantitative-morphological data. Genet Resour Crop Evol 47, 1-9.

BASSAM B.J., CAETANO-ANOLLÉS G., GRESSHOFF P.M., 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. Analyt Biochem 196, 80-83. doi:10.1016/0003-2697(91)90120-I

BORÉ J.M., FLECKINGER J., 1997. Pommiers à cidre. Variétés de France. INRA, Paris, France. 771 pp. [In French].

BROWN A.H.D., 1989. Core collections: a practical approach to genetic resources management. Genome 31, 818-824.

CLEGG M.T., 1989. Molecular diversity in plant populations. In: Plant population genetics, breeding and genetic resources (Brown A.H.D., Clegg M.T., Kahler A.L., Weir B.S., eds). Sinauer Assoc, Sunderland, USA. pp. 99-116.

CROSSA J., DELACY I.H., TABA S., 1995. The use of multivariate methods in developing a core collection. In: Core collections of plant genetic resources (Hodgkin T., Brown A.H.D., Hintum van Th.J.L., Morales E.A.V., eds). John Wiley Sons, NY, USA. pp. 77-89.

DHANARAJ A.L., RAO E.V.V.B., SWAMI K.R.M., BHAT M.G., PRASAD D.T., SONDUR S.N., 2002. Using RAPDs to assess the diversity in Indian cashew (*Anacardium occidentale* L.) germplasm. J Hort Sci Biotech 77, 41-47.

ESCRIBANO P., VIRUEL M.A., HORMAZA J.J., 2008. Comparison of different methods to construct a core germplasm collection in woody perennial species with simple sequence repeat markers. A case study in cherimoya (*Annona cherimola*, Annonaceae), an underutilised subtropical fruit tree species. Ann Appl Biol 153, 25-32. doi:10.1111/j.1744-7348.2008.00232.x

FULTON T.M., CHUNWONGSE J., TANKSLEY S.D., 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13, 207-209. doi:10.1007/BF02670897

GALLI Z., HALÁSZ G., KISS E., HESZKY L., DOBRÁNSZKI J., 2005. Molecular identification of commercial apple cultivars with microsatellite markers. HortScience 40, 1974-1977.

GIANFRANCESCHI L., SEGLIAS N., TARCHINI R., KOMJANC M., GESSLER C., 1998. Simple sequence repeats for the genetic analysis of apple. Theor Appl Genet 96, 1069-1076. doi:10.1007/s001220050841

GRENIER C., DEU M., KRESOVICH S., BRAMEL-COX P.J., HAMON P., 2000. Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum collection using random *vs* non-random sampling procedures. B. Using molecular markers. Theor Appl Genet 101, 197-202. doi:10.1007/s001220051468

GUARINO C., SANTORO S., DE SIMONE L., LAIN O., CIPRIANI G., TESTOLIN R., 2006. Genetic diversity in a collection of ancient cultivars of apple (*Malus* × *domestica* Borkh.) as revealed by SSR-based fingerprinting. J Hort Sci Biotech 81, 39-44.

GUILFORD P., PRAKASH S., ZHU J.M., RIKKERINK E., GARDINER S., BASSET H., FOSTER R., 1997. Microsatellites in *Malus* × *domestica* (apple) abundance, polymorphism and cultivar identification. Theor Appl Genet 94, 249-254. doi:10.1007/s001220050407.

HOKANSON S.C., SZEWC-MCFADDEN A.K., LAMBOY W.F., MCFERSON J.R., 1998. Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus* × *domestica* Borkh. core subset collection. Theor Appl Genet 97, 671-683. doi:10.1007/s001220050943.

HU J., ZHU J., XU H.M., 2000. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. Theor Appl Genet 101, 264-268. doi:10.1007/s001220051478.

HUTCHESON K., 1970. A test for comparing diversities based on the Shannon formula. Theor Biol 29, 151-154. doi:10.1016/0022-5193(70)90124-4.

IPGBR, 1982. Descriptor lists for apple (*Malus*). Intl Board Plant Genet Res, Rome. 49 pp.

ITOIZ R., ROYO J.B., 2003. Isoenzymatic variability in an apple germplasm bank. Genet Resour Crop Evol 50, 391-400.

JIMÉNEZ-CANTIZANO A., VARGAS A.M., DE ANDRÉS M.T., GARCÍA DE LUJÁN A., MARTÍNEZ-ZAPATER

J.M., IBÁÑEZ J., 2008. Colecciones nucleares anidadas de las variedades de uva de mesa existentes en España. Actas de Horticultura 51, 47-48. [In Spanish].

KARP A., KRESOVICH S., BHAT K.V., AYAD W.G., HODGKIN T., 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies. IPGRI Technical Bulletin Nº 2. 47 pp.

KAUFMAN L., ROSSEEUW P.J., 1990. Finding groups in data: An introduction to cluster analysis. Wiley, NY, USA, 342 pp.

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 analyses and SSR markers. J Amer Soc Hort Sci 130, 885-892.

LAQUIDAIN M.J., MIRANDA C., SANTESTEBAN L.G., ROYO J.B., 2005. Determinación de la colección nuclear del banco de germoplasma de manzano autóctono de la UPNa. Actas Portuguesas de Horticultura 8, 107-112. [In Spanish].

LAURENS F., DUREL C.E., LASCOSTES M., 2004. Molecular characterization of French local apple cultivars using SSRs. Acta Hort 663, 639-642.

LI Y., SHI Y., CAO Y., WANG T., 2004. Establishment of a core collection for maize germplasm preserved in Chinese National Genebank using geographic distribution and characterization data. Genet Resour Crop Evol 51, 845-852.

LIEBHARD R., GIANFRANCESCHI L., KOLLER B., RYDER C.D., TARCHINI R., VAN DE WEG E., GESSLER C., 2002. Development and characterization of 140 new microsatellites in apple (*Malus* × *domestica* Borkh.). Mol Breed 10, 217-241.

LIU F., SUN G.L., SALOMON B., VON BOTHMER R., 2001. Distribution of allozymic alleles and genetic diversity in the American barley core collection. Theor Appl Genet 102, 606-615.

MAGGIONI L., FISCHER M., LATEUR M., LAMONT E.J., LIPMAN E. (eds), 2004. Report of a Working Group on *Malus/Pyrus*. Second Meeting, 2-4 May 2002, Dresden-Pillnitz, Germany. IPGRI, Rome. Italy. 96 pp.

NEI M., 1987. Molecular evolutionary genetics. Columbia University Press, NY, USA. 512 pp.

PECETTI L., ANNICCHAIARICO P., DAMANIA A.B., 1992. Biodiversity in a germplasm collection of durum-wheat. Euphytica 60, 229-238.

PEREIRA S., RAMOS A.M., ASCASIBAR J., PIÑEIRO J., 2003. Analysis of apple germplasm in northwestern Spain. J Amer Soc Hort Sci 128, 67-84.

POOLE R.W., 1974. An introduction to quantitative ecology. McGraw-Hill, NY, USA., 532 pp.

ROHLF F.J., 1993. NTSYS-PC. Numerical taxonomy and multivariate analysis system. Exeter software, NY.

ROYO J.B., ITOIZ R., 2004. Evaluation of the discriminance capacity of RAPD, isoenzymes and morphologic markers in apple (*Malus* × *domestica* Borkh.) and the congruence among classifications. Genet Resour Crop Evol 51, 153-160.

SANTESTEBAN L.G., URRESTARAZU J., URBINA V., ROYO J.B., 2008. Determinación de la colección nuclear del Banco de Germoplasma de Peral Autóctono de la UdL. Actas de Horticultura 51, 55-56. [In Spanish].

SCHLÖTTERER C., 2004. The evolution of molecular markers – just a matter of fashion? Nature Rev Gen 5, 63-69. doi:10.1038/nrg1249

SHASHIDHARA G., HEMA M,V., KOSHY B., FAROOQI A.A., 2003. Assessment of genetic diversity and identification of core collection in sandalwood germplasm using RAPDs. J Hort Sci Biotech 78, 528-536.

SNEATH P.H.A., SOKAL R.R., 1973. Numerical taxonomy. Freeman. San Francisco, USA. 573 pp.

UPOV, 1981. Guidelines for the conduct of tests for distinctness, homogeneity and stability (Apple). TG/14/5. Intl Union Protection New Varieties Plants. Geneve, Switzerland. 25 pp.

UPOV, 1989. Guidelines for the conduct of tests for distinctness, homogeneity and stability (Sorghum). TG/122/3. Intl Union Protection New Varieties Plants. 25 pp.

UPOV, 1995. Guidelines for the conduct of tests for distinctness, homogeneity and stability (Apple). TG/14/8. Intl Union Protection New Varieties Plants. Geneve, Switzerland. 29 pp.

UPOV, 1996. Guidelines for the conduct of tests for distinctness, homogeneity and stability (Wheat). TG/3/11. Intl Union Protection New Varieties Plants. Geneve, Switzerland. 55 pp.

UPOV, 1998. Guidelines for the conduct of tests for distinctness, homogeneity and stability (Soya bean). TG/80/6. Intl Union Protection New Varieties Plants. Geneve, Switzerland. 39 pp.

UPOV, 2000. Guidelines for the conduct of tests for distinctness, homogeneity and stability (Pear). TG/15/3. Intl Union Protection New Varieties Plants. Geneve, Switzerland. 42 pp.

UPOV, 2005. Guidelines for the conduct of tests for distinctness, homogeneity and stability (Apple). TG/14/9. Intl Union Protection New Varieties Plants. Geneve, Switzerland. 45 pp.

VAN HINTUM TH.J.L., BROWN A.H.D., SPILLANE C., HODGKIN T., 2000. Core collections of plant genetic resources. IPGRI Technical Bulletin No. 3. Rome, Italy. 48 pp.

VOLK G.M., RICHARDS C.M., REILLEY A.A., HENK A.D., FORSLINE P.L., ALDWINKLE H.S., 2005. Ex situ conservation of vegetatively propagated species: development of a seed-based core collection for *Malus sieversii.* J Amer Soc Hort Sci 130, 203-210.

WAGNER I., SCHMITT H.P., MAURER W., TABEL U., 2004. Isozyme polymorphism and genetic structure of *Malus sylvestris* (L.) Mill. native in western areas of Germany with respect to *Malus ×domestica* Borkh. Acta Hort 663, 545-550.

YONEZAWA K., NOMURA T., MORISHIMA H., 1995. Sampling strategies for use in stratified germplasm collections. In: Core collections of plant genetic resources (Hodgkin T, Brown A.H.D., Hintum van Th.J.L., Morales E.A.V., eds). John Wiley Sons, NY, USA. pp. 35-53.