HORTSCIENCE 40(1):15–17. 2005.

Using Microsatellite Analysis to Verify Breeding Records: A study of 'Honeycrisp' and Other Cold-hardy Apple Cultivars

Paul R. Cabe[1]
Biology Department, Washington and Lee University, Lexington, VA 24450

Andrew Baumgarten and Kyle Onan
Biology Department, St. Olaf College, Northfield MN 55057

James J. Luby
Department of Horticulture Science, University of Minnesota, 1970 Folwell Avenue, St. Paul, MN 55108

David S. Bedford
Department of Horticulture Science, Horticultural Research Center, 600 Arboretum Boulevard, Excelsior, MN 55331

Additional index words. 'Haralson,' 'Connell Red,' 'Macoun,' 'Honeygold,' DNA, SSR, STR

Abstract. **We used microsatellite loci to investigate the parentage of the apple cultivar 'Honeycrisp', a patented University of Minnesota introduction. In an attempt to find the correct parents, we also examined other apple varieties associated with the University of Minnesota apple breeding program. Based on written records from the 1960s, the presumed parents of 'Honeycrisp' were 'Honeygold' and 'Macoun'. We were able to exclude both of these as parents, but found that 'Keepsake' was consistent as one of the parents. A second potential parent could not be discovered. 'Haralson', another commercially important cultivar from the University of Minnesota, is likely from a cross between 'Malinda' and 'Wealthy'.**

Despite anecdotal evidence, and even breeding records, the origin of many important plant cultivars remains unknown or uncertain. Current genetic techniques now allow specific testing of putative parents, and can be used to identify parents in cases where nothing is known. Microsatellites, a type of DNA sequence also termed simple-sequence repeat (SSR) or short tandem repeat (STR) markers, are ideal for this purpose. For instance, microsatellite analysis has been used to identify the parentage of many wine grape varieties (Bowers et al., 1999). Microsatellite profiles can also be used as fingerprints of varieties that have uses such as avoiding mislabeling of nursery stock. In this study, we used microsatellite loci previously developed for use in apples (*Malus x domestica*) to develop fingerprints that permit unique identity of several cultivars developed in the University of Minnesota apple breeding program since the 1920s.

We began this research to confirm the parents of the University of Minnesota apple 'Honeycrisp' using these genetic techniques. Breeding records for 'Honeycrisp' indicated that this cultivar came from a cross between 'Honeygold' (another University of Minnesota cultivar) and 'Macoun'. The dissimilarity of 'Honeycrisp' to these reported parents based on several fruit characteristics (Tong et al., 1999) caused us to speculate that this recorded parentage might be incorrect.

In the course of this research we were also able to confirm or reject the parentage indicated by breeding records for several varieties and to identify putative parents in some cases where parentage was unknown or suspect, notably for 'Haralson,' one of the other most important introductions from the University of Minnesota breeding program. Breeding records indicated that 'Malinda' was the female parent of 'Haralson' but the male parent was unknown as the seed resulted from open pollination.

Materials and Methods

Cultivars examined and DNA Extraction. We examined 'Honeycrisp', its presumed parents 'Honeygold' and 'Macoun', and a number of other cultivars and breeding selections from the University of Minnesota breeding program. DNA was extracted from fresh leaf tissue following protocols in Doyle and Doyle (1990). Leaves were obtained from the University of Minnesota Horticultural Research Center in Excelsior, Minnesota or the National Germplasm Repository in Geneva, N.Y. For each extraction, one gram of tissue was ground under liquid nitrogen.

Amplification of microsatellite loci. We amplified eleven microsatellite loci (Table 1) using the polymerase chain reaction (PCR). We used microsatellite primers GD12, GD15, GD96, GD142 (Hokanson et al., 1998), 02b1, 05g8, 23g4 (Guilford et al., 1997), CH02B10, CH02C06, CH01G12, and CH01H02 (Gianfranceschi et al., 1998); primers were chosen with preference for loci

Table 1. Primer sequences and dilutions used for each locus in this report.

Locus	Primer sequence	Reference	Fluorescent label	Dilution PCR : H_2O
GD12	F: TTgAggTgTTTCTCCCATTggA			
	R: CTAACgAAgCCgCCATTTCTTT	1	6-FAM	1:50
GD15	F: CgAAAgTgAgCAACgAACTCC			
	R: ACTCCATCATCgggTggTg	1	HEX	1:40
GD96	F: CggCggAAAgCAATCACCT			
	R: gCCAgCCCTCTATggTTCCAgA	1	TET	1:30
GD142	F: ggCACCCAAgCCCCTAA			
	R:ggAACCTACgACAgCAAAgTTACA	1	6-FAM	1:40
02b1	F: CCgTgATgACAAAgTgCA TgA			
	R: ATgAgTTTgATgCCCTTggA	2	6-FAM	1:50
05g8	F: CggCCATCgATTATCTTACTCTT			
	R:ggATCAATGCACTgAAATAAACg	2	HEX	1:50
23g4	F: TTTCTCTCTCTTTCCCAACTC			
	R: AgCCgCCTTgCATTAAATAC	2	HEX	1:5
CH02B10	L: CAAggAAATCATCAAAgATTCAAg			
	R:CAAgTggCTTCggATAgTTg	3	TET	1:32
CH02C06	L: TgACgAAATCCACTACTAATgCA			
	R: gATTgCgCgCTTTTTAACAT	3	6-FAM	1:16
CH01G12	F: CCC ACC AAT CAA AAA TCA CC			
	R: TgA AgT ATg gTg gTg CgT TC	3	TET	1:32
CH01H02	L: AgA gCT TCg AgC TTC gTT Tg			
	R: ATC TTT Tgg TgC TCC CAC AC	3	6-FAM	1:32

[1]Hokanson et al., 1998.
[2]Guilford et al., 1997.
[3]Gianfreschi et al. 1998.

Received for publication 8 Mar. 2004. Accepted for publication 31 May 2004. We thank Stan Hokanson and Amy Szewc-McFadden for technical advice on their microsatellite protocols, Lisa Connors for help with the microsatellite data collection, and the staff of the USDA-ARS National Plant Germplasm Repository for help in obtaining samples. This work was funded in part by St. Olaf College, Washington and Lee University, and the Office of Patents and Technology Marketing, University of Minnesota. Mention of a trademark, proprietary product, or vendor does not constitute a guarantee or warranty of the product by the U.S. Dept. of Agriculture and does not imply its approval to the exclusion of other products or vendors that also may be suitable.
[1]To whom reprint requests should be addressed; e-mail cabep@wlu.edu.

with high heterozygosity and allelic diversity. Forward primers were labeled with either HEX, TET, or FAM (IDT DNA, Coralville, Iowa).

The primers were used in conjunction with standard PCR kits from Promega Corporation (Madison, Wis.). The reactions contained 1.5 mM $MgCl_2$, 15 pmol of each primer, and 0.8 µL of template DNA (as described above), in a total reaction volume of 20 µL. Reactions used a manual hot start approach: the reaction mix was heated to 95 °C for 5 min, and 1 unit (0.2 µL) of *Taq* DNA polymerase was added to each reaction.

In general, amplification conditions were based on published protocols. Primers from Guilford et al. (1997) and Hokanson et al. (1998) used a manual hot start with cycling parameters given below. Primer sets GD12, GD15, and GD96 were amplified with identical cycling parameters: initial denaturation at 96 °C, followed by 35 cycles of 94 °C for 1 min (denaturation), 52 °C for 2 min (annealing), and 72 °C for 2 min (elongation), followed by a final elongation period of 72 °C for 10 min. Primer sets 02b1, 05g8, and 23g4 were initially denatured at 94 °C, followed by 35 cycles of 94 °C for 40 s, 52 °C for 40 s, 72 °C for 20 s, followed by a 10-min extension at 72 °C. Primer set GD142 used a touchdown profile following an initial denaturation at 94 °C: denaturation cycles of 94 °C for 1 min and elongation cycles of 72 °C for 45 s were used with an annealing temperature of 65 °C for the first two cycles, a drop of 0.5 degree per cycle for the next 18 cycles, and the final 5 cycles at 55 °C. Primer sets CH02B10 and CH02C06 used a touchdown protocol (94 °C for 2.5 min, 5 cycles of 94 °C for 30 s, 65 °C for 1 min (decreasing 1 °C per cycle), 72 °C for 1 min, 30 cycles of 94 °C for 30 s, 60 °C for 1 min, 72 °C for 1 min, and 72 °C for 5 min). Primer sets CH01H02 and CH01G12 used different annealing temperatures (61.8 and 55.2 °C respectively) but identical cycling parameters (94 °C for 2 min, 30 cycles of 94 °C for 30 s, annealing temperature for 45 s, 72 °C for 30 s, and a final extension at 72 °C for 5 min).

To assess success of the PCR reaction, 5 µL of the reaction were loaded onto 1.8% agarose gels, along with a 100-bp size ladder. Gels were stained using ethidium bromide, and digitally photographed under UV light. For successful reactions, PCR products of the appropriate size range were clearly visible (data not shown).

Electrophoresis of PCR-generated DNA fragments. Products from successful PCR reactions were diluted (Table 1), and sent to a commercial laboratory (Advanced Center for Genetic Analysis, University of Minnesota, USA) for size analysis using an ABI Prism 377 sequencer.

Data analysis. Fragment sizing was done using the program Genescan, and genotypes determined using Genotyper (Applied Biosystems, Foster City, Calif.). Fragment sizes were rounded to the nearest whole integer, and grouped into allele groups. Each putative or possible parentage assignment was manually inspected; an assignment was considered supported if all loci were consistent with the hypothesis of parentage. We were cognizant of the problems introduced by null alleles, but did not find any evidence of these in the data.

The statistical power to exclude parents is based on the number and frequencies of the alleles and the number of loci examined; typically, loci with many alleles at about equal frequency give the greatest power to exclude parents. We calculated the average exclusion probability over the ten variable loci to exclude a single parent where the other parent is not known using the approach of Jamieson and Taylor (1997). This is the most conservative approach since it is easier to exclude one potential parent if the other is known. In this calculation, we used allele frequency data from the original descriptions of these loci (Gianfranceschi et al., 1998; Guilford et al., 1997; Hokanson et al., 1998).

Results

The 'Golden Delicious' genotype was used as a control and for validation of protocols. Allele sizes for this cultivar were compared to published (Gianfranceschi et al., 1998; Guilford et al., 1997) or publicly available data (Hokanson et al., 1998 data are presented at http://grain.jouy.inra.fr/cgi-bin/webace/webace?db=rosedb) for all loci. In all cases, data from this project were consistent with published values. Genetic variation at each locus was consistent with the original descriptions. Locus GD15 was monomorphic for cultivars examined, and so data from this locus are omitted. Genotypes for 'Honeycrisp' and a number of other varieties are given in Table 2. The average exclusion probability over the ten loci presented is 0.998, meaning only 0.2% of randomly drawn cultivars could not be excluded in any given comparison.

Discussion

For 'Honeycrisp', a brief inspection of genotypes casts immediate doubt on the putative attribution. 'Honeygold' can be excluded as a parent by two loci, and 'Macoun' can be excluded by at least five loci. After obtaining this result, a number of other potential parents were checked; potential parents were identified by searching University of Minnesota breeding archives for other cultivars used in crosses. A number of potential cultivars were excluded, including MN1607, 'Spartan', 'Goodland' (data not shown), and 'Connell Red', all parents of seed progenies that were planted adjacent to the original seedling tree of 'Honeycrisp' according to field maps (data not shown). Eventually, 'Keepsake' was found to be genetically consistent as a parent of 'Honeycrisp,' and this University of Minnesota cultivar would have been available for breeding at that time. Additional information can be gleaned by examining the microsatellite genotype of MN1708, which, according to selection records, was located only a few trees away in the same orchard row and derived from the same original cross which also produced the seedling of 'Honeycrisp.' The genotype of MN 1708 is also consistent with 'Keepsake' parentage. This suggests that the mistake that led to the erroneous attribution of 'Honeycrisp' to the 'Macoun' × 'Honeygold' cross was likely a mistake in handling multiple seeds or plants prior to planting of the seedlings in the orchard. If we assume that both 'Honeycrisp' and MN1708 are offspring of 'Keepsake,' much of the genotype of the other parent can be deduced. Unfortunately, we were not able to locate any cultivar with such a genotype, or any other single cultivar (other than 'Keepsake') which is consistent as a parent to both 'Honeycrisp' and MN 1708.

'Keepsake' was derived, in turn, from a cross between 'Northern Spy' and MN447. For loci available for comparison, this matches perfectly at all but one locus, CH02B10. In this case, an allele from 'Northern Spy' (125) is close in size to an allele present in 'Keepsake' (123). This may be attributable to mutation,

Table 2. Genotypes for some of the sampled cultivars. The last 10 columns are microsatellite loci (see Table 1). Alleles are specified based on average fragment length. Genotypes with only a single allele specified are presumed homozygous.

Cultivar	Hypothesized parentage	GD12	GD96	GD142	02b1	05g8	23g4	CH02B10	CH02C06	CH01G12	CH01H02
'Honeycrisp'	'Honeygold' × 'Macoun'	153	182/184	129/159	226/231	116	89/95	121/123	230/254	147/152	237/245
'Honeygold'	'Golden Delicious' × 'Haralson'	153/192	182/188	135/144	219/231	116/121	84/115	121/125	236/240	138/147	237/251
'Macoun'	'Black Jersey' × 'McIntosh'	153/155	178/182	137/157	234/242	122/124	84/101	125/129	236/254	138/152	249
'Keepsake'	'Northern Spy' × MN 447	153/155	180/184	140/159	226/240	116/122	89/107	123/133	230/236	138/152	245/249
MN1708	'Honeygold' × 'Macoun'	153	180/182	129/140	216/226	122/122	84/89	121/123	236/254	138/147	237/245
'Haralson'	'Malinda' open pollinated	153/157	176/188	129/135	231	116/150	84/115	125/125	236/236	132/138	237/257
'Malinda'	?	153/157	178/188	135/149	219/231	116/143	84/115	123/125	236	132/156	237
'Wealthy'	?	153/159	176/186	129/138	219/231	150	84	121/125	236/254	108/138	249/257
'Golden Delicious'	'Grimes Golden' × ?	153/192	174/182	144/144	219/231	122	84/89	121/125	236/240	106/147	249
'Zestar'	'State Fair' × MN 1691	153/157	176/188	129/135	n/a	116	85/115	125	236	132/138	237/257
'State Fair'		153	153/188	129/142	219/230	124/150	95/114	125/131	236	111/138	236/245
'Grimes Golden'	?	155/192	174/182	144/157	219	122	84/110	121/123	240/244	147/151	236/249
'McIntosh'	?	153	182	135/137	234/234	124/128	84/101	129/146	230/254	132/152	249
'Fireside'	?	153/159	174/184	138/149	219/242	122/150	84	121/133	232/234	108/110	237/249

since one-step mutations are the most common type at microsatellite loci. It is also possible this is a scoring error (see discussion below). For these reasons, we accept this attribution.

In the course of this research, data were generated which allowed the testing of several other hypotheses of cultivar origins. 'Haralson,' a University of Minnesota cultivar introduced in 1922 and the most widely planted in the state for >50 years, is recorded as a cross of open pollinated 'Malinda.' This attribution matches over ten loci. Furthermore, 'Wealthy,' known to have been present in the orchard from which the 'Malinda' seed was collected (Dorsey, 1919), is perfectly consistent as the second parent. 'Chestnut' and MN447 are also recorded as open pollinated Malinda, but in these cases the data do not support these attributions.

'Honeygold' was produced, according to records, from a cross between 'Haralson' and 'Golden Delicious.' This hypothesis was completely supported over all but one of the scored loci, CH01H02. For this locus, 'Honeygold' shares the 237 allele with 'Haralson,' but has a 251 allele which Golden Delicious appears to lack. For dinucleotide microsatellites such as this locus, a well-known scoring problem sometimes arises with alleles separated by only 2 bases; the second allele may be misinterpreted as artifactual stutter band. In fact, Gianfranceschi et al. (1998) report the genotype for this locus in 'Golden Delicious' consistent with a 249/251 genotype. Thus, we interpret this inconsistency as a minor scoring error, and not reason to question this parentage assignment.

The relatively new University of Minnesota introduction 'Zestar!' (Minnewashta cultivar) has putative parents 'State Fair' and MN 1691. Although we did not collect data on MN1691, the genotype data support 'State Fair' as one of the parents. We were also able to confirm that 'Connell Red' is a sport of 'Fireside' (data for 'Connell Red' not shown), that 'Grime's Golden' is consistent as one of the parents of 'Golden Delicious,' and that 'McIntosh' is consistent as a parent of 'Macoun.'

It is worth noting that when most of these crosses were made, simple and effective genetic tools were not available to monitor breeding programs. With the advent of DNA markers, breeders can now easily confirm parentage at any step in the process. The most efficient time to do this might be when a seedling is first selected and asexually propagated, about five to eight years after the initial cross. Since only about 1% of seedlings are selected, earlier testing would be a waste of resources. Waiting much later in the testing phase (which may take another decade or more) might mean that in the case of a mistake such as 'Honeycrisp,' the original parents may have been lost or discarded.

The information gained from microsatellite markers in this study provides support for critical revisions of the parentage of cultivars with regional and international importance, namely 'Haralson' and 'Honeycrisp.' This information is not only of historic interest, but the putative pedigrees can also assist breeders in planning future crosses.

Literature Cited

Bowers, J., J. Boursiquot, P. This, K. Chu, H. Johansson, and C. Meredith. 1999. Historical genetics: the Parentage of chardonnay, gamay, and other wine grapes of northeastern France. Science 285:1562–1565.

Dorsey, M.J. 1919. Some characteristics of open-pollinated seedlings of the Malinda apple. Proc. Amer. Soc. Hort. Sci. 16:36–42.

Doyle J. and J. Doyle. 1990. Isolation of DNA from fresh tissue. Focus 12:13–15.

Gianfranceschi, L., N. Seglias, R. Tarchini, M. Komjanc, and C. Gessler. 1998. Simple sequence repeats for the genetic analysis of apple. Theor. Appl. Genet. 96:1069–1076.

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

Hokanson, S.C., A.K. Szewc-McFadden, W.F. Lamboy, and J.R. McFerson. 1998. Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus* ×*domestica* Borkh. core subset collection. Theor. Appl. Genet. 97(5–6):671–683.

Jamieson, A. and S. S. Taylor. 1997. Comparisons of three probability formulae for parentage exclusion. Animal Genet. 28:397–400.

Tong, C., D. Krueger, Z. Vickers, D. Bedford, J. Luby, A. El-Shiekh, K. Shackel, and H. Ahmadi. 1999. Comparison of softening-related changes during storage of the 'Honeycrisp' apple, its parents, and 'Delicious'. J. Amer. Soc. Hort. Sci. 124:407–415.