Jiří Drábek[1]
Michaela Smolíková[2]
Ruslan Kalendar[3,4]
Fernando A. Lopes Pinto[5]
Pavel Pavloušek[6]
Karel Klepárník[7]
Ivo Frébort[8]

[1]Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University, Olomouc, Czech Republic
[2]Department of Biochemistry, Palacký University, Olomouc, Czech Republic
[3]Institute of Biotechnology, LUKE/BI Plant Genome Dynamics, University of Helsinki, Helsinki, Finland
[4]RSE "National Center for Biotechnology" under the Science Committee, Ministry of Education and Science of the Republic of Kazakhstan, Astana, Kazakhstan
[5]Department of Photochemistry and Molecular Science, Uppsala University, Uppsala, Sweden
[6]Department of Viticulture and Enology, Mendel University, Brno, Czech Republic
[7]Institute of Analytical Chemistry of the ASCR, Brno, Czech Republic
[8]Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacky University, Olomouc, Czech Republic

Received February 16, 2016
Revised August 30, 2016
Accepted September 13, 2016

Research Article

Design and validation of an STR hexaplex assay for DNA profiling of grapevine cultivars

Although the analysis of length polymorphism at STR loci has become a method of choice for grape cultivar identification, the standardization of methods for this purpose lags behind that of methods for DNA profiling in human and animal forensic genetics. The aim of this study was thus to design and validate a grapevine STR protocol with a practically useful level of multiplexing. Using free bioinformatics tools, published primer sequences, and nucleotide databases, we constructed and optimized a primer set for the simultaneous analysis of six STR loci (VVIi51, scu08vv, scu05vv, VVMD17, VrZAG47, and VrZAG83) by multiplex PCR and CE with laser-induced fluorescence, and tested it on 90 grape cultivars. The new protocol requires subnanogram quantities of the DNA template and enables automated, high-throughput genetic analysis with reasonable discriminatory power. As such, it represents a step toward further standardization of grape DNA profiling.

Keywords:
Grapevine DNA analysis / Multiplex PCR / STRs / *Vitis vinifera* L
DOI 10.1002/elps.201600068

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

DNA profiling based on microsatellites (STR or simple sequence repeats, SSR) is the preferred way of identifying the cultivar of *Vitis vinifera* L. samples [1–6] because the data generated are objective, more precise than those obtained using traditional ampelographic techniques [7], and less expensive than analysis using SNP arrays [8] or genotyping by next-generation sequencing [9]. It also has advantages over profiling of grapevine must proteins [10], anthocyanines [11], amino acids [12], aromatic compounds [13, 14], and chemical elements [15] because the microsatellite profile is not affected by environmental factors such as a soil composition, weather conditions, or vinification [16]. Extrapolating from forensic human applications [17], microsatellite DNA profiling has the potential to be the most reliable and powerful technique for identifying specific cultivar genotypes in order to test seed purity or kinship, or to protect individual property [18–20].

Six microsatellite markers, VVMD5, VVMD7, VVMD27, VVS2, VrZAG62, and VrZAG79, have been selected by the European consortium GenRes081 for use in generating standard profiles of grapevine DNA [21]. However, the laboratories that participated in this consortium employed different combinations of PCR protocols, instruments, and methods for assessing allele length, leading to discordant results. In addition, the consortium's efforts did not exploit the full potential of high throughput genotyping: the GenRes081

Correspondence: Dr. Jiří Drábek, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University, Hněvotínská 5, 779 00 Olomouc, Czech Republic
E-mail: jiri.drabek@upol.cz
Fax: +420-585-632-180

Abbreviation: SSR, simple sequence repeat

Colour Online: See the article online to view Fig. 1 in colour.

protocol uses large amounts of template DNA and requires post-PCR amplicon pooling before electrophoresis. Similar protocols were adopted by Dangl et al. [22] and Laucou et al. [23] who genotyped 4370 grapevine accessions for 20 microsatellites using eight PCR amplifications and three sequencing runs [23, 24]. Such a sequential protocol is not amenable to DNA profiling of low number of samples and from sources where grapevine DNA is present in low quantities, i.e. in archeological excavations [25], must [26], and wine [27].

To address these deficiencies, we here present a new optimized multiplex reaction protocol for six microsatellite markers (VVIi51, scu08vv, scu05vv, VVMD17, VrZAG47, and VrZAG83) that enables highly sensitive, specific, and balanced profiling. We used this new protocol to genotype a set of reference grapevine varieties from our DNA bank, tested the system's sensitivity to variation of the protocol, and investigated its ability to profile compromised samples.

2 Materials and methods

2.1 Samples

Ninety grapevine leaves and/or pips were obtained from France ($n = 31$, S. Lalet, INRA, Marseillan-Plage), Japan ($n = 6$, N. Goto-Yamamoto, Higashi-Hiroshima), Greece ($n = 3$, Nikos, Santorini), Spain ($n = 1$, M. Galimberti, Madrid), Czech Republic ($n = 49$, J. Beneš, České Budějovice, O. Jandurová, Karlštejn, P. Kočár, Pilsen University, P. Maděra, Brno, M. Michlovský, Rakvice, P. Pavloušek, Lednice, J. Stávek, Němčičky, and M. Pražák, Olomouc-Tovéř). A sample of wine bearing a Gruener Veltliner label was bought from a local store. DNA cuvées were prepared by mixing DNA extracted from Portugieser blau and Castel 216-3 grapes in desired ratios.

2.2 DNA extraction

DNA was extracted from grapevine leaves and wine using CTAB and guanidium isothiocyanate. Briefly, 15 mL of the CTAB extracting solution was added to grapevine leaves (1–7 g), or wine (concentrated from 40 to 7 mL by vacuum evaporation), vigorously mixed, and incubated at 65°C for 30 min [28]. The lysate was then cooled by incubating at room temperature for 5 min and vigorously mixed with 5 mL of chloroform. After centrifugation for 5 min at 4700 *g*, the upper aqueous layer was pooled, precipitated with 0.65 × volume of isopropanol, and centrifuged for 10 min at 4700 *g*. The resulting pellet was washed with 80% ethanol and dissolved in 600 μL of DNAzol (Invitrogen, Carlsbad, CA, USA), a commercial solution of guanidium isothiocyanate [29]. The remainder of the extraction was performed according to the manufacturer's protocol.

For liquid nitrogen-pulverized grapevine pips (five pieces per extraction), the same CTAB/guanidium isothiocyanate method was used with proportionally reduced volumes of extraction solutions to enable the use of 1.7 mL tubes.

2.3 DNA quantitation

Quantitative PCR was performed using the short real-time protocol of Valsesia [30] with primers and the FAM-BHQ1 probe (TaqMan, Life Technologies, Austin, TX, USA) for resveratrol synthase. Each assay was performed in triplicate using a thermocycler Mx3000 (Stratagene, La Jolla, CA, USA) in a volume of 20 μL, as previously described [27]. Quantitation was based on calibration curves obtained by serially diluting spectrophotometrically quantitated DNA from leaves of the André grapevine variety in water.

2.4 Multiplex design algorithm

For a multiplex microsatellite PCR protocol to achieve high levels of discrimination while also being suitable for use with compromised DNA templates, it is preferable to use unlinked loci with a high heterozygosity, mutually compatible primers, high genotyping performance, short amplicons [31, 32], and nonoverlapping allele length ranges [33]. To identify suitable loci, we screened 370 previously published primer sequences [34–39] using both bioinformatics techniques and laboratory experiments. We initially discarded all primers assigned a quality index score below 40 by the FastPCR 5.4.56 software package (http://primerdigital.com/fastpcr.html), as well as those having extreme melting temperatures or several compatible regions within grapevine genome [40], and primer pairs producing amplicons over 300 bp in length. We then analyzed the remaining primer pairs with the FastPCR and mixPCR programs (http://www.dnabased.com/eng/index.html), and selected the 11 candidates with the highest potential multiplexing level. One primer in each pair was tagged with a fluorophore (FAM, HEX, or NED), and PCR experiments were performed with each primer pair individually and with multiplexed combinations of pairs. Preliminary experiments resulted in the discovery of some previously unknown microsatellite alleles, which caused the allele length ranges of some of the targeted loci to overlap. To eliminate this problem, we discarded five of the 11 initially selected primer pairs. Of the remaining six pairs, four were used as described in the literature without modification and two were modified by 5' end-tagging one primer in the pair with the sequence 5′-gTT TCT TCg TTg CgT AgT g-3' to increase the distance between amplicons of adjacent microsatellite pairs that were discovered to overlap when analyzed by electrophoresis. The tag sequence was designed to lack any complementarity within the grapevine genome [41] and, concurrently, to enhance nontemplate nucleotide addition [42]. All validation tests [43] were performed after adjusting the primers' concentrations

to achieve acceptable interlocus peak balance in the hexaplex PCR.

2.5 Genotyping and electrophoresis conditions

PCR amplifications were performed using 0.25 ÷ 1 ng of standard grapevine DNA in a 6 μL reaction mixture containing 60 mU of AmpliTaqGold polymerase, 1 × Gold buffer with 1.5 mM $MgCl_2$ (Life Technologies), 0.2 mM of each dNTP (Fermentas, Vilnius, Lithuania), 0.2 μg of nonacetylated bovine serum albumin (Sigma-Aldrich, Prague, Czech Republic), 0.1% Triton X-100 (Sigma-Aldrich), nonlabeled primers (Metabion, Martinsried, Germany, concentrations given in Table 1), and fluorescence-labeled primers (Life Technologies, Table 1).

Cycling was performed in a PTC-150 cycler (MJ Research, Waltham, MA, USA) using a touch-down protocol with facilitated nontemplate adenine addition. The cycle involved 11 min denaturation at 95°C followed by ten cycles of 10 s at 96°C and 50 s at 59°C (decreasing by 1°C in each cycle), then 1 min at 72°C, *n* cycles of 10 s at 96°C, 50 s at 50°C, and 50 s at 72°C, with a final 45 min step at 65°C. Microsatellites were amplified by PCR and analyzed either separately ($n = 15$, total number of PCR cycles $= 25$) or in a multiplex ($n = 18$, total number of PCR cycles $= 28$). Each PCR experiment was accompanied by a negative control without a template. Before electrophoresis, the amplicons were diluted 1:1 with deionized water. An aliquot of 10 μL of injection mastermix containing formamide and fluorescent GeneScan 400 HD ROX internal length standard (50:1, v/v) was pipetted into 96-well plate. Then, either 0.2 μL of diluted amplicons, negative control, or allelic ladder was added into the mastermix and heat-denatured in the thermocycler (3 min at 95°C). The plate was transferred to an ABI 3130 Genetic Analyzer (Life Technologies) with 50 cm capillaries filled with performance-optimized polymer (POP7, Life Technologies). Electrophoresis was performed with the following parameter settings: separation temperature 60°C, injection voltage 1.6 kV, injection time 15 s, run voltage 15 kV, and run time 1500 s. At least two technical replicates were performed for each sample of extracted DNA.

Although denaturing CE is one of the most precise and reliable sizing techniques currently available, it can still be subject to run-to-run variation resulting from differences between instruments, running conditions, the internal standards that are used, and variations in buffer composition among other things [44]. Therefore, every CE experiment performed in this work included an analysis of an allelic ladder prepared by mixing microsatellite amplicons as suggested by Butler et al. for human microsatellites [45] and by Santos et al. for grapevines [46].

The data produced by the Genetic Analyzer was exported into GeneMapper v3.0 (Life Technologies), automatically sized, and analyzed using the default values for microsatellite analysis. Ladder alleles were assigned manually and used to create bins. Then, samples were genotyped by autobinning and manually corrected. Results were compared to literature data where possible.

Allelic tables were exported into PowerMarker v3.25 [47] and GENALEX 6 [48] to perform population genetics characterization of the targeted loci.

2.6 Sequencing

The VrZAG47 locus in the Furmint, Muscat a petits grains blancs, and Rondinella grape cultivars was analyzed by bidirectional Sanger sequencing performed by SEQme, Czech Republic, using the primers AAA CgT gCC CTT AAC gAg CAg AgC A and TAC CTC gTC gTT TCC ACT TCg gTg C.

3 Results

3.1 Optimized protocol

On the basis of protocols adapted from human forensics [33] and nonhuman forensics guidelines [49] where applicable, we designed a PCR multiplex based on six grapevine

Table 1. PCR primers information

Locus	F/R	Fluorophor	Sequence	Linkage group	Concentration (μM)	Stutter level
VVIi51	F	5-FAM	gATCCCAAgAgAACCAAgAAACT	14	0.4	8%
	R		AggCTgATCTCAgTgCATATgTTg		0.4	
ssrVrZAG47	F		gTTCTTggTCTgAATACATCCgTAAgT	5	0.4	10%
	R	5-FAM	ACggTgTgCTCTCATTgTCATTg		0.4	
ssrVrZAG83	F		*gTTTCTTCgTCgTTgCgTAgTg*ACggAggCggTAgATgAgAgg	4	0.04	15%
	R	5-HEX	ggCAACggCTAgTAAATACAACg		0.4	
scu08vv	F	5-HEX	gAgACCCAgCATCgTTTCAAg	11	0.2	9%
	R		AgCAAAATCCTCCCCgTACAAgTC		0.2	
VVMD17	F		*gTTTCTTCgTCgTTgCgTAgTg*AgTCATgACTCgCCAAAATCTgACg	18	0.04	15%
	R	5-NED	CACACATATCATCACCACACgg		0.4	
scu05vv	F	5-NED	CAAgCAgTTATTgAAgCTgCAAgg	12	0.8	12%
	R		ATCATCCATCACACAggAAACAgTg		0.8	
tag	F		*gTTTCTTCgTCgTTgCgTAgTg*		0.5	

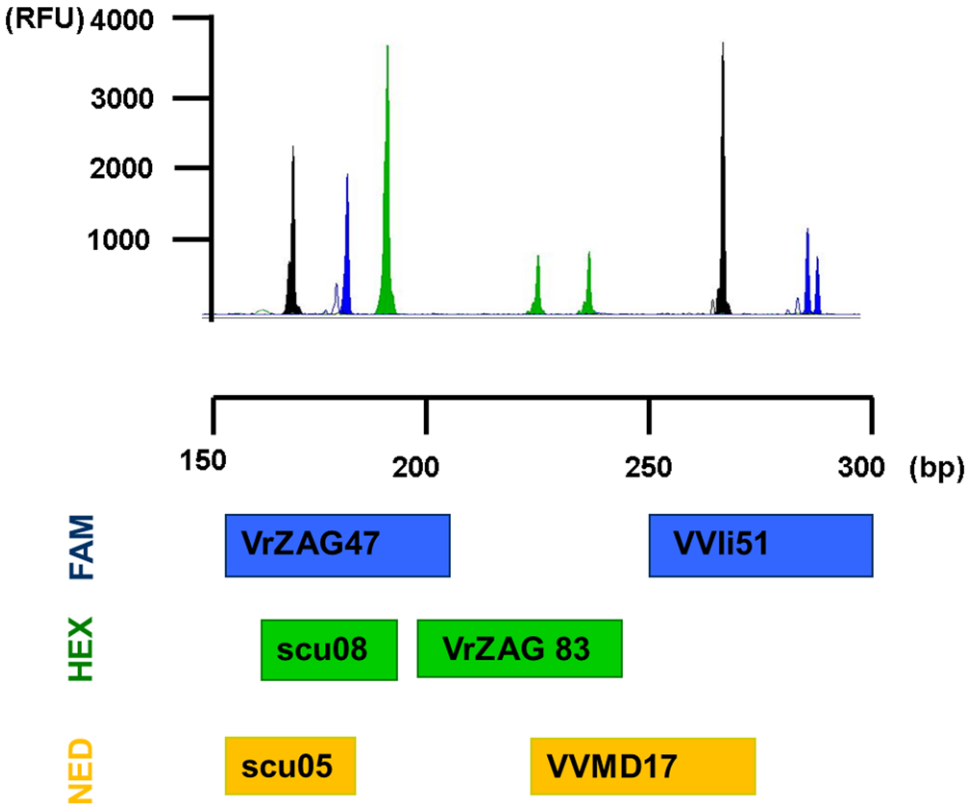

Figure 1. DNA profile of the Verdejo grapevine variety. Upper part: representative PCR profile obtained from Verdejo (VrZAg47 173/173, VVli51 264/266, scu08vv 181/181, VrZAg83 210/220, scu05vv 163/163, VVMD17 247/247). Bottom part: assignment of fluorescent tags to loci, shown on the amplicon length scale.

microsatellite loci that uses the FAM, HEX, and NED fluorescent tags. The final hexaplex consists of the microsatellites VVIi51 [35], scu08vv [39], scu05vv [39], VVMD17 [38], VrZAG47 [36], and VrZAG83 [36] (Fig. 1), with primer concentrations adjusted to achieve interlocus peak balance (Table 1).

As noted in the experimental section, one primer from each pair was tagged with a fluorophore. In addition, one primer from each of the pairs targeting the VrZAG83 and VVMD17 microsatellite loci was tailed by adding 22 nucleotides to enable the binding of a common tag primer during PCR. This modification made it possible to avoid allele overlap between the scu08vv/VrZAG83 and scu05vv/VVMD17 microsatellite pairs (Fig. 1) while retaining the relative length differences between alleles of the same locus.

3.2 Specificity and population study

Profiling data for the tested grapevine varieties in amplicon length form is presented in Supporting Information 1a and in relative length form ($n + x$) in Supporting Information 1b. In all cases, the results obtained from singleplex reactions were fully consistent with those for multiplex reactions ($n = 45$). Genotyping of yeast (*Saccharomyces cerevisiae*) and bacteria (*Escherichia coli*) samples yielded no DNA profiles, while genotyping of a close relative of *V. vinifera* (the Virginia creeper, *Parthenocissus quinquefolia* L.) yielded amplicons for VVIi51 (248/248), scu05vv (151/153), and VVMD17 (232/232).

Despite the high total exclusion probability of the hexaplex (0.988) and low total probability of identity (3.60×10^{-5}), several grapevine varieties yielded identical profiles. The following indistinguishable combinations were identified: the table varieties Aivas (progeny of Moldova × Cardinal), BV18-109 (Estafeta × Augustovskij), and Kodrianka (Moldova × Marszalski); the Muscat varieties Muscotaly sarga and its offspring Muscat Ottonel (Chasselas × Muscotaly sarga); and Pinot noir and its bud mutation Pinot noir praecox.

An unknown sample in our DNA bank that was believed to be a Merlot (progeny of Cabernet Franc and Magdeleine Noire des Charentes [50]) was revealed to be Sauvignon blanc. Other samples without varietal identification, including a grapevine grown by Gregor Mendel in Brno, a wild grapevine from southern Moravia, a grapevine believed to be of the Olsava variety and another believed to be of the Palava variety, could not be identified by our database searches. Mendelian segregation analysis in the Cabernet Sauvignon family (Cabernet franc × Sauvignon blanc) revealed scu05vv discrepancies (CS 175/175, CF 192/192, SB 166/175).

We were unable to find verified hexaplex microsatellite data for all of the grapevine varieties considered in this work in the literature [35, 36, 38, 39]. Consequently, direct comparisons to previously reported results could only be made for the VrZAG47/VVMD27 locus. We compared our

Table 2. Population genetics parameters of multiplex loci

	Genotypes	Alleles	Effective alleles	Estimated frequency of null alleles	Shannon's information index	Observed heterozygosity	Expected heterozygosity	Unbiased expected heterozygosity	Fixation index
VrZAG47	28.00	17.00	5.33	0.18	2.03	0.49	0.81	0.82	0.40
VVIi51	33.00	13.00	5.98	0.00	2.06	0.86	0.83	0.84	−0.03
scu08vv	4.00	3.00	1.23	0.02	0.39	0.16	0.18	0.19	0.16
VrZAG83	18.00	8.00	5.16	0.04	1.74	0.73	0.81	0.81	0.09
scu05vv	23.00	15.00	3.02	0.22	1.64	0.30	0.67	0.67	0.55
VVMD17	13.00	7.00	3.01	0.01	1.31	0.66	0.67	0.67	0.02

The negative fixation index of VVIi51 may indicate an excess of heterozygosity due to negative assortative mating, or selection for heterozygotes.

data for 35 INRA varieties to data presented by ten research groups, which are here named for their principal investigators: Ibanez, Grando, Botta, Peterlunger, Regner, Meredith, Monteiro, Crespan, This, and Maul. To compare our results to those of Ibanez and Grando, we had to add 14 bp to our results to compensate for the use of different primers; for Borra and Peterlunger, we added 15 bp, for Regner, Meredith, Monteiro, Crespan, and This we added 16 bp, and for Maul we added 17 bp [21]. Our genotyping data agreed with the consensus genotypes with the following exceptions. In the cases of Furmint, Muscat a petit grains, Romorantin, and Rondinella, our hexaplex missed allele 163; in the case of Vialla it missed allele 171; in the case of Kober 5BB, the consensus alleles were 175/194 while our data suggested 159/194; and in the case of Jacquez the consensus alleles were 163/173 while our data suggested 169/169.

Sequencing of the primer-binding regions in Furmint, Muscat a petit grains, and Rondinella revealed no mismatch with the forward primer (ACg gTg TgC TCT CAT TgT CAT Tg) while five nucleotides were mismatched in the reverse primer-binding region (gT**A** C**CA** g**AT** CTg AAT ACA TCC gTA Ag**Y** instead of gTT CTT ggT CTg AAT ACA TCC gTA AgT).

The population genetics parameters of our microsatellites [51] calculated from our data are summarized in the Table 2. The most discriminatory locus in our set is VVIi51, with 33 genotypes and an observed heterozygosity of 0.86, while the least discriminatory locus was scu08vv, with four genotypes and an observed heterozygosity of 0.16.

3.3 Robustness study

The smallest quantity of template that reproducibly yielded a full profile with peak heights above the analysis threshold of 80 RFU while also achieving satisfactory peak height balance (PHB = height of the first peak/height of the second peak $\geq$ 0.45) was 0.13 ng. Template amounts between 0.25 and 1 ng per reaction, corresponding to 250 and 1000 copies of the diploid grapevine genome, respectively, produced high peaks with PHB $\geq$ 0.6 without extra alleles and were chosen as the standard inputs.

Although several brands of Taq polymerase produced a full DNA profile under our standard conditions (data not shown), we used the AmpliTaq Gold polymerase (Life Technologies) with built-in hot-start [52], which has a wide efficient working concentration range (concentrations of 30 to 90 mU per reaction yielded full profiles without allele drop-in or drop-out). Our multiplex was quite insensitive to variation in primer concentrations: full profiles were obtained when using anywhere between half and twice the specified values. Reducing the magnesium chloride concentration to 1.0 mM reduced the peak heights of the VVIi51 locus such that they became close to the sensitivity threshold but left other loci unaffected, while increasing the magnesium concentration up to 2.0 mM did not produce any artefacts. Therefore, magnesium concentrations between 1.5 and 2.0 mM should be used in the assay. The reaction volume cannot be reduced below 6 μL; below this level, it is likely that evaporation significantly reduces the PCR yield and may cause allele drop-out. Conversely, increasing the volume to 15 μL did not affect the results in any direction. Our multiplex protocol is not sensitive to thermocycler well-to-well temperature inconsistencies because the temperatures used in the touch-down cycling program could be increased or lowered by 2°C without changing the obtained profiles.

To assess the scope for using our protocol to perform DNA profiling of grapevine mixtures in cuvées or other mixed samples, experiments were performed using the following ratios of Portugieser blau and Castel 216-3 DNA templates: 1:0, 9:1, 4:1, 2:1, 1:1. 1:2, 1:4, 1:9, 0:1. Duplicate PCR amplifications were performed for each ratio, using a total of 1 ng of template DNA per 10 μL of amplification reaction mixture. Full profiles for both major and minor fractions were obtained when the ratio of major to minor component was 4:1 or less.

3.4 Compromised samples

To assess the ability of our multiplex to analyze samples that may be compromised (by fragmentation, low DNA concentrations, and/or excess quantities of nongrapevine DNA), we performed profiling experiments using wine, a rotten leaf

sample from a wild-grown grapevine, and a several hundred year-old grapevine pip from an archeological excavation.

An inexpensive varietal wine (40 mL) whose label indicated that it was made from Gruener Veltliner grapes was found to possess some Gruener Veltliner alleles (VrZAG47 173; VVIi51 262; scu08vv 181/181; VrZAG83 214/220; scu05vv 163/166; and VVMD17 247/249), but it lacked the Gruener Veltliner allele VVIi51 264 and also contained several alleles associated with other varieties (VrZAG47 163/165/167/171/175; VVIi51 244/250; VrZAG83 208/210; scu05vv 161/181/190/192; and VVMD17 237/243/248).

Analysis of pip sample ČB 1/20 20005 from the Early Iron Age České Budějovice excavation site yielded a full DNA profile (VrZAG47 169/173; VVIi51 264/266; scu08vv 181/188; VrZAG83 208/220; scu05vv 163/163; and VVMD17 239/249).

Analysis of a leaf sample from a wild grapevine found in Czech Republic [53] that had been almost completely destroyed by a mold infection during transport to the laboratory yielded the profile VrZAG47 173/173, VVIi51 266/266, scu08vv 181/181, VrZAG83 210/210, scu05vv 192/200, and VVMD17 239/248.

4 Discussion

In this work, we developed a new multiplexed profiling profile based on the use of six microsatellite primer pairs for simultaneous DNA profiling of six *Vitis spp*. loci.

4.1 Specificity and population study

Because of the "random amplification of polymorphic DNA effect" [54, 55] that occurs when two primers designed for different microsatellites unexpectedly bind within the productive length of the PCR amplification, multiplexed PCR experiments are more likely to generate artefact peaks than singleplex experiments. However, we did not observe any discrepancies between our singleplex and multiplex results, or any indication of L1/L2 meristem cell layer chimerism [56,57].

Our analysis missed allele 163 at the VVMD27 locus (which is also known as VrZAg47) in the Furmint, Muscat a petit grains, Romorantin, and Rondinella varieties, and allele 171 in the Vialla variety. Sequencing of available DNA sequences (Furmint, Muscat a petit grains, and Rondinella) revealed five mismatches between the reverse primer and its binding sequence: gT**A** C**CA** g**A**T CTg AAT ACA TCC gTA Ag**Y** instead of gTT CTT ggT CTg AAT ACA TCC gTA AgT. The nucleotide mismatches at the 5′ end of the reverse primer exist because the primer was originally designed to amplify sequences from the *Vitis riparia* genome rather than *Vitis vinifera* [34]. This problem has already been addressed by Laucou et al. [23], who developed a new *V. vinifera*-specific primer (gTA CCA gAT CTg AAT ACA TCC gTA Ag**T)**. However, even the new Laucou primer would not enable amplification of the second allele in Furmint, Muscat a petit grains, and Rondinella using a proofreading-free polymerase because of the polymorphism at the primer's 3' end. We were unable to further investigate discrepancies in the results for Romorantin (missed allele 163), Vialla (missed allele 171), Cabernet franc (missed scu05vv allele 175), Kober 5BB (consensus 175/194, our data 159/194), and Jacquez (consensus 163/173, our data 169/169) because the corresponding biological material had been wholly consumed. While the missed alleles were probably due to the primer's 3' end polymorphism as discussed above, the results for the Kober 5BB and Jacquez lack a similarly obvious biological explanation and are perhaps more likely to be the result of labeling errors. A comparison of our VrZAG47 data to the European Vitis Database (http://www.eu-vitis.de/index.php) and Vitis International Variety Catalogue (VIVC, http://www.vivc.de/index.php) revealed additional differences (Supporting Information 1C), suggesting that results obtained with our hexaplex should be used with caution if the aim is to integrate them with external DNA profile databases.

The cross-reactivity of our primers with *P. quinquefolia* L. should not cause any difficulties in oenological interpretations of grapevine DNA profiles due to the distinctive habitats of these two species.

Identical DNA profiles were obtained for some varieties in our grapevine collection (Aivas/Kodrianka/BV18-109, Muscotaly sarga/Muscat Ottonel, and Pinot noir/Pinot noir praecox) when analyzed with the hexaplex, demonstrating the limitations imposed by using only six microsatellites for identification and/or kinship assignment. This lack of discrimination is unsurprising given that the Combined DNA index system for human identification uses 13 microsatellite loci [58] and at least nine microsatellites are recommended for kinship studies in grapevine [50, 59]. A panel of 45 multiplex microsatellite PCRs was recently suggested by Zarouri et al. [60] but needs validation.

In this work, we have used the amplicon length nomenclature and the $n + x$ [21] nomenclature for alleles. The preferred nomenclature, based on the repeat number [61], necessitates the sequencing of all alleles, which would have been beyond the scope of this project.

4.2 Robustness

The optimal conditions for our multiplex are specified in Table 1. However, there is a reasonably wide experimental window for the PCR parameters in the assay: the template amount can be varied between 0.25 ng and 1 ng/reaction; the $MgCl_2$ concentration between 1.5 and 2.0 mM; the primer mix concentration between 0.5 and 2.0 times the standard values; the touch-down annealing temperature cycling can be varied by ± 2°C relative to the standard setting of 59°C to 50°C; and the reaction volume can be reduced to as little as 6 μL. Our multiplex has higher level of multiplexing (6 vs. 3) than that of Alba et al. [62] but a lower level than that of Migliaro et al. [63], which uses a genetically engineered polymerase from the KAPA3g Plant PCR kit (Kapa

Biosystems, USA) and was only tested on 12 samples. Additionally, our protocol is more robust and has a better detection limit (28 PCR cycles compared to 40 cycles) than other published grapevine multiplexes [64, 65]. It can distinguish mixtures of two samples where the minor fraction accounts for at least a quarter of the total grapevine DNA content. Discrimination when the minor fraction accounts for only a tenth of the target genetic material is routinely achieved with human microsatellites but could not be achieved in this case due to comparatively high levels of stutter (Table 1). This can be attributed to the use of dinucleotide instead of tetranucleotide repeats. Tetranucleotide repeats are scarce in the grapevine genome but a list of potentially useful microsatellites with a core length of four nucleotides was recently identified [46, 66]. While SNPs [67, 68] avoid stuttering problems and do not require an allelic ladder for standardized genotyping, we concur with Santos [46] that long core repeat microsatellites and a forensic standardization approach [49] to their profiling still has considerable unexploited potential in grapevine genetics. It remains to be seen whether massively parallel sequencing-based technologies [69] offer sufficient advantages over STRs.

4.3 Compromised samples

We successfully obtained data from all of the targeted multiplex loci in our experiments using the moldy sample, the archeological sample, and the wine sample. Some extra alleles were identified in the wine sample, suggesting that varieties outside of our grapevine collection were used for winemaking. In many countries, a wine can be designated monovarietal wines if up to 15% of its grape content is from varieties other than that on the label.

The yield and quality of the grapevine DNA obtained from wine can vary widely depending on the grape variety that was used to make the wine and the vinification process [27]. Similarly, the yield and quality of DNA from ancient samples depends heavily on the temperatures, moisture, and bacteria to which the sample has been exposed. Dedicated studies are needed to determine whether our multiplex can produce reliable results when challenged with such variation, and whether it can compete with alternative emerging technologies for wine analysis such as electronic noses [70].

5 Concluding remarks

We designed and validated a PCR multiplex of six STRs for profiling grapevine DNA templates. Our mixture of fluorescently labeled primers and PCR reagents for the VrZAG47, VVIi51, scu08vv, VrZAG83, scu05vv, and VVMD17 loci uses 0.5 ng of *Vitis* DNA per multiplex reaction with 28 touchdown PCR cycles, and can tolerate small deviations from the specified PCR protocol. We successfully genotyped 92 samples, including mixtures, archeological pips, and wine. The development of this protocol represents a further step toward the establishment of a reliable and general *Vitis* multiplex PCR procedure for forensic authentication of grapevines, must, and wine.

This work was supported by grants LO1304, LM2010004, and TE02000058. We thank Drs. S. Lalet, N. Goto-Yamamoto, M. Galimberti, Nikos, J. Beneš, O. Jandurová, P. Kočár, P. Maděra, M. Michlovský, P. Pavloušek, J. Stávek, and M. Pražák for grapevine samples. We also thank Profs. L. Bocák, M. Navrátil, and Dr. M. Valárik (Olomouc, the Czech Republic) for generously providing access to instruments, and Veronika Ťoupalíková, MSc. for technical help in addressing reviewers' comments.

Authors have no conflict of interest to declare.

6 References

[1] Garcia-Beneytez, E., Moreno-Arribas, M. V., Borrego, J., Polo, M. C., Ibanez, J., *J. Agric. Food Chem.* 2002, *50*, 6090–6096.

[2] Siret, R., Gigaud, O., Rosec, J. P., This, P., *J. Agric. Food Chem.* 2002, *50*, 3822–3827.

[3] Thomas, M. R., Cain, P., Scott, N. S., *Plant Mol. Biol.* 1994, *25*, 939–949.

[4] Regner, F., Wiedeck, E., Stadlbauer, A., *Vitis* 2000, *39*, 103–107.

[5] Bodor, P., Szoke, A., Toth-Lencses, K., Veres, A., Deak, T., Kozma, P., Bisztray, G. D., Kiss, E., *Biotechnol. Biotechnol. Equip.* 2014, *28*, 14–20.

[6] Kayesh, E., Bilkish, N., Liu, G. S., Chen, W., Leng, X. P., Fang, J. G., *Genet. Mol. Res.* 2014, *13*, 2220–2230.

[7] Mancuso, S., *Vitis* 2001, *40*, 59–63.

[8] De, L. G., Chipashvili, R., Failla, O., Maghradze, D., *BMC Plant Biol.* 2015, *15*, 154, doi: 10.1186/s.12870-015-0510-9.

[9] Hyma, K. E., Barba, P., Wang, M., Londo, J. P., Acharya, C. B., Mitchell, S. E., Sun, Q., Reisch, B., Cadle-Davidson, L., *Plos One* 2015, *10*, e0134880.

[10] Moreno-Arribas, M. V., Cabello, F., Polo, M. C., Martin-Alvarez, P. J., Pueyo, E., *J. Agric. Food Chem.* 1999, *47*, 114–120.

[11] Garcia-Beneytez, E., Cabello, F., Revilla, E., *J. Agric. Food Chem.* 2003, *51*, 5622–5629.

[12] Vasconcellos, A. M. P., Chaves das Neves, H. J., *J. Agric. Food Chem.* 1989, *37*, 931–937.

[13] Munoz-Organero, G., Ortiz, J. M., *Riv. Vitic. Enol.* 2004, *3*, 55–64.

[14] Makris, D. P., Kallithraka, S., Mamalos, A., *Talanta* 2006, *70*, 1143–1152.

[15] Almeida, C. M., Vasconcelos, M. T., *J. Agric. Food Chem.* 2003, *51*, 4788–4798.

[16] Imazio, S., Labra, M., Grassi, F., Winfield, M., Bardini, M., Scienza, A., *Plant Breeding* 2002, *121*, 531–535.

[17] Brettell, T. A., Butler, J. M., Almirall, J. R., *Anal. Chem.* 2007, *79*, 4365–4384.

[18] Lijavetzky, D., Cabezas, J. A., Ibanez, A., Rodriguez, V., Martinez-Zapater, J. M., *BMC Genomics* 2007, *8*, 424, doi: 10.1186/1471-2164-8-424.

[19] Pindo, M., Vezzulli, S., Coppola, G., Cartwright, D. A., Zharkikh, A., Velasco, R., Troggio, M., *BMC Plant Biol.* 2008, *8*, 12, doi: 10.1186/1471-2229-8-12.

[20] Santiago, J. L., Gonzalez, I., Gago, P., onso-Villaverde, V., Boso, S., Martinez, M. C., *Austral. J. Grape Wine Res.* 2008, *14*, 223–229.

[21] This, P., Jung, A., Boccacci, P., Borrego, J., Botta, R., Costantini, L., Crespan, M., Dangl, G. S., Eisenheld, C., Ferreira-Monteiro, F., Grando, S., Ibanez, J., Lacombe, T., Laucou, V., Magalhaes, R., Meredith, C. P., Milani, N., Peterlunger, E., Regner, F., Zulini, L., Maul, E., *Theor. Appl. Genet.* 2004, *109*, 1448–1458.

[22] Dangl, G. S., Mendum, M. L., Yang, J., Walker, M. A., Preece, J. E., *Ecol. Evol.* 2015, *5*, 5671–5684.

[23] Laucou, V., Lacombe, T., Dechesne, F., Siret, R., Bruno, J. P., Dessup, M., Dessup, T., Ortigosa, P., Parra, P., Roux, C., Santoni, S., Vares, D., Peros, J. P., Boursiquot, J. M., This, P., *Theor. Appl. Genet.* 2011, *122*, 1233–1245.

[24] Bacilieri, R., Lacombe, T., Le, C. L., Di Vecchi-Staraz, M., Laucou, V., Genna, B., Peros, J. P., This, P., Boursiquot, J. M., *BMC Plant Biol.* 2013, *13*, 25, doi: 10.1186/1471-2229-13-25.

[25] Jacquat, C., Martinoli, D., *Veg. Hist. Archaeobot.* 1999, *8*, 25–30.

[26] Faria, M. A., Nunes, E., Oliveira, M. B. P. P., *Eur. Food. Res. Technol.* 2008, *227*, 845–850.

[27] Drabek, J., Stavek, J., Jaluvkova, M., Jurcek, T., Frebort, I., *Eur. Food. Res. Technol.* 2008, *226*, 491–497.

[28] Lefort, F., Douglas, G. C., *Ann. Forest Sci.* 1999, *56*, 259–263.

[29] Chomczynski, P., Mackey, K., Drews, R., Wilfinger, W., *Biotechniques* 1997, *22*, 550–553.

[30] Valsesia, G., Gobbin, D., Patocchi, A., Vecchione, A., Pertot, I., Gessler, C., *Phytopathology* 2005, *95*, 672–678.

[31] Butler, J. M., Shen, Y., McCord, B. R., *J. Forensic Sci.* 2003, *48*, 1054–1064.

[32] Drabek, J., Chung, D., Butler, J. M., McCord, B. R., *J. Forensic Sci.* 2004, *49*, 859–860.

[33] Butler, J. M., *Methods Mol. Biol.* 2005, *297*, 53–66.

[34] Sefc, K. M., Regner, F., Turetschek, E., Glossl, J., Steinkellner, H., *Genome* 1999, *42*, 367–373.

[35] Merdinoglu, D., Butterlin, G., Bevilacqua, L., Chiquet, V., Adam-Blondon, A. F., Decroocq, S., *Mol. Breeding* 2005, *15*, 349–366.

[36] Sefc, K. M., Regner, F., Turetschek, E., Glossl, J., Steinkellner, H., *Genome* 1999, *42*, 367–373.

[37] Adam-Blondon, A. F., Roux, C., Claux, D., Butterlin, G., Merdinoglu, D., This, P., *Theor. Appl. Genet.* 2004, *109*, 1017–1027.

[38] Bowers, J. E., Dangl, G. S., Meredith, C. P., *Am. J. Enol. Vitic.* 1999, *50*, 243–246.

[39] Scott, K. D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E. M., Lee, L. S., Henry, R. J., *Theor. Appl. Genet.* 2000, *100*, 723–726.

[40] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., *J. Mol. Biol.* 1990, *215*, 403–410.

[41] Pinto, F. L., Svensson, H., Lindblad, P., *BMC Biotechnol.* 2006, *6*, 31, doi: 10.1186/1472-6750-6-31.

[42] Brownstein, M. J., Carpten, J. D., Smith, J. R., *Biotechniques* 1996, *20*, 1004–1010.

[43] Micka, K. A., Sprecher, C. J., Lins, A. M., Theisen Comey, C., Koons, B. W., Crouse, C., Endean, D., Pirelli, K., Lee, S. B., Duda, N., Ma, M., Schumm, J. W., *J. Forensic Sci.* 1996, *41*, 582–590.

[44] Haberl, M., Tautz, D., *Mol. Ecol.* 1999, *8*, 1347–1349.

[45] Butler, J. M., McCord, B. R., Jung, J. M., Lee, J. A., Budowle, B., Allen, R. O., *Electrophoresis* 1995, *16*, 974–980.

[46] Santos, S., Oliveira, M., Amorim, A., van, A. B., *Electrophoresis* 2014, *35*(21–22), 3201–3207.

[47] Liu, K. J., Muse, S. V., *Bioinformatics* 2005, *21*, 2128–2129.

[48] Peakall, R., Smouse, P. E., *Mol. Ecol. Notes* 2006, *6*, 288–295.

[49] Linacre, A., Gusmao, L., Hecht, W., Hellmann, A. P., Mayr, W. R., Parson, W., Prinz, M., Schneider, P. M., Morling, N., *Forensic Sci. Int. Genet.* 2011, *5*, 501–505.

[50] Boursiquot, J.-M., Lacombe, T., Laucou, V., Julliard, S., Perrin, F.-X., Lanier, N., Legrand, D., Meredith, C., This, P., *Austral. J. Grape Wine Res.* 2009, *15*(2), 144–155.

[51] Weir, B. S., *Genetic Data Analysis II*, Sinauer Associates, Inc., Sunderland, MA 1996.

[52] Moretti, T., Koons, B., Budowle, B., *Biotechniques* 1998, *25*, 716–722.

[53] Madera, P., Martinkova, M., *J. Forest Sci.* 2002, *48*, 482–485.

[54] Deng, H. W., Zhou, Y., Recker, R. R., Johnson, M. L., Li, J., *Biotechniques* 2000, *29*, 298–304, 307-8.

[55] Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., Tingey, S. V., *Nucleic Acids Res.* 1990, *18*, 6531–6535.

[56] Franks, T., Botta, R., Thomas, M. R., Franks, J., *Theor. Appl. Genet.* 2002, *104*, 192–199.

[57] Hocquigny, S., Pelsy, F., Dumas, V., Kindt, S., Heloir, M. C., Merdinoglu, D., *Genome* 2004, *47*, 579–589.

[58] Miller, K. W. P., Brown, B. L., Budowle, B., *Progress Forensic Genet. 9* 2003, *1239*, 617–620.

[59] Laucou, V., Boursiquot, J. M., Lacombe, T., Bordenave, L., Decroocq, S., Ollat, N., *Vitis* 2008, *47*, 163–167.

[60] Zarouri, B., Maria Vargas, A., Gaforio, L., Aller, M., Teresa de Andres, M., Antonio Cabezas, J., *Tree Genet. Genomes* 2015, *11*(17), doi: 10.1007/s11295-015-0843-4.

[61] Olaisen, B., Bar, W., Brinkmann, B., Budowle, B., Carracedo, A., Gill, P., Lincoln, P., Mayr, W. R., Rand, S., *Vox Sang.* 1998, *74*, 61–63.

[62] Alba, V., Bergamini, C., Genghi, R., Gasparro, M., Perniola, R., Antonacci, D., *Mol. Biotechnol.* 2015, *57*, 709–719.

[63] Migliaro, D., Morreale, G., Gardiman, M., Landolfo, S., Crespan, M., *Plant Genet. Resour.* 2013,*11*(2), 182–185.

[64] Frei, A., Baumgartner, D., Frey, J. E., Gafner, J., *Bot. Helv.* 2004, *114*, 169–179.

[65] Rodriguez-Plaza, P., Gonzalez, R., Moreno-Arribas, M. V., Polo, M. C., Bravo, G., Martinez-Zapater, J. M., Martinez, M. C., Cifuentes, A., *Eur. Food. Res. Technol.* 2006, *223*, 625–631.

[66] Cipriani, G., Marrazzo, M. T., Di, G. G., Pfeiffer, A., Morgante, M., Testolin, R., *BMC Plant Biol*. 2008, *8*, 127, doi: 10.1186/1471-2229-8-12.

[67] Cabezas, J. A., Ibanez, J., Lijavetzky, D., Velez, D., Bravo, G., Rodriguez, V., Carreno, I., Jermakow, A. M., Carreno, J., Ruiz-Garcia, L., Thomas, M. R., Martinez-Zapater, J. M., *BMC Plant Biol*. 2011, *11*(153), doi: 10.1186/1471-2229-11-153.

[68] Mercati, F., De Lorenzis, G., Brancadoro, L., Lupini, A., Abenavoli, M. R., Barbagallo, M. G., Di Lorenzo, R., Scienza, A., Sunseri, F., *Tree Genet. Genomes* 2016, *12*(59), doi: 10.1007/s11295-016-1021-z.

[69] Yang, S. S., Fresnedo-Ramirez, J., Wang, M. H., Cote, L., Schweitzer, P., Barba, P., Takacs, E. M., Clark, M., Luby, J., Manns, D. C., Sacks, G., Mansfield, A. K., Londo, J., Fennell, A., Gadoury, D., Reisch, B., Cadle-Davidson, L., Sun, Q., *Hort. Res.* 2016, *3*(16002), doi: 10.1038/hortres.2016.2. eCollection 2016

[70] Cozzolino, D., Smyth, H. E., Cynkar, W., Dambergs, R. G., Gishen, M., *Talanta* 2005, *68*, 382–387.

