Plant Breeding **123**, 554—557 (2004)

ISSN 0179-9541

Efficient low-cost DNA extraction and multiplex fluorescent PCR method for marker-assisted selection in breeding

J. E. Frey, B. Frey, C. Sauer and M. Kellerhals

Agroscope FAW Waedenswil, Department of Crop Protection, Molecular Diagnostics Laboratory, Swiss Federal Research Station for Horticulture, CH-8820 Waedenswil, Switzerland; E-mail: juerg.frey@faw.admin.ch

With 1 figure and 3 tables

Received January 15, 2004/Accepted May 21, 2004
Communicated by W. E. Weber

Abstract

Marker-assisted selection (MAS) is an increasingly important tool in current breeding efforts for improved crop plants and animal breeds. It enables detection of favourable alleles in early developmental stages and thus may result in substantial cost savings. Until now, however, the high costs of the required chemicals and materials, together with the still very labour-intensive methods, have been an obstacle to widespread application of MAS. A new multiplex-polymerase chain reaction (PCR)-based method has been developed for reliable low-cost, high-throughput screening. By its use 3366 apple seedlings were screened with an average hands-on time from DNA extraction to data ready for analysis of <4 h per 96 plants, and at a cost below US$ 0.5 per marker per plant. Factors that have a strong effect on segregation ratios such as elevated levels of outcrossing are easily detected, as a significant correlation was observed between deviation from expected segregation ratios in some affected markers and the level of outcrossing in a cross. The new method is suitable for many crop species and, provided that suitable buffers are used for DNA extraction, for animals too.

Key words: *Malus* × *domestica* — *Podosphaera leucotricha* — *Venturia inaequalis* — fluorescent primers — marker-assisted selection — multiplex-polymerase chain reaction

Marker-assisted selection (MAS) enables the detection of favourable alleles in the early stages of development and therefore allows a significant reduction of the breeding population (Liebhard et al. 2003). MAS thus has the potential to greatly increase the efficiency of crop plant and animal breeding (Hospital et al. 1997, Ribaut and Hoisington 1998, Dekkers and Hospital 2002). This is especially true if large populations can be screened for several molecular markers (Hospital et al. 1997, Davis and DeNise 1998, Young 1999). However, to date, economic constraints of the prohibitive costs of the laboratory work involved in large-scale molecular screening are a major obstacle to the widespread incorporation of MAS in breeding programmes (Moreau et al. 1998, Dekkers and Hospital 2002, Dreher et al. 2003, Koebner and Summers 2003). It is therefore essential to develop MAS methods that deliver a maximum of information with a minimum of input in time and money. From a technical perspective this requires development of efficient DNA extraction procedures and robust multiplex-polymerase chain reaction (PCR) protocols, which are currently not available, and the use of sophisticated software for data analysis.

In crop breeding, the greatest benefits from MAS will be in perennial crops because many important traits are expressed only after years of costly field maintenance (Liebhard et al. 2003). For example, MAS is currently an important tool in apple breeding (Kellerhals et al. 2000). One of the primary aims of apple breeding programmes is the development of disease-resistant varieties. Two of the major diseases of apple are scab (*Venturia inaequalis*) and powdery mildew (*Podosphaera leucotricha*). A number of molecular markers have been developed for major genes coding for resistance to these two pathogens (e.g. Urbanietz et al. 1999). A selection of these markers was used routinely to pyramid (i.e. accumulate; Dekkers and Hospital 2002) apple scab and powdery mildew genes in individual progeny of apple crosses.

Here, a method is presented that facilitates the screening of large plant populations of several thousand individuals with reasonably low effort and at economically tolerable costs. As required for MAS, this method enables the detection of young individuals among the progeny of crosses that have inherited all the desirable traits from the parents. It also enables the detection of effects that may affect allele frequencies, such as outcrossing, selfing and/or low viability alleles. This new method dramatically decreases the amount of time and effort required when using classical breeding methods.

Materials and Methods

Plant materials: Young apple leaves (2–5 cm length) were collected from greenhouse grown seedlings at the 3–4-leaf stage. Leaves were placed in small (90 × 135 mm) paper bags with a 15 mm flap (order number 14.346.04; Mueller und Krempel AG, Buelach, Switzerland), labelled and stored at −20°C. Each leaf sample consisted of a 4.7 mm diameter disc taken using the collar of the top of a 0.2 ml flat cap PCR tube (Molecular BioProducts, Bioconcept, Allschwil, Switzerland) to puncture the leaf when placed flat on the flap of the envelope in which it was stored. This sampling technique helped to avoid possible sources of cross-contamination. Sample discs were stored in PCR tubes at −20°C.

DNA extraction and multiplex-PCR: The DNA was extracted from the leaf discs using the extraction and dilution buffers of the Extract-N-Amp™ Plant PCR Kit [Fluka, Buchs, Switzerland; Sigma order no. XNAP2; buffer order nos E7526 (extraction solution) and D5688 (dilution solution)] according to the manufacturer's recommendations. The extract was then diluted by a factor of 10 and 2 µl of this crude extract was used in a 10 µl PCR reaction. For the PCR reaction, the Qiagen Multiplex Kit (Qiagen, Basel, Switzerland) with fluorescently labelled primers was used (Table 1). For a full microtitre plate of 96 samples, 928 µl of master mix was prepared consisting of 580 µl 2X

Table 1: Microsatellite (MS) and sequence characterized amplified region (SCAR) markers used in this study

Marker	Gene	Forward primer (5′–3′)	Reverse primer (5′–3′)	Marker type/number	Expected linked allele size	Observed linked allele size	Reference
AT20-450	*PL1*	f-atcagccccacatgaatctcatacc	acatcagccctcaaagatgagaagt	SCAR	450	447	Markussen et al. (1995)
AL07	*Vf*	h-tccttactgaggaggaaaccag	caagggaactgatctttcgttg	SCAR	450	464	Tartarini et al. (1999)
CHVf1	*Vf*	atcaccaccagcagcaaag	h-catacaaatcaaagcacaaccc	MS 1	159	157	C. Gessler, personal communication
CH04H02	*PL2*	ggaagctgcatgatgagacc	h-ctcaaggatttcatgcccac	MS 2	185	np (181)	A. Patocchi, personal communication
CH05e03	*Vbj*	n-cgaatattttcactctgactggg	(g)caagttgttgtactgctccga	MS 3	150	154	M. Gygax, personal communication
CH02C02a	–	f-cttcaagttcagcatcaagacaa	(g)tagggcacacttgctggtc	MS 4	176	177	A. Patocchi, personal communication
CH03C02	–	n-tcactatttacgggatcaagca	gtgcagagtctttgacaaggc	MS 5	128	130[1]	Liebhard et al. (2002)
CH02B10	*Vr*	f-caaggaaatcatcaaagattcaag	caagtggcttcggatagt	MS 6	122	np (118)	Hemmat et al. (2002)

[1] Tentative association.
Fluorescent labels are at the 5′-end of one primer and are indicated by the code for the label (i.e. colour), followed by a dash (f = 6FAM; h = HEX; n = NED). Bases in parentheses were appended to better suit the attached labels as suggested by the manufacturer (Applied Biosystems). Linkage to a resistance-conferring gene has so far not been established for the markers CH02C02a and CH03C02. These markers were included to increase the degree of multiplexing, the published and observed fragments are as far as is known not disease linked. np = resistance linked allele not present in the crosses assessed (the fragment size indicated in parenthesis was observed in plants not used in the present study).

Qiagen Multiplex PCR Master Mix [Qiagen Multiplex PCR Buffer, pH 8.7, containing dNTPs, Qiagen HotStar *Taq* DNA Polymerase, and 6 mM $MgCl_2$ (for a final concentration of 3 mM)], 116 µl Q-Solution (5X concentrated proprietary Qiagen PCR additive), 116 µl of a primer mix with 2 µM of each primer (for a 0.2 µM final concentration of each primer) and 116 µl of highly pure water obtained from a Milli-Q Synthesis A10 (Millipore, Molsheim, France). The master mix was then placed in a Multiprobe II robotic liquid handling system (Perkin-Elmer, Rodgau, Germany) for single-tube pipetting, i.e. distribution of the master-mix and dilution and distribution of the DNA as mentioned above. PCR was performed on Techne Genius thermocyclers (Witec AG, Littau, Switzerland) using the recommended multiplex-PCR protocol (Qiagen) with the following cycling conditions: 2 min at 40°C (to preheat the cover) and 15 min at 95°C for initial denaturation; then 40 cycles of 40 s at 94°C, 90 s at 57°C and 90 s at 72°C, followed by 30 min at 60°C and a final hold at 10°C.

Electrophoresis: After PCR amplification, 40 µl of water was added to the amplification product and 0.5 µl was transferred to 15 µl of formamide containing the fluorescent GeneScan™-500 ROX™ standard (Applied Biosystems, Rotkreuz, Switzerland) using 8-channel multipipettors, followed by 2 min at 96°C for heat denaturation and rapid cooling in the freezer. The plates were then transferred to a 3100 Genetic Analyzer (Applied Biosystems) and run on a high-resolution polymer (POP-6; Applied Biosystems) with the following modifications to the factory default settings for genotyping (GeneScan36_POP4 DefaultModule parameters): run temperature of 60°C, injection time of 10 s, and a run time of 6500 s. The run time was decreased to 3500 s if only short fragments were analysed.

Data analysis: After the runs, the files produced by the Genetic Analyzer were ready for the software GENEMAPPER™ Version 3.0 (Applied Biosystems) and analysed using the default values. Selfers, i.e. plants that originated from self-pollination, were identified by the presence of both alleles from the mother plant and outcrossers, i.e. plants that originated from pollination by an unknown male father (not the one whose pollen was used in the artificial pollination), by the presence of alleles not present in either parent.

Table 2: Number of F_1 progeny tested (N) for each cross, percentage of failed analyses, self-pollinations and outcrossings among all crosses

		Percentage failed analysis			
Cross	F_1 (N)	Percentage presence/absence analysis impossible	Percentage genotyping impossible	Percentage selfers	Percentage outcross
0201	409	0.00	0.24	0.73	0.24
0202	171	1.17	1.17	2.92	1.75
0203	91	3.30	3.30	5.49	30.77
0204	375	0.53	0.53	0.80	6.93
0101	88	0.00	1.14	0.00	0.00
0102	206	0.49	0.49	0.49	0.97
0103	187	0.53	1.60	0.00	6.95
0104	58	0.00	5.17	0.00	32.76
0105	146	0.68	3.42	2.74	10.27
0117	800	0.38	0.38	0.25	0.25
0118	800	0.00	0.00	0.25	3.38
0119	35	0.00	0.00	2.86	11.43
Total	3366	AVG = 0.39	AVG = 0.71	AVG = 0.77	AVG = 4.16

AVG, averages; 'Percentage presence/absence analysis impossible' indicates the percentage of multiplex polymerase chain reactions (PCR) that could not be analysed due to PCR failure; 'Percentage genotyping impossible' indicates the percentage of multiplex analyses in which amplification failure of individual alleles prevented genotyping.

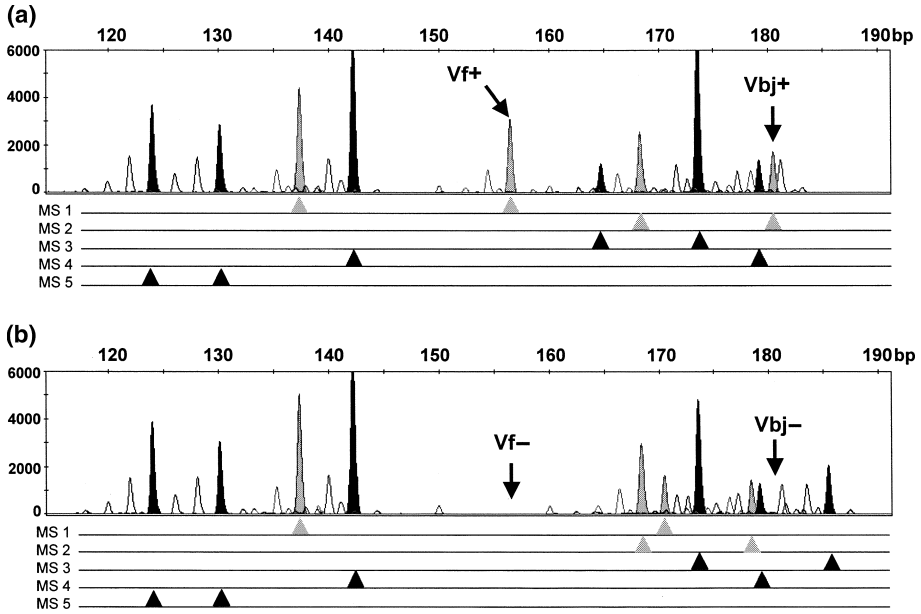

Fig. 1: Two pentaplex-polymerase chain reactions (PCR) of two F_1 genotypes of cross 0204. The individual displayed in panel (a) shows both alleles indicating the presence of the *Vf* and *Vbj* genes (*Vf*+, *Vbj*+), whereas the individual in panel (b) lacks both alleles (*Vf*–, *vbj*–). MS1–5 = microsatellites 1–5 (see Table 1). Triangles indicate alleles for each microsatellite. The *X*-axis indicates fragment length in base pairs, and the *Y*-axis indicates arbitrary fluorescence values

Results

In total, 3366 plants were screened with up to eight different molecular markers in a single multiplex-PCR reaction for each plant. Six of the markers used in this MAS project were microsatellites and two were sequence characterized amplified regions (SCARs) (Table 1). The screened plants originated from 12 crosses as listed in Table 2. Figure 1 shows the result of an analysis of two different progeny genotypes of cross 0204 using pentaplex-PCR amplification for five microsatellite loci in 10 µl reaction volumes. Multiplexing of microsatellites together with SCARs required doubling the amount of SCAR primers compared with the amount of microsatellite primers. No other modifications to the manufacturer's protocol (Qiagen) were necessary. Only 13 of the 3366 plants could not be assigned with respect to presence/absence of the resistance alleles of one or more of the markers in the first attempt using the modified protocol, corresponding to a failure rate of $<0.4\%$.

The microsatellite data were analysed for presence/absence of the diagnostic fragment and also used to determine the genotype (the combination of alleles for all microsatellites) of all progeny plants. Thus, one could assess, for each cross, the percentage of selfers (i.e. progeny from self-fertilizations) as well as the percentage of outcrossers (progeny accidentally pollinated by plants other than the regular pollen source). The frequency of selfers did not vary much between the different crosses and was generally low, rarely exceeding 3% (Table 2). Low levels of selfing in apples are also reported by breeders (Kellerhals et al. 1997). In contrast, the percentage of outcrossers varied greatly between crosses and covered a range from 0 to over 30% (Table 2). One-third of all crosses showed more than 10% outcrossing (Table 2).

The new method allowed a considerable reduction in material costs and labour time compared with a method used earlier which was mainly relying on an efficient DNA extraction procedure (Dilworth and Frey 2000; Table 3). It now allows the analysis of five microsatellite markers in 96 individual plant samples for less than US$ 211 and with <4 h hands-on time.

Discussion

The new MAS method described here required six steps: (1) plant leaf collection, (2) labelling and storage, (3) leaf sample collection and DNA extraction, (4) multiplex-PCR amplification of fluorescently labelled microsatellite markers, (5) PCR product preparation and capillary electrophoresis, and finally

Table 3: Comparison of requirements in labour time and material costs of the new method with a previously used method

	2000	2003	Change (%)
Number of samples	650	3366	+520
Multiplexing (degree)	1–2	5[1]	+470
DNA extraction time (min for 96 samples)	120	30	−75
Error rate (% in first round)	27	0.4	−99
Material costs per plant (US$)	1.83	2.19	+20
Material costs per marker (five markers in 2003)	0.92	0.44	−52
Hands-on time (min) per run (96 samples)	420	230	−45
Hands-on time per plant (min)	4.38	2.40	−45

[1] Successfully tested with eight multiplex markers.

(6) data analysis. The hands-on time required for sample handling could be significantly reduced. For 96 plant samples, leaf disc sampling required 2 h, DNA extraction 30 min, and setting up the robotic station another 30 min. Thus, hands-on time for the first three steps took only 2 min per plant. No more than 5 min were then required to take the plate containing the 96 samples from the robotic station, briefly shake and centrifuge, load the plate in the thermal cycler and put the samples back into storage. After the PCR amplification, about 50 min were required to prepare 96 samples and to set up the ABI 3100 for fragment analysis. In the final step, about 1 h was required for computer analysis of the results for 96 samples produced by the capillary electrophoresis. The total hands-on time from leaf sampling in the greenhouse to genotyping-ready data was thus under 4 h (230 min) per 96 sample reaction plate, or 2.4 min per individual plant. This is very short as, until recently, DNA extraction alone often took up to 1 day for 50–100 individual plants. Since labour is by far the most expensive single item in any cost calculation, this makes the new method a very economic solution to high-throughput MAS and any other plant analysis requiring large-sample sizes together with many markers. The time required for the genotyping process depends on the depth of analysis, i.e. for scoring presence/absence of markers, only approximately 1 h per 96 plants is needed, whereas haplotyping may be distinctly more time-consuming, depending on the complexity of the markers (Liebhard et al. 2002).

Although the new method requires more sophisticated materials (e.g. fluorescently labelled primers, capillary electrophoresis materials) than the previously used method (Dilworth and Frey 2000), the costs of disposable materials and chemistry could be kept low because of reduced reaction volumes and low failure rates (Table 3). Material costs per plant amount to less than US$ 2.20 (i.e. for a pentaplex analysis only US$ 0.44 per marker). Moreover, there is a great potential to further reduce costs by increasing the degree of multiplexing. Octaplex-PCR has been successfully used for several hundred analyses and if routine decaplex analysis is possible this would reduce costs per marker to US$ 0.22. Therefore, compared with the already fairly economic method previously developed (Dilworth and Frey 2000), the high-throughput MAS method presented here represents important progress. Both labour time as well as material costs per marker and plant could be dramatically reduced. The method compares favourably with other currently used MAS systems. For example, the costs established by Dreher et al. (2003) for multiple simple sequence repeat (SSR) marker analysis for reagents and supplies are estimated at US$ 0.59 (45% of US$ 1.31; see Table 2 in their paper) per marker and plant with a sample size of 1000 plants, and the time used for genotypic screening of 100 plants is 8 h which is about 30% more than is required for the method proposed here. The DNA extraction is claimed by the manufacturer of the extraction buffers (Sigma) to work for many different plant species. This was confirmed by successfully amplifying the chloroplast *psbA* gene from 15 important crop plants using method (primers and amplification protocol) described in Frey et al. (1999). Furthermore, other buffers were recently released for DNA extraction from animals with the same strategy as for plants (buffers in the kits Extract-N-Amp™ Tissue PCR Kit and Extract-N-Amp™ Blood PCR Kit, respectively; both from Sigma), enabling the method to be used without major modification in animal breeding. The savings in cost and labour using the newest method presents an important contribution to widening the use of MAS in crop breeding.

Acknowledgements

The authors wish to thank Brion Duffy, Andrea Patocchi, Monika Pfunder, two anonymous reviewers and W.E. Weber for valuable comments, and Matthias Gaberthüel and Syngenta for providing samples of 15 crop species.

References

Davis, G. P., and S. K. DeNise, 1998: The impact of genetic markers on selection. J. Anim. Sci. **76,** 2331—2339.

Dekkers, J. C. M., and F. Hospital, 2002: The use of molecular genetics in the improvement of agricultural populations. Nat. Rev. Genet. **3,** 22—32.

Dilworth, E., and J. E. Frey, 2000: A rapid method for high throughput DNA extraction from plant material for PCR amplification. Plant Mol. Biol. Rep. **18,** 61—64. See also Erratum in Issue 18, 157.

Dreher, K., M. Khairallah, J.-M. Ribaut, and M. Morris, 2003: Money matters (I): costs of field and laboratory procedures associated with conventional and marker-assisted maize breeding at CIMMYT. Mol. Breed. **11,** 221—234.

Frey, J. E., H. Müller-Schärer, B. Frey, and D. Frei, 1999: Complex relation between triazine-susceptible phenotype and genotype in the weed *Senecio vulgaris* may be caused by chloroplast DNA polymorphism. Theor. Appl. Genet. **99,** 578—586.

Hemmat, M., S. K. Brown, and N. F. Weeden, 2002: Tagging and mapping scab resistance genes from R1270-7a apple. J. Am. Soc. Hort. Sci. **127,** 365—370.

Hospital, F., L. Moreau, F. Lacoudre, A. Charcosset, and A. Gallais, 1997: More on the efficiency of marker-assisted selection. Theor. Appl. Genet. **95,** 1181—1189.

Kellerhals, M., W. Mueller, L. Bertschinger, C. Darbellay, and W. Pfammatter, 1997: Obstbau. Landwirtschaftliche Lehrmittelzentrale, Zollikofen, Switzerland.

Kellerhals, M., B. Koller, E. Dilworth, and C. Gessler, 2000: Advances in marker-assisted apple breeding. Acta Hort. **538,** 535—540.

Koebner, R. M. D., and R. W. Summers, 2003: 21st century wheat breeding: plot selection or plate detection? Trends Biotechnol. **21,** 59—63.

Liebhard, R., L. Gianfranceschi, B. Koller, C. D. Ryder, R. Tarchini, E. Van De Weg, and C. Gessler, 2002: Development and characterisation of 140 new microsatellites in apple (*Malus* × *domestica* Borkh.). Mol. Breed. **10,** 217—241.

Liebhard, R., B. Koller, L. Gianfranceschi, and C. Gessler, 2003: Creating a saturated reference map for the apple (*Malus* × *domestica* Borkh.) genome. Theor. Appl. Genet. **106,** 1497—1508.

Markussen, T., J. Krüger, H. Schmidt, and F. Dunemann, 1995: Identification of PCR-based markers linked to the powdery-mildew-resistance gene *Pl1* from *Malus robusta* in cultivated apple. Plant Breeding **114,** 530—534.

Moreau, L., A. Charcosset, F. Hospital, and A. Gallais, 1998: Marker-assisted selection efficiency in populations of finite size. Genetics **148,** 1353—1365.

Ribaut, J.-M., and D. Hoisington, 1998: Marker-assisted selection: new tools and strategies. Trends Plant Sci. **3,** 236—239.

Tartarini, S., L. Gianfranceschi, S. Sansavini, and C. Gessler, 1999: Development of reliable PCR markers for the selection of the *Vf* gene conferring scab resistance in apple. Plant Breeding **118,** 183—186.

Urbanietz, A., H. Schmidt, and F. Dunemann, 1999: Molecular markers in early seedling tests for scab and mildew in apples. In: K. R. Tobutt, and E. H. Alston (eds), Proc. Eucarpia Symp. Fruit Breeding and Genetics. Acta Horticulturae **484,** 429—434.

Young, N. D., 1999: A cautiously optimistic vision for marker-assisted breeding. Mol. Breed. **5,** 505—510.