International Journal of Horticultural Science 2007, 13 (3): 37–42
Agroinform Publishing House, Budapest, Printed in Hungary
ISSN 1585-0404

Molecular identification of old Hungarian apple varieties

Wichmann, B.[1], Galli, Zs.[1], Molnár, S.[1], Galbács, Zs.[1], Kiss, E.[1], Szabó, T.[2] & Heszky, L.[1]

[1]St. István University (SIU), Institute of Genetics and Biotechnology, Gödöllő H-2103, Hungary
[2]Research and Extension Centre for Fruit Growing, Újfehértó H-4244, Hungary

Summary: Altogether 40, mainly old Hungarian apple varieties were screened with six previously described microsatellite markers. A total of 71 polymorphic alleles were detected (average 11.8 alleles/locus) and the heterozygosity of markers averaged very high (0.8). The genetic variability among the genotypes proved to be so remarkable that as few as three markers from the applied six were enough to distinguish between the 40 varieties. This was also confirmed by the cumulative probability of obtaining identical allele patterns for two randomly chosen apple genotypes for all loci, which value was quite low: 2.53×10^{-5}. The molecular identification of these genetically very different old apple genotypes could be very useful in future breeding programs.

Key words: microsatellite fingerprinting, genetic variability, optimal primer combination

Introduction

The diversity of agriculture has significantly decreased, mainly as a result of globalization. As a result of evolution and of the approximately ten thousand years of selection work completed by mankind, only 15 plant and 8 animal species were chosen to produce up to 90% of our current food products. This reduction can be felt in all areas of agricultural, and, accordingly, in pomology, as well. Although there are presumably more than a thousand apple cultivars in Europe, including regional old varieties, only a few international cultivars are currently used in apple production. The main dominant cultivar is Golden Delicious, followed by Red Delicious, Granny Smith, and Gala. The share of other cultivars is less than 30%.

The main purposes of apple breeding are to produce fruits of size, colour, and flavour that are desirable and lead to customer satisfaction, and besides, to produce resistant/tolerant apple cultivars against powdery mildew (causative agent: *Podosphaera leucotricha*), apple scab (causative agent: *Venturia inaequalis*), and against the bacteria *Erwinia amylovora,* which causes fireblight (*Crosby* et al., 1992). The necessary genetic background for these purposes can be found in wild apple species and in regional old varieties. Crossing with wild cultivars and following selection steps are time consuming processes, because the removal of unfavourable features makes several back-cross steps necessary (*Fischer,* 1993; *Fischer & Fischer,* 1996). Some old varieties that grow only in private gardens have excellent fruit quality and disease tolerance (*Holb*, 2000). These have adapted extremely well to the different circumstances of soil and climate of the given region. Using these varieties in apple breeding will hopefully come to the front in the future.

The large number of regional old varieties makes the use of reliable molecular markers for characterization, identification, and distinction necessary. For identification of apple cultivars, RAPD (*Mulcahy* et al. 1993) and ISSR (*Goulão & Oliveira,* 2001) techniques and primers planned for the long terminal repeat region of retrotransposons (*Antonius-Klemola* et al. 2006) have been already used. Even so, microsatellites (SSR's: simple sequence repeats) seem to be the best markers for this purpose since they show high polymorphism of repetitive elements among genotypes and have co-dominant inheritance. At present, almost 200 SSR markers are available in apple (*Guilford* et al. 1997, *Gianfranceschie* et al. 1998, *Hokanson* et al. 1998, *Liebhard* et al. 2002). SSR markers have been proven to be highly informative and useful for distinguishing genotypes and for determining genetic relationships among *Malus* cultivars and species (*Goulão & Oliveira* 2001, *Hokanson* et al. 2001, *Liebhard* et al. 2002, *Laurens* et al. 2004). Developed SSR profiles can be very useful in practice, e.g. for selection of distant parents to obtain higher genetic variation in progenies, to identify outcrosses or self-pollinated individuals which do not belong to the progeny of the applied cross, for nursery control and protection of breeders' right, etc.

The main purpose of this study was to distinguish between the old and a couple of new apple varieties using microsatellite markers. In the first step, samples were collected from 40 genotypes at the Research and Extension Centre for Fruit Growing, Újfehértó. Our goal was to find the most simple and inexpensive – using the fewest markers – technique for molecular distinction of all genotypes. Six previously described SSR markers (*Liebhard* et al. 2002) were applied, which were already used in our previous work for identification of 66 commercial apple cultivars (*Galli* et al. 2005).

Materials and Methods

Genomic DNA was extracted using a DNeasy® Plant Mini Kit (Qiagen) and PCR amplifications were carried out

Table 1 Microsatellite primer sequences, number, and size range of alleles and Polymorhism Information Content (PIC) values found by *Liebhard* et al. (2002) and in the present study. The confusion probabilities (C) are also shown, which were used to find the optimal combination of SSR markers for identification.

SSR	*forward primer sequence* *reverse primer sequence* *5'→3'*	Allele size	Allele number	PIC	Allele size	Allele number	PIC	C
		Liebhard et al. 2002			Own result			
CH03g07	*aat aag cat tca aag caa tcc g* *ttt ttc caa atc gag ttt cgt t*	119–181	5	0.77	119–179	8	0.75	0.233
CH04e03	*ttg aag atg ttt ggc tgt gc* *tgc atg tct gtc tcc tcc at*	179–222	11	0.88	161–222	13	0.81	0.166
CH04g10	*caa aga tgt ggt gtg aag agg a* *gga ggc aaa aag agt gaa cct*	127–168	5	0.83	123–168	11	0.79	0.192
CH05c02	*tta aac tgt cac caa atc cac a* *gcg aag ctt tag aga gac atc c*	168–200	4	0.60	160–200	12	0.77	0.214
CH05d11	*cac aac ctg ata tcc ggg ac* *gag aag gtc gta cat tcc tca a*	171–211	5	0.69	169–227	12	0.86	0.113
CH05e03	*cga ata ttt tca ctc tga ctg gg* *caa gtt gtt gta ctg ctc cga c*	158–190	10	0.87	149–193	15	0.84	0.128

by a Perkin Elmer 9700 thermocycler using Cy-5 labelled SSR primers. The six applied pairs of oligonucleotide primers (CH03g07, CH04e03, CH04g10, CH05c02, CH05d11, CH05e03) have already been identified and published (Liebhard et al. 2002). The reactions were performed in a final volume of 20 µl, the reaction mixture consisting of the following components: 50 ng template DNA, 1 × PCR buffer, 0.9 mM $MgCl_2$, 0.3 µM of both primers, 0.2 mM each of dNTPs and 1.2 U of *Taq* DNA polymerase (WestTeam). The cycling profile consisted of an initial denaturation step of 2 min followed by 35 cycles of 20 seconds at 94 °C, 30 seconds at 59 °C, and 60 seconds at 72 °C. The amplification process was finished with 5 min at 72 °C. Products were preliminarily tested by running on 1.2% agarose gel, necessarily diluted with TE1 buffer (10 mM Tris, 1 mM EDTA) up to 30 times and microsatellite alleles were visualized and determined in ALF*express*-II DNA analyzer (Amersham BioSciences), following standard protocols. PCR amplifications using the same extracted DNA as template and allele size determination were repeated at least twice.

Based on the frequencies of observed microsatellite alleles, Polymorphic Information Content (PIC) was calculated by the formula: PIC=1-Σp_i^2 where 'p_i' is the frequency of the *i*th allele of each SSR marker (Anderson et al. 1993). It is essentially the same as the index of heterozygosity. To find the optimal combination and minimum set of markers for identification, the 40 different varieties (N=40) were analyzed according to *Tessier* et al. (1999). At first, the confusion probability (C) was calculated, i.e. the probability that two randomly chosen individuals from the sample of 40 genotypes have identical banding patterns. The confusion probability is for the *i*th pattern of the given *j*th primer present at frequency p_i in this set of varieties is: $c_i = p_i \frac{(Np_i - 1)}{N - 1}$. For the *j*th primer, the confusion probability C_j is equal to the sum of the different c_i for all I patterns generated by the primer: $C_j = \sum_{i=1}^{I} c_i$. The probability of identity (PI) of obtaining the same SSR profile for two randomly chosen apple cultivars was calculated by the $PI = \prod_{j=1}^{J} C_j$ formula. The total number of non-differentiated pairs of cultivars for the *j*th primer is given by $xj = \frac{N(N-1)}{2} Cj$. For a given combination of k primers, this number is: $X_k = \frac{N(N-1)}{2} \prod_{j=1}^{k} C_j$ (*Tessier* et al. 1999).

For the cluster analyses, each detected allele was scored as present (1) or absent (0) in all varieties. The unweighted pair group method using arithmetic means (UPGMA) was used and a dendrogram was constructed based on Jaccard's similarity coefficient (*Jaccard*, 1908) by the statistic software package SPSS 11.0 for Windows (SPSS Inc., USA).

Results

Every SSR primer-pair generated reliable microsatellite alleles in all 40 apple varieties and proved to be suitable to distinguish between all of them. A total of 71 polymorphic alleles were amplified and the number of alleles per marker ranged from

Table 2 Allele sizes (bp) and their frequencies (%) of 6 SSR loci distributed on 40 tested apple varieties

CH03g07		CH04e03		CH04g10		CH05c02		CH05d11		CH05e03	
Size	**%**	**Size**	**%**	**Size**	**%**	**Size**	**%**	**Size**	**%**	**Size**	**%**
119	31.5	**161**	1.4	**123**	9.0	**160**	11.7	**169**	10.3	**149**	1.4
121	1.4	**176**	1.4	**125**	1.5	**164**	1.3	**171**	16.7	**155**	1.4
123	9.6	**184**	6.8	**127**	7.5	**168**	40.3	**173**	23.1	**157**	1.4
127	11.0	**186**	4.1	**135**	38.8	**170**	11.7	**175**	14.1	**161**	5.4
129	35.6	**192**	1.4	**137**	17.9	**172**	19.5	**177**	6.4	**163**	27.0
153	4.1	**196**	37.8	**139**	1.5	**174**	1.3	**181**	3.8	**165**	2.7
165	5.5	**198**	10.8	**141**	3.0	**176**	6.5	**187**	11.5	**167**	4.1
179	1.4	**200**	2.7	**143**	9.0	**180**	1.3	**195**	3.8	**171**	2.7
		202	5.4	**147**	1.5	**190**	1.3	**203**	1.3	**173**	20.3
		204	9.5	**155**	3.0	**196**	1.3	**205**	5.1	**175**	1.4
		208	6.8	**168**	7.5	**198**	1.3	**225**	2.6	**179**	4.1
		210	4.1			**200**	2.6	**227**	1.3	**183**	1.4
		222	8.1							**185**	5.4
										191	18.9
										193	2.7

8 (CH03g07) to 15 (CH05e03), with an average of 11.8 (*Table 1*). The repeated PCR reactions show the same results.

Some new allelic forms were detected which were not identified in the previous work (*Liebhard* et al. 2002). This probably could be explained by the use of only 8 cultivars there, while we used many more. This is the explanation for the difference in PIC values, as well.

It is worth comparing these results with our previous one (*Galli* et al. 2005), where 66 commercial apple cultivars were screened with the same markers. In the previous study, the PIC values were lower, especially in the case of the CH04g10 and CH05d11 markers, where the rising of the value is 30%. This proves that variability is quite large between the land varieties, i.e. those markers which were determined as not appropriate for differentiation in commercial apple cultivars work excellently in this study.

A higher number of alleles usually produce higher PIC values. However this statement is not always true, as the distribution of allele frequencies is also crucial (*Table 2*). The CH05d11 marker proved to be best with its 12 alleles for distinction, better than the CH05e03 marker with 15 alleles. Each SSR marker had at least one or two characteristic allele(s) with high occurrence. These alleles can possibly be considered as base values for SSR repeats, and new alleles might derive from the decrease or increase of the number of repeats during DNA replication.

The average amount of heterozygotes is 76% with a range between 41% and 83%. These are similar to the previously described values of *Gianfranceschi* et al. 1998, *Hokanson* et al. 1998, *Liebhard* et al. 2002, and *Laurens* et al. 2004. CHO4g10 produced the most homozygote alleles (47%), but interestingly its discriminative capability is still great.

Table 3 shows the allele distribution of all apple varieties at six microsatellite loci. In the cases where only one band was visible, its size is reported twice since the genotype is presumed to be homozygous. To determine whether a real homozygous allele or a null allele is involved, access to a segregating population is needed. (The intensity of the specific band usually helps in practice. The peak of the real homozygous allele is much higher.)

Table 3 Allelic composition of 40 apple varieties analyzed by 6 SSR markers. Numbers are in base-pairs.

	CH03g07	CH04e03	CH04g10	CH05c02	CH05d11	CH05e03
Alant alma	119,129	196,204	125,135	172,200	175,177	165,165
Amália	119,129	184,200	123,143,155	160,168	171,177	163,173
Asztrahán fehér	129,165	196,198	135,141	170,176	173,175	163,173
Asztrahán piros	129,129	198,204	135,135	170,176	173,175	191,191
Burgundi	123,129	186,196	135,135	168,168	171,173	165,191
Bogovits alma	123,127	184,202	135,135	168,196	171,171	170,170
Dániel-féle renet	119,119	196,204	135,168	168,172	169,175	163,163
Chieftain	129,153	196,196	135,168	168,200	171,173	163,191
Nyári csíkos fűszeres	119,129	161,161	135,147	160,170	169,187,195	163,191
Éva	119,129	196,204	135,135	168,176	173,173	185,191
Fekete tányér alma	119,129	196,202	135,135	168,172	195,225	173,191
Fertődi téli	119,123,129	184,196	135,135	168,168	173,173	163,168
Húsvéti rozmaring	119,129	192,196,202	135,135	168,172	173,187,205	191,191
Gravensteini	123,129	196,208	135,137	168,172,176	169,187	163,193
Jakab alma	119,119	186,208	135,135	168,170	173,205	163,173
Kenézi piros	123,129	196,204	137,141	168,172	173,186	163,173
Keszthelyi kúpos alma	119,119	196,208	135,135	168,172	173,175	173,191
Középfajta renet	119,129	186,196	123,143,155	160,168,172	173,177	155,167,183
Kubany	123,127	184,198	127,135	168,172	171,171	173,185
Liptói-féle rozmaring	119,127	196,198	127,137	170,176	177,181	163,167,179
Mutsu II.	119,127	198,222	127,135	168,174	171,173	163,179,185
Nyári fontos	119,129	196,222	137,137	160,168,172	171,181,187	161,173
Nyári sóvári	119,129	176,196	127,139	168,172	169,225	163,173
Őszi borizű	127,127	196,196	135,135	160,168	169,195	157,191
Őszi piros renet	129,129	196,208	123,137,143	160,170	173,173	163,163
Puha sóvári	119,129	196,210	137,137	170,170	169,205	163,173
Rózsa alma	119,129	210,222	135,168	168,168	173,186	173,191
Sárga édes	121,127	202,222	137,137	168,172	171,187	149,173
Sikulai	129,165	196,210	137,137	168,190	175,177	163,193
Simonffy piros	119,129	196,200	123,135,143	170,170	173,227	160,173
Széchenyi renet	129,153,165	196,222	123,137,143	168,198	171,175	171,179
Téli arany parmen	129,153	196,196	135,135	168,170	169,175	163,163
Téli banán	119,127	184,196	127,137	164,168	169,205	163,185
Téli fehér kálvil	119,129	198,208	135,168	168,168	171,173	163,191
Téli fehér tafota	127,127	196,196	135,135	168,172	173,175	163,163
Tordai piros kálvil	119,129	196,204	137,168	168,168	171,175,187	173,191
Újvári őszi alma	119,129	196,222	135,135	160,168,172	171,181,187	160,173
Vajalma	123,179	196,196	137,137	160,168	173,203	161,191
Világ dicsősége	119,119	198, 204	135,135	172,180	171,171	191,191
Vista Bella	165,165	198,198	123,135,143	160,168	175,175	163,173

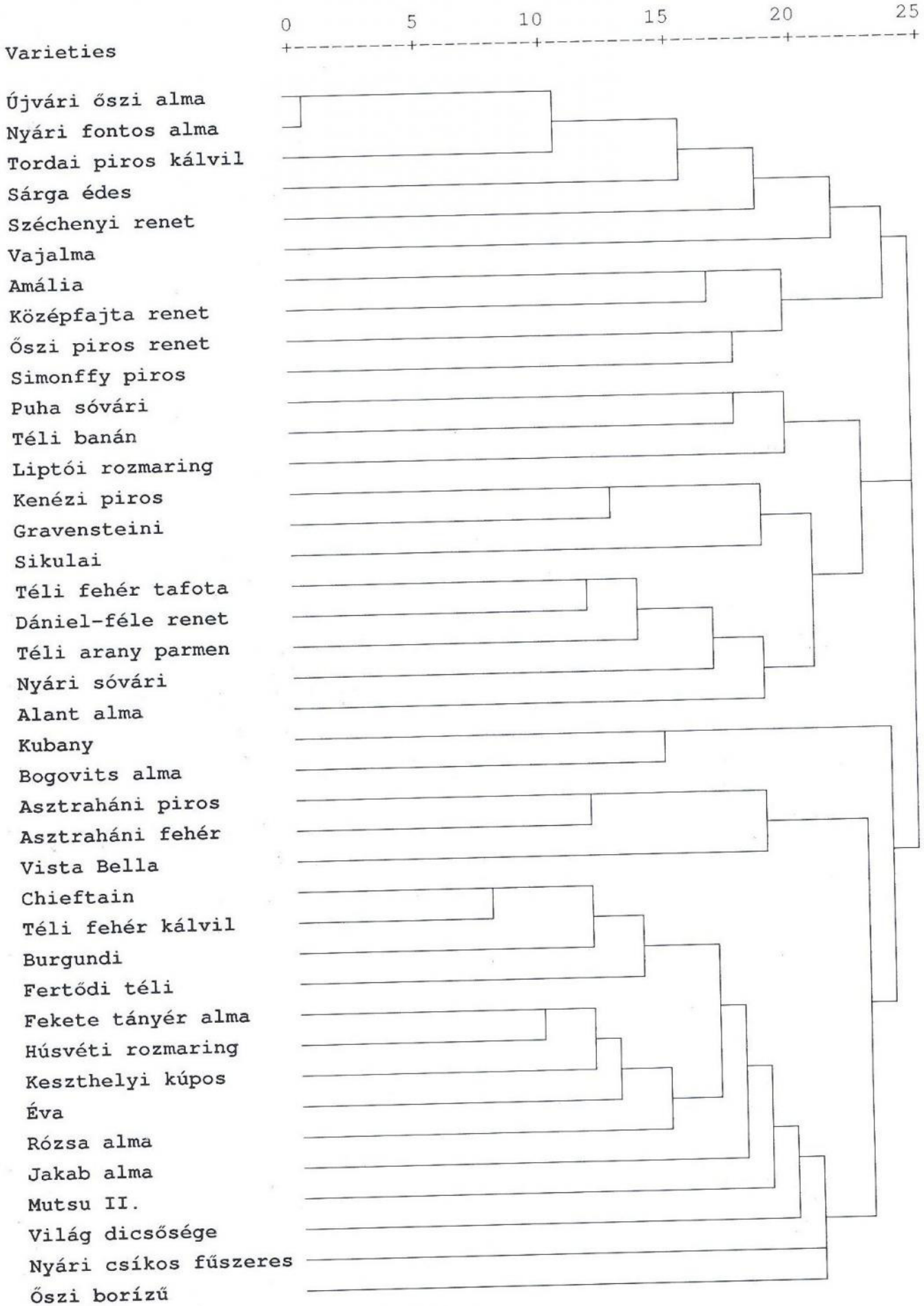

Figure 1 The dendrogram of 40 apple varieties differentiated by 6 SSR markers.

Three distinguishable alleles were determined with at least one primer-pair in case of 15 varieties, indicating triploidy. Presumably, the third bands are not amplification artefacts since these patterns were visible in all replications. Unfortunately, the pedigree of these varieties is unknown, but it is quite possible that triploids were selected a number of times in the past because of their better fruit characteristics. (Triploids rarely arose spontaneously under natural circumstances, when a haploid pollen grain fertilizes a diploid ovule produced by meiotic non-disjunction.)

Successful differentiation of all apple genotypes was accomplished by using only three (CH05d11, CH05e03, CH04e03,) of the SSR markers (*Table 4*). Using the two most polymorphic SSR markers (CH05d11, CH05e03) only 2 genotype-pairs ('Dániel-féle renet' – 'Téli arany parmen' and 'Chieftain' – 'Téli fehér kálvil') were indistinguishable, but it was achieved by the third (CH04e03) marker.

Table 4 Selection of the most efficient minimum set of SSR markers for the identification of 40 different varieties. (Comparison of the experimentally observed and theoretical efficient marker combinations)

Marker combination	Numbers of unseparability pairs	
	Real	Calculated
CH05d11	46	88.3
CH05d11 + CH05e03	4	12.4
CH05d11 + CH05e03 + CH04e03	**0**	2.1
CH05d11 + CH05e03 + CH04e03 + CH04g10	**0**	0.4

The probability that two different varieties displayed the same genotype at all investigated loci (Probability of Identity) was calculated to be very low (2.53×10^{-5} which means 1 : 39525). This confirms the high potential of SSR's for differentiation of apple varieties and the huge genetic variation within land varieties, as well. (In our previous study (*Galli* et al. 2005) this value was much lower (1.79×10^{-4}); none the less 66 commercial apple cultivars were distinguished.)

Similarity matrices were calculated from SSR data. The UPGMA-based dendrogram obtained is shown in *Figure 1*. Whilst the number of markers used for molecular identification should be as low as possible for genotype characterization purposes, for genetic relationship studies and similarity detection between accessions the more SSR markers are used the better. For this reason the dendrogram relating to all the 40 apple varieties is based on data of all 6 microsatellite loci, but these data are still slight enough for a dendrogram with good resolution. However, the genetic similarity between varieties can be seen. The closest relationship (more than 73%) was found between varieties 'Újvári őszi alma' and 'Nyári fontos alma'. There are some genotype-pair ('Világ dicsősége' – 'Őszi piros renet', 'Bogovits alma' – 'Puha sóvári', 'Asztrahán piros' – 'Sárga édes'), where absolutely no similarity could be detected. The 'Asztrahán piros' and 'Asztrahán fehér' varieties – which could be thought of as sport mutants – show only 50% similarities, accordingly the derivation by somatic mutation is impossible etc.

The detected high genetic variation between land varieties also demonstrates the usefulness of germplasm collections. Developed SSR profile of these genotypes is essential for efficient breeding methods to create new apple cultivars with increased disease resistance, better fruit quality, or other traits.

Acknowledgements

We would like to thank Mr. Dávid Balázs M.Sc. student for his technical assistance. Support for this work was provided by the Hungarian Scientific Research Fund (OTKA PF63582).

References

Anderson, J. A., Churchill, G. A., Autrique, J. E., Tanksley, S. D. & Sorrels, M.E. (1993): Optimizing parental selection for genetic linkage maps. Genome 36:181–186.

Antonius-Klemola, K., Kalendar, R. & Schulman, A. H. (2006): TRIM retrotransposons occur in apple and are polymorphic between varieties but not sports. Theor. Appl. Genet. 112:999–1008.

Crosby, J. A., Janik, J., Pecknold, P. C., Korban, S. S., O'Connor, P. A., Ries, S. M., Goffreda, J. & Voordeckers, A. (1992): Breeding apples for scab resistance: 1945–1990. Fruit Var. J. 46: 145–166.

Fischer, C. Fischer, M. (1996): Ergebnisse der Apfelzüchtung in Dresden-Pillnitz. Gartenbauwissenschaft. 61: 139–146.

Fischer, C. (1993): Results from the resistance breeding programme for apple in Dresden-Pillnitz. Acta Hort. 347: 163–168.

Galli, Z., Halász, G., Kiss, E., Heszky, L. & Dobránszki, J. (2005): Molecular identification of commercial apple cultivars with microsatellite markers. HortScience 40(7):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.

Goulão, L. & Oliveira, C. M. (2001): Molecular characterisation of cultivars of apple (*Malus* × *domestica* Borkh.) using microsatellite (SSR and ISSR) markers. Euphytica 122:81–89.

Guilford, P., Prakash, S., Zhu, J. M., Rikkerink, E., Gardiner, S., Bassett, H. & Forster, R. (1997): Microsatellites in *Malus* × *domestica* (apple): abundance, polymorphism, and cultivar identification. Theor. Appl. Genet. 94:249–254.

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.

Hokanson, S.C., Lamboy, W. F., Szewc-McFadden, A. K. & McFerson, J. R. (2001): Microsatellite (SSR) variation in a collection of *Malus* (apple) species and hybrids. Euphytica 118:281–294.

Holb I. (2000): Az alma ventúriás varasodásának mértéke integrált és ökológikus védekezési programokban. Kertgazdaság (Horticulture) 32. (2): 25–35.

Jaccard, P. (1908): Nouvelles recherches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44:223–270.

Laurens, F., Durel, C. E. & Lascostes, M. (2004): Molecular characterization of French local apple cultivars using SSRs. Acta Hort. 663:639–642.

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.

Mulcahy, D.L., Cresti, M., Sansavini, S., Douglas, G. C., Linskens, H. F., Bergamini Mulcahy, G., Vignani, R. & Pancaldi, M. (1993): The use of random amplified polymorphic DNAs to fingerprint apple genotypes. Scientia Horticulturae, 54(2):89–96

Tessier, C., J. David, P., This, J., Boursiquot, M. & Charrier, A. (1999): Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. Theor. Appl. Genet. 98:171–177.