Volume 61(1):77-83, 2017
Acta Biologica Szegediensis
http://www2.sci.u-szeged.hu/ABS

ARTICLE

Discrimination of grape varieties by Start Codon Targeted genotyping using partially degenerate primers

Krisztina Miró, Tibor Nagy, Edit Korom, Ferenc Marincs*

Agricultural Biotechnology Institute, National Agricultural Research and Innovation Center, Gödöllő, Hungary

ABSTRACT DNA fingerprinting of crop species should be technically simple and easy-to-perform, reproducible, and should provide sufficient amount of information. PCR-based methods can meet one or more of these criteria, but they often employ multiple discrete primers or require to test large number of arbitrary primers to provide enough information, which make these methods technically complicated. Our aim was to develop a simple, reproducible, PCR-based method for grape genotyping, which overcomes these limitations. We tested twelve, partly degenerate primers to genotype 14 Hungarian and international grape varieties and found one primer producing 17 polymorphic bands after data normalization, which was sufficient to separate the varieties. The discriminating power of this primer, in term of the number of polymorphic bands, PIC and Rp values, was the same or better than the SCoT primers with definite sequences described in previous studies. The phylogenetic tree obtained using sequences amplified with this primer was reliably consistent with the publicly available information about the genetic origin of some of the tested varieties. We developed a simple and accurate method to genotype grapevine, which provided sufficient amount of data to discriminate 14 varieties.

Acta Biol Szeged 61(1):77-83 (2017)

KEY WORDS

degenerate primer
genotype
grapevine cultivar
PCR
SCoT marker

Introduction

PCR-based DNA fingerprinting techniques are popular in plant biology for genotyping, species/variety separation and phylogenetic studies, because they are relatively easy to perform. These methods can be broadly categorized into three groups (Poczai et al. 2013). Conserved DNA-, gene family-, transposon-, resistance gene- and RNA-based fingerprinting techniques require prior knowledge of the target gene sequences for primer design. RAPD, ISSR and AFLP fingerprinting methods use arbitrary primers and therefore no knowledge about the amplified sequences is needed. Targeted fingerprinting methods, such as SRAP (Sequence-Related Amplified Polymorphism) and SCoT (Start Codon Targeted) are between the arbitrary and other targeted techniques, since they require no or limited sequence information but still target genomic regions more or less specifically.

Of the targeted fingerprinting methods, SRAP genotyping employs 16-18 bp long forward and reverse primer pairs. The forward and reverse primers have constant, but different sequence, and their 3' triplet can be varied, which provide a certain level of amplification specificity. The SRAP molecular marker system was developed for *Brassica* (Li and Quiros 2001) and has been used in turf grass (Budak et al. 2004), bean (Alghamdi et al. 2012) and in a number of other crop species (Aneja et al. 2012). SCoT genotyping (Collard and Mackill 2009) is based on the amplification of genomic DNA using a single primer, which targets the conserved region around the start codon of highly (for example LEA, seed storage, mitochondrial, ribosomal, proline-rich and glycine-rich proteins, histones, globulins, calmodulins, just to name a few) and lowly (for example regulatory, signal transduction and cell wall proteins) expressed plant genes (Sawant et al. 1999). SCoT has been used to fingerprint rice (Collard and Mackill 2009), mango (Luo et al. 2010), potato (Gorji et al. 2011), peanut (Xiong et al. 2011), and ramie (Satya et al. 2015).

In grape (*Vitis vinifera*), which is a major crop worldwide with numerous international and very large number of local varieties, RAPD has been quite widely used for variety identification, polymorphism and diversity studies (Kocsis et al. 2005; Karataş and Ağaoğlu 2008; Salayeva et al. 2010; Butiuc-Keul et al. 2011; Zhao et al. 2011). SRAP and SCot genotypings were employed less frequently to study the relationship of cultivated grape varieties and/or wild grape species (Guo et al. 2012a; Liu et al. 2012).

Submitted November 16, 2016; Accepted January 31, 2017
*Corresponding author. E-mail: marincs.ferenc@abc.naik.hu

Ideally, a PCR-based genotypic method should produce a reasonable number of polymorphic bands, as only those have a diagnostic value. Moreover, this goal should be achieved using as few as possible primers or primer combinations. In methods such as RAPD, which use arbitrary primers, or in targeted fingerprinting, such as SRAP and SCoT, the design of primers is easy, since it can be done *in silico*. However, the problem with these methods is the large number of possible primer variations and consequently a large-scale testing of them is a major concern. For example, in RAPD random hexamers are used as primers and there are 4,096 variations of such oligonucleotides. In SRAP, the number of the 3' triplet variations for one primer can be 64, and because SRAP primers are used in pairs, 4,096 (64 x 64) combinations are possible. This number increases further by varying the core region of the primers. Even in SCoT genotyping, in which the number of possible primers is far less than in RAPD and SRAP, several dozens of primers have been tested (Collard and Mackill 2009).

In this study, we describe a simple and efficient SCoT genotyping of grapevine, in which we tested only twelve partially degenerate primers, and showed that a single primer was able to discriminate 14 cultivated international and local varieties being important for the Hungarian wine industry.

Materials and Methods

Young leaf samples of 14 Hungarian and international wine grape varieties (11 local 'Olaszrizling' ['Italian Riesling'] clones plus 'Furmint', 'Hárslevelű', 'Leányka', 'Ezerjó', 'Szürkebarát' ['Pinot Gris'], 'Rajnai Rizling' ['Rheine Riesling'], 'Chardonnay', 'Kékfrankos', 'Kadarka', 'Portugieser', 'Merlot', 'Pinot Noir' and 'Cabernet Savignon') were obtained from the vineyards of the Research Institute for Viticulture and Enology, National Agricultural Research and Innovation Center, Badacsony, Hungary.

DNA was isolated using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the instructions of the manufacturer, followed by an additional purification step using the Zymo OneStep PCR Inhibitor Removal Kit (Zymo Research, Irvine, CA, USA). DNA samples were quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA) and then diluted to 10 ng/µl for PCR.

PCR was performed in a 20 µl reaction volume comprising 40 ng DNA, 1 µM primer, 200 µM of each dNTPs, 1× final concentration of DreamTaq Green Buffer and 1.25 U DreamTaq Green DNA Polymerase (Thermo Fisher Scientific, Wilmington, DE, USA). The following PCR conditions were used: 1 cycle for 3 min at 94 ºC; 35 cycles of 94 ºC, 1 min, 50 ºC, 1 min, 72 ºC, 1 min; 1 cycle for 10 min at 72 ºC. PCR products were separated on 1% (w/v) agarose gel in 1× TBE buffer and the gels were photographed under UV illumination. Primer sequences (Table 1) were designed from the consensus sequences around the ATG start codon of highly and lowly expressed plant genes (Sawant et al. 1999), and were custom synthesized by Integrated DNA Technologies (Leuven, Belgium).

Gel images were analyzed by the freeware GelAnalyzer (www.gelanalyzer.com) using the "Valley-to-valley" option for background subtraction. Since diffuse bands can have a large total peak area, the peak height of the bands above the background was determined instead of the value of the total peak area. Peak height values in each sample were Z-score normalized in Excel, by which process the mean value and the standard deviation of the normalized intensities became 0 and 1, making the samples comparable. After normalization, the bands in each varieties with Z-score larger than the standard deviation of the normalized intensities (*i.e.* 1) were retained and allocated a value of 1, while the bands that were not present in a particular variety or had a Z-score smaller than 1 were dismissed and allocated the value of 0. The bands, which had only zero values in all varieties, were considered monomorphic, and the bands having both 1 and 0 values across varieties were considered as polymorphic. By this way, seventeen polymorphic bands were identified and the binary matrix obtained from their 0 and 1 values was used to perform cluster analysis of the samples using the R programming environment. The Euclidean distances between grape varieties were calculated using the "dist" function. Distance values were used to visualize the linkage-based clustering by the "hclust" function. Primer's Resolving power (Rp) and the Polymorphic Information Content (PIC) was calculated according to Gorji et al. (2011).

Results

The aim of this study was to design a simplified SCoT genotyping method by reducing the number of potential primers, which should be tested in order to receive polymorphic bands. To do this, six sequences were extracted from each of the conserved nucleotide sequence around the start codon of genes expressed at high or low levels, respectively (Sawant et al. 1999). Each sequence carries the ATG start codon at different position and the surrounding sequences, in which some nucleotides are degenerate (Sawant et al. 1999). Thus, in contrast to definite primers used by others for SCoT genotyping, we used partly degenerate primers in our study (Table 1).

Using single primers, PCRs were performed and the products were analyzed by agarose gel-electrophoresis. No products were detected with primers LE3, LE4 and LE5. For all other primers, the total number of bands for any primer

Table 1. Primers used in this study.

Primer [a]	Sequence (5′-3′) [b]
HE1	AATGGCTNCCT/ACNAC/TA/CCC
HE2	CAATGGCTNCCT/ACNAC/TA/CC
HE3	ACAATGGCTNCCT/ACNAC/TA/C
HE4	AACAATGGCTNCCT/ACNAC/T
HE5	A/CAACAATGGCTNCCT/ACNA
HE6	TA/CAACAATGGCTNCCT/ACN
LE1	NATGGNGNNGNNANANCC
LE2	ANATGGNGNNGNNANANC
LE3	G/AANATGGNGNNGNNANAN
LE4	NG/AANATGGNGNNGNNANA
LE5	NNG/AANATGGNGNNGNNAN
LE6	NNNG/AANATGGNGNNGNNA

[a] HE and LE, primers from around the start codon sequence of highly and lowly expressed plant genes (Sawant et al. 1999); [b] The ATG start codon is underlined.

ranged from twelve (primer LE2) to 44 (primer HE3). With the exception of primer LE2, for which all bands were monomorphic, polymorphic bands outnumbered monomorphic ones (Table 2), although the polymorphic/monomorphic ratio varied substantially from 1.4 to 5.0 (Fig. 1). Band intensities were normalized as described in the Materials and Methods section to make samples comparable. Elimination of the bands whose normalized intensity was below the set threshold reduced the total number of bands in each sample (Table 2). Only polymorphic bands remained after normalization for primer HE6 (Table 2), but the obtained seven bands did not discriminate the 14 varieties (data not shown). Other primers did not discriminate the varieties either, due to the low number of polymorphic bands. For primer HE3, for which 44 bands were detected in the genotyping (Fig. 1), 17 polymorphic bands remained after the thresholding (Table 2 and Supplementary Table 1). This polymorphism obtained by primer HE3 was sufficient to completely discriminate the varieties, but did not discriminate between the 'Olaszrizling' ('Italian Riesling') clones. The intra-varieties discrimination power of primer HE3 is reflected by its high PIC and Rp values of 0.997 and 5.86, respectively.

Cluster analysis of the 14 varieties using the binary matrix generated from polymorphic bands obtained with primer HE3 revealed three major clusters (Fig. 3). In the top cluster, three well-separated clades were observed, containing 'Chardonnay', a mixture of both white and red varieties and 'Merlot', respectively. The middle cluster contains all of the 'Olaszrizling' ('Italian Riesling') clones and two Hungarian white grape varieties, 'Hárslevelű' and 'Leányka', and an international red variety, 'Portugieser'. In the bottom cluster, two clades were observed, containing white ('Pinot Gris' and 'Rheine Riesling') and red ('Kadarka' and 'Pinot Noir') varieties, respectively.

Table 2. Numbers of bands obtained in grapes using degenerate SCoT primers.

Primer	Number of bands					
	Before normalization			After normalization		
	Total	Mono-morphic	Poly-morphic	Total	Mono-morphic	Polymor-phic
HE1	17	4	13	2	1	1
HE2	22	9	13	4	3	1
HE3	44	7	37	17	0	17
HE4	30	8	22	6	1	5
HE5	42	7	35	6	1	5
HE6	25	7	18	7	0	7
LE1	17	5	12	1	1	0
LE2	12	12	0	3	3	0
LE3	0	n.a.	n.a.	n.a.	n.a.	n.a.
LE4	0	n.a.	n.a.	n.a.	n.a.	n.a.
LE5	0	n.a.	n.a.	n.a.	n.a.	n.a.
LE6	26	6	20	4	1	3
Total	235	65	170	50	11	39

[a] n.a. = not applicable

Discussion

In this study, we described the Start Codon Targeted (SCoT) genotyping of 14 Hungarian and international grape varieties. We tested only twelve degenerate primers, which represent a significant reduction compared to other studies. In potato, Gorji et al. (2011) used 12 SCoT primers in 15 combination, although the theoretical number of combination for twelve primers is 144. In both rice and Chinese grape varieties, 36 primers were tested (Collard and Mackill 2009; Guo et al. 2012b), while in mango, peanut and ramie, 33, 18, and 20

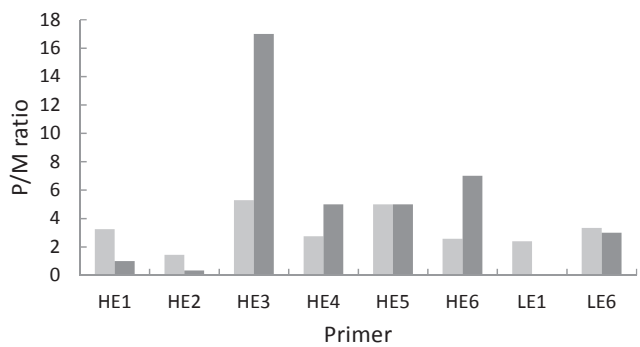

Figure 1. Ratios of the polymorphic/monomorphic (P/M) bands. ■: before normalization; ■: after normalization. For both before and after normalization, for primers with no polymorphic bands, no ratio is shown and for primers with no monomorphic bands, the number of polymorphic bands is shown (Table 2).

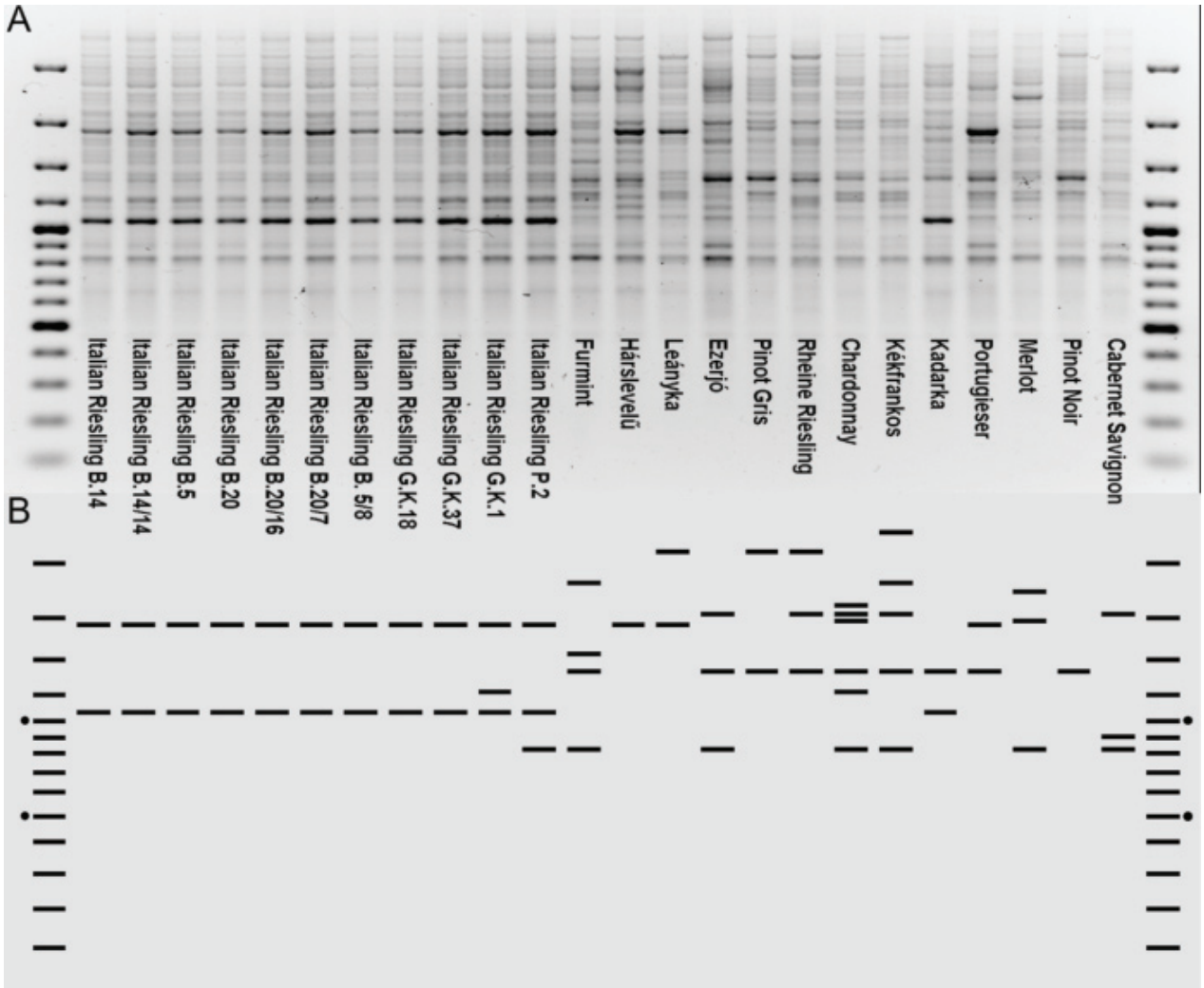

Figure 2. The genotyping profile of grape varieties obtained by the SCoT primer HE3. Panel A: image of the agarose gel electrophoretic separation of the PCR products. Panel B: simulated electrophoretic image of the polymorphic bands after normalisation and thresholding (Supplementary Table 1). The molecular marker is a 100 bp Plus DNA ladder. In Panel B, dots label the 500 and 1000 bp bands with stronger appearance in Panel A.

primers were evaluated, respectively (Luo et al. 2010; Xiong et al. 2011; Satya et al. 2015). In theory, the number of primers with definite sequence can be very high in SCoT genotyping, due to the degenerate nucleotides in the consensus sequence (Sawant et al. 1999). For example, if the degenerate primer HE3 used in this study were converted into definite-sequence primers, 160 primer sequences would be obtained. Therefore, by using degenerate primers, the same number of samples can be tested with less effort than by definite primers, or more samples can be tested, increasing the cost-effectiveness of the method.

Our study with the degenerate primers produced reliable results, compared to that of others with definite primers (Table 3) although, the 72% ratio of the polymorphic bands obtained in this study was only the fourth highest amongst compared studies. The number of polymorphic bands per degenerate primer was much higher than those achieved using definite primers (Table 3). Since three primers did not produce any band in our study, we achieved 235 bands in total by using only nine primers. Of these, 170 were polymorphic, and even after a strict normalization we obtained 39 polymorphic bands in total (Table 2). Here we would like to emphasize the importance of the normalization step in our method, because this can make data interpretation more objective. In contrast to the non-normalized data, where the mean value and the standard deviation vary from sample to sample, the applied Z-score normalization resulted in the same mean value and standard deviation, respectively, of the normalized band in-

Table 3. Comparison of the performance of SCoT genotyping studies.

Number of primers	Number of total bands[a]	Number of polymorphic bands	Polymorphic ratio (%)[b]	Average number of polymorphic bands per primer	Reference
13	n.g.	50	n.a.	3.8	Collard and Mackill (2009)
33	273	208	76	6.3	Luo et al. (2010)
15	130	26	20	1.7	Gorji et al. (2011)
18	157	60	38	3.3	Xiong et al. (2011)
17	131	122	93	7.2	Guo et al. (2012b)
20	136	119	87	5.9	Satya et al. (2015)
9	235	170	72	18.9	This study

[a] n.g. = not given; [b] n.a. = not applicable; in the "Number of primers" column, the boldface and underlined number indicates degenerate primers, normal lettering indicates definite primers.

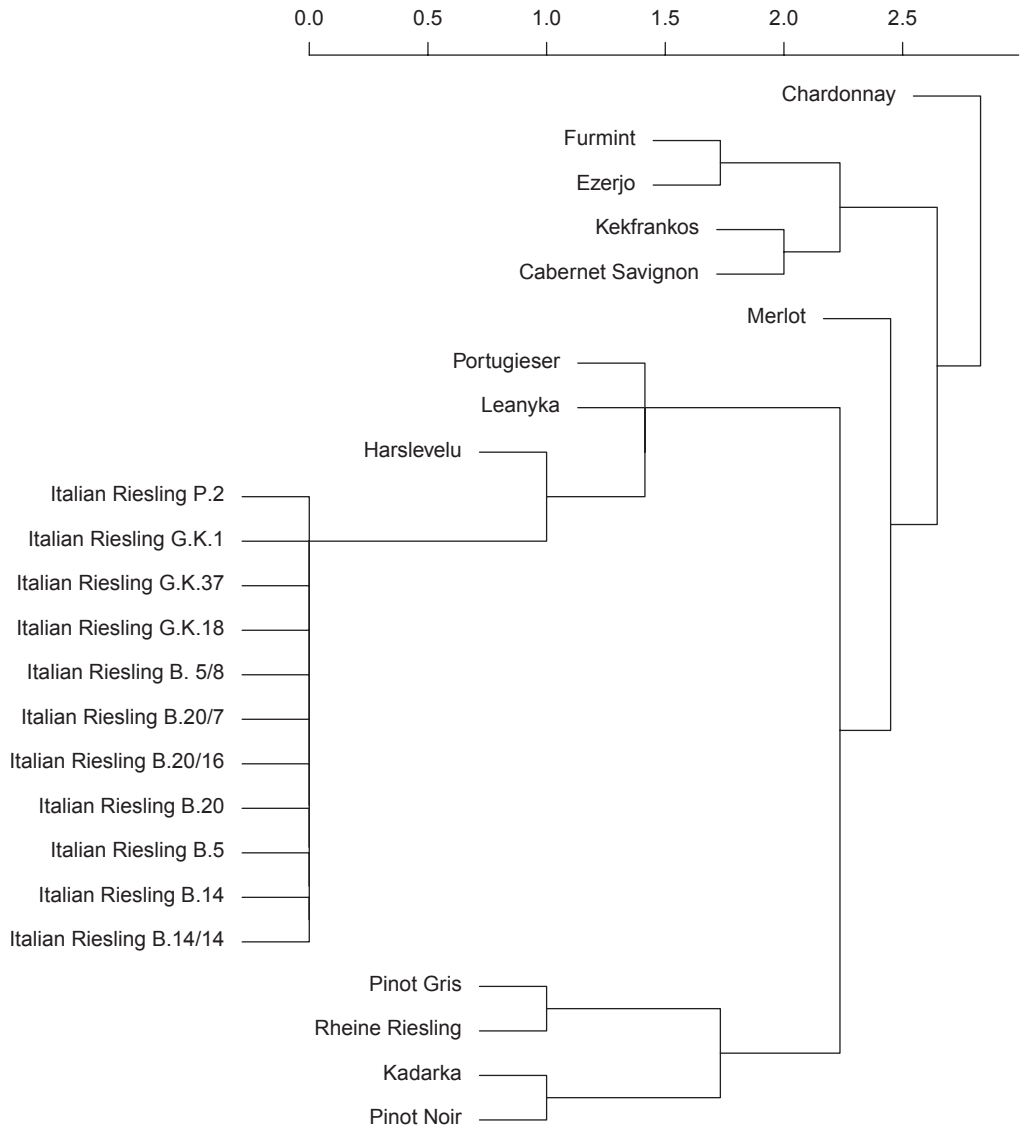

Figure 3. Cluster analysis of 14 grape varieties based on SCoT fingerprinting with primer HE3. Scale indicates Euclidean distance between the varieties.

tensities across varieties (see Materials and Methods). This makes data more comparable and by setting a threshold, under which normalized band intensity values were excluded from the analysis, the visual examination for the absence and the presence of bands can also be avoided. After normalization of data, one primer, HE3, produced 17 polymorphic bands (Supplementary Table 1) and a unique band pattern for each variety (Fig. 2). Thus it had a strong discrimination power reflected by its high PIC and Rp values of 0.997 and 5.86, respectively. For comparison, the highest PIC and Rp values reported by Gorji et al. (2011) and Satya et al. (2015) for a SCoT primer were 0.324 and 3.31, and 0.93 and 5.0, respectively. Although, a PIC value of 1.65 was reported in peanut, the highest number of polymorphic bands was only seven (Xiong et al. 2011). The best PIC value reported for a SCoT primer in grape was 0.91 (Guo et al. 2012b). From these comparisons we can conclude that the degenerate primer HE3 performed the same or even better than definite SCoT primers of other studies.

We wanted to know whether the phylogenetic tree produced by the 17 polymorphic markers is reliable, so we mined the literature to find any data, which might be consistent with the relationships between varieties in the tree. 'Cabernet Franc' was reported as one of the parents for both 'Merlot' and 'Cabernet Sauvignon' (Boursiquot et al. 2009) and this relationship was indicated by their position in the same cluster (Fig. 3). In the 'Pinot' family, the white 'Pinot Gris' variety is a mutant of the black variety 'Pinot Noir' (Yakushiji et al. 2006; Vezzulli et al. 2012), and these varieties were positioned in the same cluster (Fig. 3). 'Portugieser' might be related genetically to almost 100 red and white grape varieties (Regner et al. 1999). Although the Hungarian varieties, 'Hárslevelű' and 'Leányka', being in the same clade as Portugieser (Fig. 3), are not known as putative relatives of 'Portugieser' (Regner et al. 1999), a relationship between them is possible since white grapes have arisen from red varieties by mutations in MYB transcription factors regulating berry color (Walker et al. 2007). The exact genetic origin of the 'Riesling' and the Hungarian varieties included in this study is not known. It was reported, based on a RAPD study, that the Hungarian 'Ezerjó' and 'Hárslevelű' varieties are related (Kocsis et al. 2005), which, however, was not confirmed by our study. Guo et al. (2012b) analysed a number of Chinese and several international grape varieties using definite SCoT primers. Their analysis indicated that two varieties, 'Merlot' and 'Cabernet Sauvignon' are in closely related but separated clusters. Our above-described results are consistent with this and both analyses reflect to the written parentage of the two varieties (Boursiquot et al. 2009).

By employing degenerate primers, we were able to achieve a technically simple and easy to perform SCoT genotyping method, which discriminated 14 grape varieties and may be useful for genetic studies in grapes.

Acknowledgments

This work was supported by the Hungarian Ministry of Agriculture. The authors thank Dr. Gizella Jahnke for the leaf samples.

References

Alghamdi SS, Al-Faifi SA, Migdadi HM, Khan MA, EL-Harty EH, Ammar MH (2012) Molecular diversity assessment using sequence related amplified polymorphism (SRAP) markers in *Vicia faba* L. Int J Mol Sci 13:16457-16471.

Aneja B, Yadav NR, Chawla V, Yadav RC (2012) Sequence-related amplified polymorphism (SRAP) molecular marker system and its applications in crop improvement. Mol Breed 30:1635-1648.

Boursiquot JM, Lacombe T, Laucou V, Julliard S, Perrin F-X, Lanier N, Legrand D, Meredith C, This P (2009) Parentage of merlot and related winegrape cultivars of southwestern France: Discovery of the missing link. Aust J Grape Wine Res 15:144-155.

Budak H, Shearman RC, Gaussoin RE, Dweikat I (2004) Application of sequence-related amplified polymorphism markers for characterization of turfgrass species. HortScience 39:955-958.

Butiuc-Keul AL, Crăciunaş C, Coste A, Farago M (2010) Discrimination and genetic polymorphism in several cultivar of grapevine by RAPD markers. Rom Biotech Lett 15(Suppl):110-115.

Collard BCY, Mackill DJ (2009) Start Codon Targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol Biol Report 27:86-93.

Gorji AM, Poczai P, Polgar Z, Taller J (2011) Efficiency of arbitrarily amplified dominant markers (SCOT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. Am J Potato Res 88:226-237.

Guo D, Zhang J, Liu C, Zhang G, Li M, Zhang Q (2012a) Genetic variability and relationships between and within grape cultivated varieties and wild species based on SRAP markers. Tree Genet Genomes 8:789-800.

Guo DL, Zhang JY, Liu CH (2012b) Genetic diversity in some grape varieties revealed by SCoT analyses. Mol Biol Rep 39:5307-5313.

Karataş H, Ağaoğlu YS (2010) RAPD analysis of selected local Turkish grape cultivars (*Vitis vinifera*). Genet Mol Res 9:1980-1986.

Kocsis M, Járomi L, Putnoky P, Kozma P, Borhidi A (2005) Genetic diversity among twelve grape cultivars indigenous to the Carpathian Basin revealed by RAPD markers.

Vitis 44:87-91.

Li G, Quiros CF (2011) Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica*. Theor Appl Genet 103:455-461.

Liu C, Fan X, Jiang J, Guo D, Sun H, Zhang Y, Feng J (2012) Genetic diversity of Chinese wild grape species by SSR and SRAP markers. Biotechnol Biotec Eq 26:2899-2903.

Luo C, He XH, Chen H, Ou SJ, Gao MP (2010) Analysis of diversity and relationships among mango cultivars using Start Codon Targeted (SCoT) markers. Biochem Syst Ecol 38:1176-1184.

Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JPT, Hyvönen J (2013) Advances in plant gene-targeted and functional markers: a review. Plant Methods 9:6.

Regner F, Eiras-Dias JE, Stadlbauer A, Blahous D (1999) Blauer Portugieser, the dissemination of a grapevine. Cienc Tec Vitivinic 14:37-44.

Salayeva S, Akhundova E, Mammadov A (2010) Evaluation of DNA polymorphism among cultivated and wild grapevine accessions from Azerbaijan. Czech J Genet Plant Breed 46:75-84.

Satya P, Karan M, Jana S, Mitra S, Sharma A, Karmakar PG, Ray DP (2015) Start codon targeted (SCoT) polymorphism reveals genetic diversity in wild and domesticated populations of ramie (*Boehmeria nivea* L. Gaudich.), a premium textile fiber producing species. Meta Gene 3:62-70.

Sawant SV, Singh PK, Gupta SK, Madnala R, Tuli R (1999) Conserved nucleotide sequences in highly expressed genes in plants. J Genet 78:123-131.

Vezzulli S, Leonardelli L, Malossini U, Stefanini M, Velasco R, Moser C (2012) Pinot blanc and Pinot gris arose as independent somatic mutations of Pinot noir. J Exp Bot 63:6359-6369.

Walker AR, Lee E, Bogs J, McDavid DAJ, Thomas MR, Robinson SP (2007) White grapes arose through the mutation of two similar and adjacent regulatory genes. Plant J 49:772-785.

Xiong F, Zhong R, Han Z, Jiang J, He L, Zhuang W, Tang R (2011) Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachis hypogaea* L.) genotypes. Mol Biol Rep 38:3487-3494.

Yakushiji H, Kobayashi S, Goto-Yamamoto N, Jeong ST, Sueta T, Mitani N, Azuma A (2006) A skin color mutation of grapevine, from black-skinned Pinot Noir to white-skinned Pinot Blanc, is caused by deletion of the functional VvmybA1 allele. Biosci Biotechnol Biochem 70:1506-1508.

Zhao MZ, Zhang YP, Wu WM, Wang C, Qian YM, Yang G, Fang JG (2011) A new strategy for complete identification of 69 grapevine cultivars using random amplified polymorphic DNA (RAPD) markers. Afr J Plant Sci 5:273-280.