

Development of a cost-effective pyrosequencing approach for SNP genotyping in barley

C. Silvar^{1,2}, D. Perovic^{2,*}, A. M. Casas¹, E. Igartua¹ and F. Ordon²

¹Department of Genetics and Plant Production, Aula Dei Experimental Station, CSIC, E-50080 Zaragoza, Spain.

²Julius Kühn-Institute (JKI), Federal Research Institute for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Erwin-Baur-Str. 27, 06484 Quedlinburg, Germany

* Address for correspondence:

Dragan Perovic, Julius Kuehn-Institute (JKI), Federal Research Institute for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Erwin-Baur-Str. 27, 06484 Quedlinburg, Germany

E-mail: dragan.perovic@jki.bund.de

Telephone: +49(0)394647611

Fax: +49(0)394647600

Abstract

An improved, efficient, reliable and cost-effective pyrosequencing protocol for SNP genotyping in plants is described. Labelling of the PCR products, required for the single-stranded pyrosequencing assay is carried out in a one-step PCR reaction with a universal biotinylated M13 primer of 19-bp. A ratio of 1:10 of tailed-primer to M13 primer in a 15 µl reaction volume turned out to be best suited for a successful identification of different alleles based on known SNPs. This technology allows cost effective SNP genotyping in plant genomes for loci of interest at an overall low cost in a short period of time and is therefore best suited for marker based selection procedures.

Key words: SNP, pyrosequencing, M13 tail, barley, breeding, MAS, cost-effective

Introduction

Molecular markers have been widely used in plant breeding for phylogenetic studies, comparative genomics, mapping of genes or quantitative trait locus (QTL) and marker assisted selection (MAS) (Varshney et al. 2005, Ganal et al. 2009). The most abundant molecular markers in plant and animal genomes are Single Nucleotide Polymorphism (SNP), which constitute single base-pair variations in the DNA sequence. They are becoming the marker of choice in plant breeding, as they are easily prone to automation and high-throughput. Consequently, genomic resources based on SNPs for crop plant species are growing fast (e.g. Close et al. 2009). Nevertheless, for breeding purposes it is often the case that one specific SNP has to be analysed on thousands of genotypes, because a certain allele is indicative for a desired trait (e.g. Tyrka et al. 2004). At this scale, the most popular method for this purpose is still SNP genotyping using Cleaved Amplified Polymorphic Sequences (CAPS) (Konieczny and Ausubel 1993), which is based on the digestion with a restriction endonuclease. This technology, however, has some drawbacks, as the possibility of false scorings due to partial restriction enzyme digestion and the limitation to SNPs for which a restriction enzyme is available. Therefore, the development of an accurate and low-cost methodology to assay SNP-based markers in plant breeding is still necessary.

For this purpose we propose an improvement on the SNP detection procedure through pyrosequencing (Ronaghi et al. 1998). Pyrosequencing is a real-time sequencing technique based on the detection of released pyrophosphate during DNA synthesis. This method involves the conversion of the PCR product into single-strand DNA and isolation of one of the strands, usually through labelling with biotin. Pyrosequencing is accurate, flexible, and can be easily automated (Langae and Ronaghi 2005). Several applications of this technique for SNP genotyping in plant genetic studies have been proposed (Mochida et al. 2003, Varshney et al. 2008, Hanemann et al. 2009, Wang et al. 2010). However, one major disadvantage of this technology, which notably increases the expenses, is the need for primers typically labelled with biotin.

The aim of this work is to develop a universal scheme for the conversion of SNP markers into easy-to-handle and cost-effective pyrosequencing markers suitable for application in plant breeding.

Material and Methods

Plant Material. Seedlings from the barley landraces 'SBCC097' (Igartua et al. 1998) and 'MBR1012' and the cultivars 'Plaisant' and 'Scarlett' were harvested at the two leaf stage and frozen in liquid nitrogen. Total genomic DNA was extracted using either the NucleoSpin Plant XL kit (Macherey–Nagel) or the CTAB method according to Stein et al. (2001).

Primer design and PCR optimization. A set of SNP markers, previously analyzed in an Illumina Golden Gate assay (BOPA1, Close et al. 2009), were employed to set up the new protocol (Table 1). Each pyrosequencing marker comprises a set of three primers: forward, reverse and sequencing primers. All primers were designed using the *Pyrosequencing Assay Design* Software (Biotage). The primer design parameters included a 'primer length' up to 21 bp and an 'optimal product size' between 150 and 200 bp. An M13 tail (5'-CACGACGTTGTAAAACGAC-3') of 19 bp was added to the 5' end of the forward primer. The sequencing primer was always designed on the complementary strand to the M13-tagged primer. The biotin label was incorporated to a M13 universal primer with a complementary sequence to the M13 tail.

First, we optimized the PCR conditions without the M13 primer. PCR reactions were performed in a total volume of 15 μ l, containing 25 ng of genomic DNA, 1 \times PCR reaction buffer (Solis Biodyne, Tartu, Estonia), 2.5 mM Mg²⁺, 0.2 mM of dNTPs (Fermentas, St. Leon-Rot, Germany), 0.2 mM of each primer (Biolegio, AB Nijmegen, Netherland) and 0.5 U *Taq* FIREPol polymerase (Solis Biodyne). *Taq* DNA polymerase (Qiagen), FIREPol polymerase (Solis Biodyne) and AmpliTaq® DNA Polymerase (Applied Biosystems), respectively, with their corresponding reaction buffers were used for the parallel assessment of amplification quality and quantity. All fragments were amplified using a touchdown PCR profile in which the annealing temperature was decreased in

0.5°C increments from 62°C to 56°C, followed by 35-40 cycles at 56°C. When using a Hot-Start *Taq* FIREPol polymerase (Solis Biodyne), the PCR programme was modified by increasing the first step at 94°C up to 15 min and the number of cycles up to 45. Amplicons were analysed by electrophoresis on a 1.5% agarose gel in 1×TBE buffer and visualized using ethidium bromide (5µg/ml) staining.

Once the presence of non-specific products was corroborated, the biotinylated M13 universal primer was incorporated to the reaction using different ratios (tailed forward primer: M13 primer): 1:2, 1:5 and 1:10. The amount of reverse primer was kept constant. A control reaction without the third primer was carried out at the same time. The presence of only one amplicon and the strength of amplification were checked by electrophoresis on a 1.5 % agarose gel. In instances where it was desirable to improve the PCR specificity and yield, a Hot-Start Polymerase (Solis Biodyne) was employed.

Pyrosequencing assay. The pyrosequencing assay for SNP genotyping was carried out on a Pyromark ID system (Biotage). The experimental procedure for pyrosequencing assays, including annealing plate preparation, immobilization of PCR products to streptavidin beads and the preparation of single stranded pyrosequencing template DNA were done as described on manufacturer's instructions.

Results and discussion

Two markers (11_0761 and 11_0068) from the barley Illumina BOPA1 (Close et al. 2009), were employed for the optimization of the PCR and pyrosequencing approach. DNA from 'SBCC097' and 'Plaisant' were successfully amplified with specific forward and reverse primers for each locus (data not shown). The first step of PCR represents the critical step for the further pyrosequencing assay since the reliability and reproducibility of the technique largely depends on the quality and quantity of the PCR product. The optimal concentration for M13 primer and the stoichiometry of

three primers was assayed to avoid the amplification of unspecific products and optimize the yield of labelled PCR product. A touchdown programme, as recommended in Schuelke (2000), turned out to be very suitable for obtaining good-quality and strong fragments. Amplification with three primers for these two markers gave a single and specific product of expected size. No differences were observed between different ratios in an agarose gel by visual inspection (Fig. 1a). The standard reaction volume used for PCR was 15 μ l. If a weak PCR product was obtained, the volume was increased up to 25 μ l. We did not notice differences in amplification with cheap (FIREPol) and more expensive polymerases (AmpliTaq® and Qiagen *Taq*) (data not shown); therefore all further steps were performed by using the cheaper FIREPol polymerase.

To assess which concentration of universal primer produces the highest amount of biotinylated amplicons, the above PCR products were used as templates for the pyrosequencing assay. All different ratios tested produced a pyrogram, in which each allele is clearly identified. The different primer stoichiometries did not produce a visible effect on the resulting pyrograms (Fig. 1b).

The new approach described in this paper is similar to the M13-tailed primer method reported earlier by Oetting et al. (1995). That technique has been mostly used in plant and animal genetics for microsatellite genotyping using fluorescent dye labels and capillary electrophoresis or laser detection systems (Anca et al. 2007, Santra et al. 2008). However, to our knowledge, no report exists on the application of an M13 tail for the genotyping of SNP markers. Only Royo et al. (2007) described a method for mutation detection in humans based on an M13 tail and a pyrosequencing scheme. In plants, only a short report by Pacey-Miller and Henry (2003) showed preliminary data on using a universal biotinylated primer in a pyrosequencing approach. The main drawback of the latter method is that two separate PCRs were performed to incorporate the biotinylated tag sequence to the template. In the present work, we overcome such inconvenience by labelling the PCR product with biotin in a single-tube step. Compared to the procedure reported by Royo et al. (2007), one major advantage of our method lies in the length of the M13 tail. Nowadays, many companies offer

the cheapest oligo synthesis tier up to 40 bp. The 24 bp long universal primer proposed by Royo et al (2007) leaves only 16 bp for forward primer design in the cheap tier. In most cases, longer primers will be needed, thus entailing significant cost increases of oligo synthesis. Our 19-bp sequence of M13 universal primer allows the design of forward primers up to 21 bp without increasing the cost of the primer synthesis. Additionally, smaller PCR volumes can be used in our approach compared to Royo et al. (2007) and thereby the cost of chemicals per reaction will be notably reduced.

To evaluate the reproducibility and robustness of the method, we selected the ratio 1:10 (0.2 μ M M13 primer: 0.02 μ M forward primer) for further work on another four BOPA1 SNP markers and two different lines ('MBR1012' and 'Scarlett'). No other PCR components or conditions for these markers were optimized. PCR products of expected size and expected pyrograms were obtained for all 4 different markers and a polymorphisms was detected in three of them (Fig. S1). Therefore, we can conclude that all the time-consuming preliminary steps for primer optimization do not need to be undertaken for each new primer set or DNA material. This feature constitutes a considerable improvement compared to previously available procedures and a notable profit for rapid and successful implementation of this technology into the laboratory, since individual optimization of each SNP assay is not necessary. Additionally, the method proposed successfully identified SNPs that differed in the composition of adjacent sequence (which frequently entails the development of specific procedures to assay each SNP). Therefore, this method facilitates to assay a panel of SNPs with a single procedure, avoiding the need to deploy a suite of different genotyping methods for each specific SNP.

The technology presented here provides an opportunity for laboratories possessing a pyrosequencer to optimize its use by reducing notably the expenses. This method will be best suited for all cases in which a limited number of SNPs has to be tested on a large amount of genotypes, e.g. in high resolution mapping procedures and in marker based selection schemes. Additionally, the flexibility

provided by our method to explore a large number of genotypes, together with the ‘sequencing by synthesis’ technology inherent in the pyrosequencer, opens possibilities for SNP discovery in a much cheaper way. It would be possible to identify SNPs which are known (as it has been demonstrated in MBR1012 and Scarlet) but also to discover new SNPs in the neighbourhood of known SNPs while analyzing new genotypes in germplasm collections (see SNP 11_0537 in Fig. S1). Thus, new alleles for known genes may be discovered.

In conclusion, the major advantages of the approach presented here are a reduction of the overall cost of pyrosequencing (cheap *Taq* polymerase, only one general biotinylated primer, small PCR volumes and only one single-tube reaction for labelling of PCR products) and of the requirements for individual assay optimization, while maintaining the specificity, flexibility and accuracy of a typical pyrosequencing approach.

Acknowledgments

This study was partially carried out within project AGL2007-63625, funded by the Spanish Ministry for Science and Innovation. C.S. holds an I3P contract from CSIC. C.S. was supported by mobility fellowships from the Deutsche Forschungsgemeinschaft (DFG), CSIC, Fundacion Caja Inmaculada and COST Action FA0604 (Tritigen).

References

- Anca, M. V., S. B. Andersen, D. Banga, M. Ardelean, and A. M. Torp, 2007: Genetic mapping of wheat *2bl* chromosome using SSR markers. *Buletin USAMV-CN* **64**, 1454–1459.
- Close, T. J., R. B. Prasanna, L. S. Lonardi, Y. Wu, N. Rostoks, L. Ramsay, A. Druka, N. Stein, J. T. Svensson, S. Wanamaker, S. Bozdog, M. L. Roose, M. J. Moscou, S. Chao, R. K. Varshney, P. Szűcs, K. Sato, P. M. Hayes, D. E. Matthews, A. Kleinhofs, G. J. Muehlbauer, J. DeYoung, D. F. Marshall, K. Madishetty, R. D. Fenton, P. Condamine, A. Graner, and R. Waugh, 2009: Development and implementation of high-throughput SNP genotyping in barley. *BMC Genomics* **10**, 582–594.

- Ganal, M. W., T. Altmann, and M. S. Röder, 2009: SNP identification in crop plants. *Curr. Opin. Plant Biol.* **12**, 211–217.
- Hanemann, A., G. F. Schweizer, R. Cossu, T. Wicker, and M. S. Röder, 2009: Fine mapping, physical mapping and development of diagnostic markers for the *Rrs2* scald resistance gene in barley. *Theor. Appl. Genet.* **119**, 1507–1522.
- Igartua E, M. P. Gracia, J. M. Lasa, B. Medina, J. L. Molina-Cano, J. L. Montoya, and I. Romagosa, 1998: The Spanish barley core collection. *Genet. Resour. Crop Ev.* **45**, 475–481.
- Konieczny, A., and F. M. Ausubel, 1993: A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* **4**, 403–410.
- Langaee, T., and R. M. Ronaghi, 2005: Genetic variation analyses by Pyrosequencing. *Mutation Res.* **573**, 96–102.
- Mochida, K., Y. Yamazaki, and Y. Ogihara, 2003: Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Mol. Gen. Genomics* **270**, 371–377.
- Oetting, W.S., H. K. Lee, D. J. Flanders, G. L. Wiesner, T. A. Sellers, and R. A. King, 1995: Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics* **30**, 450–458.
- Pacey-Miller, T., and R. Henry, 2002: Single-nucleotide polymorphism detection in plants using a single-stranded pyrosequencing protocol with a universal biotinylated primer. *Anal. Biochem.* **317**, 165–170
- Ronaghi, R. M., M. Uhlen, and P. Nyren, 1998: A sequencing method based on real-time pyrophosphate. *Science* **281**, 363–366.
- Royo, J. L., M. Hidalgo, and A. Ruiz, 2007: Pyrosequencing protocol using a universal biotinylated primer for mutation detection and SNP genotyping. *Nat. Protoc.* **2**, 1734–1739.
- Santra, D. K., X. M. Chen, M. Santra, K. G. Campbell, and K. K. Kidwell, 2008: Identification and mapping QTL for high-temperature adult-plant resistance to stripe rust in winter wheat (*Triticum aestivum* L.) cultivar ‘Stephens’. *Theor. Appl. Genet.* **117**, 793–802

- Schuelke, M., 2000: An economic method for the fluorescent labelling of PCR fragments. *Nature Biotech.* **18**, 233–234.
- Stein, N., G. Herren, and B. Keller, 2001: A new DNA extraction method for high-throughput marker in a large-genome species such as *Triticum aestivum*. *Plant Breed.* **120**, 354–356.
- Tyrka, M., L. Blaszyk, J. Chelkowski, V. Lind, L. Kramer, M. Weilepp, H. Wisniewska, and F. Ordon, 2004: Development of the single nucleotide polymorphism marker of the wheat *Lr1* leaf rust resistance gene. *Cell Mol. Biol. Lett.* **9**, 879–889.
- Varshney, R. K., A. Graner, and M.E. Sorrells, 2005: Genic microsatellite markers in plants: features and applications. *Trends Biotech.* **23**, 48–55
- Varshney, R. K., T. Thiel, T. Sretenovic-Rajicic, M. Baum, J. Valkoun, P. Guo, S. Grandó, S. Ceccarelli, and A. Graner, 2008: Identification and validation of a core set of informative genic SSR and SNP markers for assaying functional diversity in barley. *Mol. Breed.* **22**, 1–13.
- Wang, G., I. Schmalenbach, M. von Korff, J. Leon, B. Kilian, J. Rode, and K. Pillen, 2010: Association of barley photoperiod and vernalization genes with QTLs for flowering time and agronomic traits in a BC2DH population and a set of wild barley introgression lines. *Theor. Appl. Genet.* **120**, 1559–1574

Table 1. Marker name and primer sequences for BOPA1 SNPs employed for optimization of the pyrosequencing assay.

Marker ¹	Unigene ²	SNP	Forward primer ³ (5'–3')	Reverse primer (5'–3')	Sequencing primer (5'–3')	Sequence to analyze	PCR product (bp)
11_0761	481	C/G	GCTGTCAGAGTCATCAACAC	AATAATGGTGCACGACAAGT	CAAGCTCCATCAATCG	GC/GCATGGTACCA	157
11_0068	14616	C/G	CATTTTCTGGGCATTACG	AAGGACTACGACGAGAATG	CTGGGAGGATCTGGT	CTC/GAAGAAGAGTG	154
11_1372	17115	A/G	AATTCTTGCGGATTACTC	GCTACCATGTGATAGAGCT	TCCAGTGCCCATCAT	C/TTATTCCAGTA	188
11_0537	15977	A/C	TCCCTTGTTTCATAGAGTCG	GCGGTTTTGATAATTTAGAAG	CTCAGAGAGATTCCCG	WG/TCGGAAACAAT	187
11_0036	14496	A/G	CCAGCAGACAAATGACAC	AAGGGTGGATACACCATCT	CCATCTCCGACAACCTC	GAGC/TGGAAACAAG	153
11_1129	15178	A/G	GCCATCCATCTTGTGTC	TTACAGCACGAACACCAT	ACCCTAGCTTGGTAGTTC	C/TGCCATCAGAA	186

¹Illumina OPA markers are designated according to BOPA1 name (Close et al. 2009)

²Number of Unigene in HarvEST Assembly #35 (<http://harvest.ucr.edu>)

³The biotinylated M13 (CACGACGTTGTAACGAC) was added at the 5' end of each forward primer

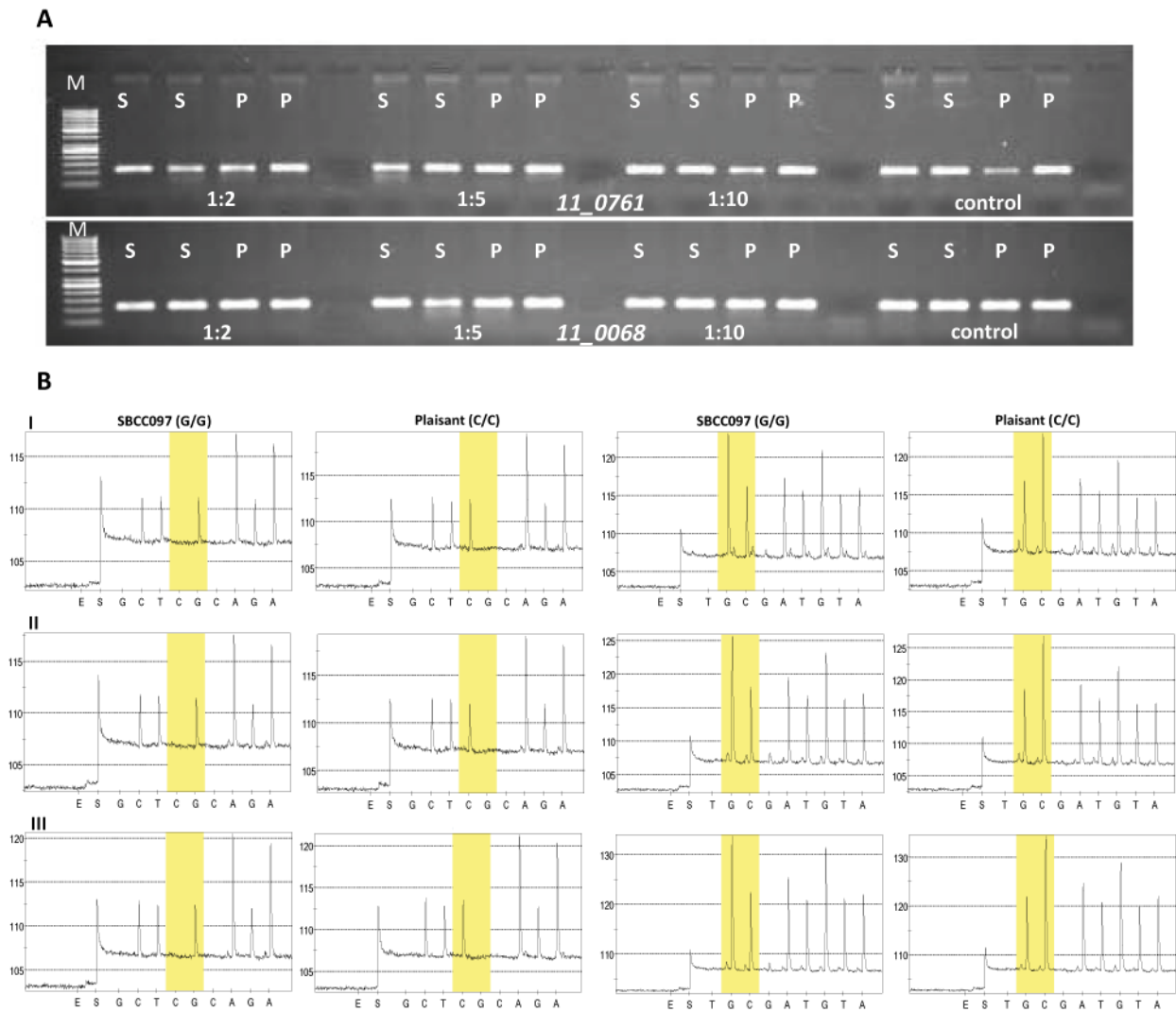


Figure 1. (A) PCR amplification of two BOPA1 markers using three primers (M13, tailed-forward and reverse primers). S, SBCC097; P, Plaisant; M, 50 bp marker. Control indicates a PCR reaction without M13 primer. (B) Pyrograms obtained with two different SNP BOPA1 markers; 11_0068 (right) and 11_0761 (left) on lines SBCC097 and Plaisant. The pyrograms observed by pyrosequencing analysis represent each possible genotype. I, II and III indicate the ratio of tailed forward primer to M13 primer; 1:2, 1:5 and 1:10, respectively. Letters on the X-axis indicate the dispensation order established for the machine according to the “sequence to analyze”. *E* and *S* stand for enzyme and substrate reagents, respectively.