

Microsatellite-based DNA Fingerprinting of Selected Grapevine Cultivars

C.J. Van Heerden¹, P. Burger², R. Prins^{*3}

(1) Department of Genetics, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

(2) ARC Infruitec-Nietvoorbij, Private Bag X5026, Stellenbosch 7599, South Africa

(3) CenGen (Pty) Ltd, 78 Fairbairn Street, Worcester 6850, South Africa

Submitted for publication: April 2017

Accepted for publication: August 2017

Keywords: SSR, *Vitis vinifera*, pedigree, cultivar identification, fungal resistance, conserved haplotype

Cultivar identification by ampelography is often difficult and is sensitive to environmental conditions, thus it can be problematic to distinguish between closely related cultivars. DNA fingerprinting offers an alternative method that is not influenced by the environment and is relatively easy to perform. However, discriminating between closely related individuals can be problematic if inadequate or insufficient markers are used. Following the estimation of null allele frequencies, an initial set of 35 microsatellite markers was reduced to 20 to generate unique DNA fingerprints for the majority of 111 different grape cultivars and breeding lines. Molecular evidence was utilised to evaluate the accuracy of the reported pedigrees for several cultivars bred in South Africa (SA). The use of markers linked to known downy mildew and powdery mildew resistance loci (*Rpv3* and *Ren3*) provided information regarding the frequency of these resistance loci in the breeding material analysed.

INTRODUCTION

Traditionally, ampelography has been used for cultivar determination in grapevines, but since it is environmentally sensitive it can lead to erroneous identification, especially in artificial conditions, and it is not useful for clones (Vignani *et al.*, 2002; This *et al.*, 2004). Molecular markers designed around simple sequence repeats (SSRs, also known as microsatellites) offer a higher differentiating power and, while they can be difficult and expensive to develop initially, SSRs have proven to be a robust and effective tool for identification, parentage assignment and genetic mapping in *Vitis* (Dangl *et al.*, 2001; Akkak *et al.*, 2007; Bellin *et al.*, 2009). In recent years, the use of automated fluorescence technologies have made the simultaneous analysis of several loci feasible and have decreased the cost per data point (Adam-Blondon *et al.*, 2004; Merdinoglu *et al.*, 2005; Bautista *et al.*, 2008; Ibáñez *et al.*, 2009).

More recently, single nucleotide polymorphisms (SNPs) have also been used to generate DNA fingerprints (Nybom *et al.*, 2014), determine genetic diversity and population structure (Emanuelli *et al.*, 2013), generate genetic maps (Troggio *et al.*, 2007), and identify candidate genes in association studies (Emanuelli *et al.*, 2010). In general, SNPs are bi-allelic, with a maximum of four alleles possible

in outbreeding populations. This level of polymorphism is lower than what is regularly found for SSR markers, making individual SNP markers less informative than most SSR markers. However, the use of high-density arrays means that very large numbers of SNPs can be genotyped at relatively low cost and, by using large numbers of SNPs, the same overall level of variation can be observed. Emanuelli *et al.* (2013) used both SSRs and SNPs to investigate genetic diversity and population structure in grapes and found that they yielded similar results, but that the greater information content of SSRs made them more suitable for core collection construction.

When dealing with the identification of grape cultivars and parentage assessments, several problematic issues must be considered. Grapevines are propagated by cuttings, and therefore all accessions of a specific cultivar are expected to be genetically identical. However, somatic mutations could result in changes to the DNA fingerprint of an individual plant without necessarily changing the phenotype (Martínez *et al.*, 2006). It is even more probable that somatic mutations will cause changes in the phenotype without resulting in changes to the DNA fingerprint (Imazio *et al.*, 2002). The latter often results in the establishment of a “clone” of the

*Corresponding author: E-mail address: cjvh@sun.ac.za

Acknowledgements: This work is based on research supported in part by the National Research Foundation (NRF) of South Africa (THRIP grants 70082, 72059 and 75125), including incentive funding for an NRF-rated scientist (Renée Prins; 85943). The NRF is thanked for funding the equipment, which is based at the Central Analytical Facilities of Stellenbosch University (UID 65258). The Deciduous Fruit Producers' Trust (DFPT), the South African Table Grape Industry Partnership (SATI) and the Department of Science and Technology (DST) are also thanked for their financial support. Danie van Schalkwyk is thanked for supplying additional plant material. Debbie Snyman and Lizaan Rademeyer (CenGen) are thanked for technical assistance. Paul Shaw (James Hutton Institute, UK) is thanked for his support in setting up Helium

original cultivar. For instance, Muscat of Alexandria Red is a red clone of the white cultivar Muscat of Alexandria (*Vitis International Variety Catalogue* database; <http://www.vivc.de/>), while Muscat of Alexandria C982 has a more elongated white berry. Furthermore, the long and complex history and distribution of grape cultivation has led to the widespread use of synonyms (more than one name for a given cultivar) and homonyms (the same name given to different cultivars in different regions). This situation is further complicated by the possibility that some cultivars may not have a monoclonal origin, as in the case of Sangiovese, of which some clones differ from other clones at one of seven SSR loci tested (Vignani *et al.*, 1996). Other complicating factors, including the possibility of chimeras (Pelsy *et al.*, 2010) and PCR artefacts, add complexity when analysing microsatellite data. Multiple studies have reported the presence of null alleles (Bautista *et al.*, 2008; Cipriani *et al.*, 2010; Laucou *et al.*, 2011), which can lead to incorrect parentage exclusions (Dakin & Avise, 2004). Null alleles are estimated by detecting a lower than expected number of heterozygous individuals in a population (Van Treuren, 1998), but low-frequency null alleles in small sets of samples can go undetected.

Historically, all *V. vinifera* cultivars were susceptible to *Plasmopara viticola* (Staudt & Kassemeyer, 1995; Cadle-Davidson, 2008), and most were susceptible to *Erysiphe necator* (Alleweldt & Possingham, 1988), the pathogens causing downy mildew and powdery mildew in grapes. Resistance loci against both pathogens have been introduced from other *Vitis* species to create resistant cultivars like Regent (Eibach & Töpfer, 2003), and the chromosomal location of these loci have been determined by various mapping studies (Fischer *et al.*, 2004; Hoffmann *et al.*, 2008; Zyprian *et al.*, 2016) to allow for marker-assisted selection (MAS) when these cultivars are used as donor sources to breed resistant cultivars.

This study aimed to implement the use of a set of fingerprint and gene-associated markers for identification and discrimination purposes, and to determine the disease resistance loci status of grapevine cultivars frequently used in table grape breeding in SA. Where possible, the parentage of cultivars was also analysed.

MATERIALS AND METHODS

Plant material and DNA extraction

One hundred and twenty-five accessions were selected from the grape gene bank collection at the Agricultural Research Council (ARC) Infruitec-Nietvoorbij, for which some pedigree data were known (VIVC, <http://www.vivc.de/>; P. Burger, personal communication, Table S1, Online Resource 1). A set of 13 cultivars was submitted more than once to act as internal controls. Thirteen of the 125 cultivars were also known to be resistant to downy mildew and/or powdery mildew.

DNA extractions from leaves were performed using the NucleoSpin® Plant II 96 kit (Macherey Nagel, Duren, Germany; Telfer *et al.*, 2013) implemented on a Tecan Genesis 200RMP (Tecan, Maennedorf, Switzerland) liquid-handling robot. In short, two 1 cm² fresh or frozen leaves were homogenized in a 2 ml microcentrifuge tube using the TissueLyser (Qiagen, Venlo, Netherlands). Homogenisation

was performed twice in the lysis buffer (supplied with the kit) for 30 seconds at 30 Hz using 5 mm stainless steel beads. The homogenised tissue was lysed for 30 min at 56°C and then centrifuged for 10 min at 14 000 g to clear the lysate of as much particulate matter as possible. The tubes were then opened and placed on the Genesis for processing. The script follows the standard vacuum-based protocol supplied by the kit manufacturer. After extraction, DNA was quantified using a NanoDrop (Thermo Fisher Scientific) and the DNA concentrations were normalised to 30 ng/μl.

Marker selection

An initial set of 35 SSR markers (Table S2, Online Resource 2) was selected for genotyping (Bowers *et al.*, 1999; Sefc *et al.*, 1999; Di Gaspero *et al.*, 2005; Lopes *et al.*, 2006; Martínez *et al.*, 2006; Vouillamoz & Grando, 2006; Vouillamoz *et al.*, 2007; Van Heerden *et al.*, 2014). This set included three of the set of markers proposed by This *et al.* (2004) for DNA fingerprinting (VVMD7, VrZAG62 and VrZAG79), as well as marker VrZAG47, which amplifies the same locus as VVMD27 (Dalbó *et al.*, 2000). Several markers from chromosomes 15 and 18 were included, as they were previously implicated in resistance to powdery mildew and downy mildew (Fischer *et al.*, 2004; Van Heerden *et al.*, 2014). Multiplex PCR reactions were performed using 0.75 U Supertherm *Taq* (Qiagen) and the buffer supplied with the enzyme. MgCl₂ and primer concentrations were empirically determined and modified to yield reproducible, clear data. Amplifications were done using 30 ng of genomic DNA as template on an ABI 9700 GeneAmp PCR instrument. The 35 markers were arranged in seven multiplexes and amplified in optimised multiplex PCR reactions.

Data scoring and analysis

Data analysis was performed using GeneMapper® V4. The following procedure was used to ensure accurate scoring of the data: a) a peak was only scored if the data had a relative fluorescence intensity of more than 500, or b) if a peak of lower intensity was repeatedly seen in various samples while it was also absent in others. In the case of uncertainty, samples were repeated. Several samples were also submitted in duplicate, and the scores for these duplicate samples were compared to verify that scoring was consistent. The table of results was exported from GeneMapper® V4 as a comma-separated values (csv) file and imported into Microsoft Excel for further manipulation and analysis.

GenAIEx 6.5 (Peakall & Smouse, 2012) was used to identify identical genotypes in the dataset. Identical genotypes that confirmed known duplicate samples were removed so that each of these genotypes was only represented once in the dataset during subsequent data analysis, while samples with identical genotypes but different names were retained. Identity V4 (Wagner & Sefc, 1999) was then used to identify markers with a high probability of null alleles. Where parent-offspring groups were available, the inheritance of markers was manually inspected to check for the presence of rare null alleles. Markers with observed or expected null allele frequencies higher than 0.05 were removed prior to further analysis. GenAIEx 6.5 was also used in a second round of analysis to determine the allele frequency, deviation from the

Hardy-Weinberg equilibrium and the inbreeding coefficient (F_{is}), and the probability of identity was calculated. GenAIEx 6.5 was also used to identify identical genotypes. Markers that deviated significantly from the expected Hardy-Weinberg equilibrium were removed, and only the genotypes for the remaining subset of 20 markers were used in Identity V4 to compile a list of potential parent-offspring groups and establish a final DNA fingerprint.

The *Rpv3* downy mildew (UDV108, VVIN16-cjvh) and *Ren3* powdery mildew (UDV116) resistance-linked marker data (Van Heerden *et al.*, 2014) were investigated to determine if common alleles could be identified in the following known downy mildew- and/or powdery mildew-resistant cultivars: Regent (*Rpv3*, *Ren3*, downy and powdery mildew resistant; Fischer *et al.*, 2004; Van Heerden *et al.*, 2014), Villard Blanc (*Rpv3*, *Ren3*, *Ren8*, downy and powdery mildew resistant; Zyprian *et al.*, 2016), Moldova (downy and powdery mildew resistant; Eibach, 1994; Ulea *et al.*, 2012), Kodryanka (downy and powdery mildew resistant; Pavloušek, 2006; 2007), Ajvaz (downy and powdery mildew resistant; Pavloušek, 2006; 2007), Marroo Seedless (downy mildew resistant, susceptible to powdery mildew; Clingeffer & Possingham, 1988), Chambourcin (*Rpv3*, downy and powdery mildew resistant; Eibach, 1994; Di Gaspero *et al.*, 2012), Sarolta (downy mildew resistant, susceptible to powdery mildew; Pernes, 2004), Poeloeskei Muskotaly (downy and powdery mildew resistant; Pavloušek, 2006; 2007), Kishmish Vatkana (*Ren1*, powdery mildew resistant; Hoffmann *et al.*, 2008), Isabella (downy and powdery mildew resistant, VIVC; Eibach, 1994), Vanessa (powdery mildew resistant; Eibach, 1994) and Seneca (powdery mildew resistant; Eibach, 1994).

Without using any of the marker data generated, previously published pedigree data (VIVC; <http://www.vivc.de/>; P Burger, personal communication; Table S1, Online Resource 1) for the study samples was graphically summarised using Helium software (Shaw *et al.*, 2014) to create a clearer picture of the relationships in this dataset. Reported ancestors not genotyped in this study were also included in instances where these individuals were needed to link related individuals (for instance, Seibel 6468 is included since it represents the relationship between Villard Blanc and Chambourcin).

RESULTS

The Helium-generated graphic depictions revealed a core family consisting of 89 out of the 125 individuals, which is shown to create a frame of reference for the results reported below (Fig. S1, Online Resource 3).

A comparison of the fingerprints obtained with 35 SSRs confirmed 11 out of 13 duplicates in the 125 samples analysed (Table S3, Online Resource 4). The two duplicates of Colombar differed from one another at three loci (VrZAG47, VMC5G1.1 and VVIQ61), while the two Isabella individuals differed at all but three loci. Individuals with identical genotypes were removed to prevent over-representation of the alleles prior to estimating the null allele frequencies. However, cultivars with the same genotype but with different names were retained, since they are independent members of the grapevine population in SA, resulting in the retention of 113 individuals in the final analysis.

The subsequent estimated null allele frequencies, as determined with Identity V4, were predicted at frequencies higher than 0.05 for markers VMC2B1.1, VMC1A5, VMC1G3.2, VVIN74-cjvh2, VVMD8-cjvh and UDV047, and they were therefore removed (Table 1). In addition, manual inspection of the inheritance patterns in the set of known pedigrees allowed the detection of rare null alleles for VVIB63 and VVIN62. However, VVIB63 had an observed null allele frequency of less than 0.05 and was therefore retained. VVIN62, with an observed null allele frequency higher than 0.05, was removed.

Eight markers that deviated from the expected Hardy-Weinberg equilibrium for genotype frequencies at $P < 0.01$ were also excluded, resulting in a total of 20 markers (183 alleles) retained for the fingerprint analysis (Table 1). The number of alleles per marker ranged from 4 to 13, with an average of 9.2 alleles per marker, while the effective number of alleles ranged from 1.8 to 6.5, with an average of 4.1. The inbreeding coefficient (F_{is}), as calculated by GenAIEx 6.5, was -0.05. Individual allele frequencies ranged from 0.004 to 0.74 (Fig. S2, Online Resource 5). The probability of identity using the set of 20 markers was $1.0e-20$, and the probability of identity for sibs was $1.9e-08$.

The final DNA fingerprints based on this subset of 20 markers yielded 105 different genotypes, of which 98 were unique to a single cultivar, while the remainder were assigned to seven groups of cultivars with identical genotypes (Table 2). In addition, two sets of samples were identical for 19 markers. A single VCM6G1 allele was the only difference between Emperor Seedless and Red Emperor, as well as the siblings G4-1437 and G4-1478. Several possible parent-offspring combinations were identified within this dataset (Table S1, Online Resource 1). The reported pedigrees of G4-1478, Regal Seedless, Sundance Seedless, Ebony Star and Pirobella were not supported by the genotypic data.

A common allele (236) was detected for the *Rpv3*-linked marker UDV108 in eight of the 10 known downy mildew-resistant cultivars, and it was absent in all other cultivars analysed in this study. At the other *Rpv3*-linked marker, VVIN16-cjvh, a common allele (245) was seen in nine of the 10 known downy mildew-resistant cultivars, but also in nine downy mildew-susceptible cultivars. At the *Ren3* locus, UDV116 had a common allele (146) in seven (Ajvaz, Chambourcin, Kodryanka, Moldova, Poeloeskei Muskotaly, Regent and Villard Blanc) of 11 powdery mildew-resistant cultivars (Table S3, Online Resource 4). The same allele (146) also occurred in three powdery mildew-susceptible cultivars (Alphonse Lavallée, Black Rose and Cardinal) as well as in three cultivars of unknown powdery mildew status (Belinj Originalnuj, Black Emerald and Victoria). The other *Ren3*-linked marker, UDV047, was removed from the analysis due to a predicted null allele frequency higher than 0.05 (Table 1).

DISCUSSION

The allele frequencies showed that, while there was a relatively high number of alleles for most markers, a small number of alleles had a high frequency in the population, a trend that has been observed in other studies (Bowers *et al.*, 1999; Cipriani *et al.*, 2010; Di Gaspero *et al.*, 2012).

TABLE 1

A summary of the statistical analysis of the 35 markers as calculated after the removal of the duplicated accessions showing predicted and observed null allele frequencies, along with the Chi-square Test for Hardy-Weinberg equilibrium for the genotype frequency distribution, number of alleles (Na), effective number of alleles (Ne), and the information index (I). Only the 20 markers in bold were used for generating the final DNA fingerprint.

Marker	Estimated null allele frequency	Observed null allele frequency	ChiSq	Signif	Na	Ne	I
UDV047	0.06	0.01	105.7	Ns	14	4.5	1.9
UDV108	-0.04	0.01	79.0	Ns	13	4.8	1.9
UDV116	-0.01	0.00	73.0	Ns	11	4.6	1.8
VMC1A5-cjvh	0.12	0.01	125.5	***	11	4.3	1.7
VMC1G3.2-cjvh	0.05	0.00	137.5	***	11	3.7	1.6
VMC2A3	-0.03	0.00	181.2	***	15	6.8	2.2
VMC2B1.1	0.08	0.01	16.6	*	4	2.5	1.0
VMC2F10-cjvh	-0.01	0.00	20.2	Ns	8	2.9	1.3
VMC2H4	0.01	0.01	130.2	**	14	6.2	2.1
VMC3E5	0.00	0.00	11.1	Ns	9	1.8	1.0
VMC5G1.1	0.03	0.01	58.2	***	8	2.2	1.1
VMC6B11	-0.07	0.00	78.0	*	11	3.9	1.6
VMC6G1	-0.01	0.00	39.5	Ns	8	6.0	1.9
VMC7F2	-0.10	0.00	51.8	**	8	3.2	1.4
VMC8A7-cjvh	0.01	0.00	84.7	*	11	5.7	1.9
VMC8B5	0.04	0.01	68.8	***	9	2.6	1.4
VMC8F4.2	0.01	0.01	119.4	***	8	2.8	1.3
VrZAG21	0.02	0.00	17.3	Ns	7	4.3	1.6
VrZAG25	0.00	0.00	57.5	*	9	4.1	1.6
VrZAG47	0.01	0.00	104.7	***	9	5.0	1.7
VrZAG62	-0.03	0.00	34.2	Ns	9	4.1	1.7
VrZAG79	-0.01	0.00	105.6	*	13	6.5	2.0
VrZAG83	-0.04	0.00	17.5	Ns	5	3.4	1.3
VVC62-cjvh	-0.03	0.00	52.3	Ns	11	4.6	1.8
VVIB63-cjvh	0.00	0.03	21.9	Ns	6	2.8	1.3
VVIM93	0.00	0.00	46.2	Ns	12	3.0	1.6
VVIN16-cjvh	-0.03	0.00	12.4	Ns	5	3.1	1.3
VVIN62	0.02	0.05	30.9	***	5	1.6	0.8
VVIN74-cjvh2	0.12	0.02	144.8	***	9	5.1	1.8
VVIP33	-0.03	0.00	51.9	Ns	10	5.4	1.9
VVIQ61	-0.06	0.00	5.0	Ns	4	2.3	1.0
VVIT68	-0.01	0.00	30.4	Ns	9	3.4	1.5
VVMD17	0.01	0.00	79.7	***	8	3.2	1.4
VVMD7	-0.01	0.00	51.2	Ns	12	4.5	1.8
VVMD8-cjvh	0.14	0.08	250.1	***	14	5.3	2.0

ns = not significant, * P < 0.05, ** P < 0.01, *** P < 0.001

TABLE 2
Groups of identical genotypes.

Group	Cultivars with identical genotypes
Group 1	Colombar-V15 Colombar-V27
Group 2	Early Sweet Prime
Group 3	G1-1074 G1-1075
Group 4	Honeybunch Muscat Seedless
Group 5	Muscat of Alexandria C982 Muscat of Alexandria Red
Group 6	Scarlet Dew G4-568 G4-655
Group 7	Sugraone Sugra 14 Ralli Seedless

Since the population typed in this study contains a large group of individuals that form a family group spanning several generations, and which was subjected to selection pressure during breeding for certain characteristics, it is not unexpected that alleles closely linked to any of these characteristics would have an increased frequency. As expected, the inbreeding coefficient (F_{is}) of -0.05 supports an outbreeding population (Riechert & Roeloffs, 1993).

Overall, the detection of 98 samples with a unique genotype within the set of 125 analysed samples underlined the discriminatory power of SSRs, with the remaining samples clustering in groups of identical genotypes (Table 2). The fingerprint data was in concordance with the reported pedigrees (Table S1, Online Resource 1), with some exceptions that involved either markers with undetected null alleles, inconsistent peak detection due to allelic dropout, or incorrect accessions and pedigree records.

Initially, the two Colombar samples could be discriminated with the set of 35 markers, but after reducing the final fingerprint set to 20 these samples were identical. The identical nature of the genotypes of Muscat of Alexandria C982 (white berries) and Muscat of Alexandria Red (red berries) was expected, as both are clones of Muscat of Alexandria (<http://www.vivc.de/>). Similarly, no genotypic differences were observed between Sugraone, Sugra 14 and Ralli Seedless. Both Sugra 14 (red berries) and Ralli Seedless (red berries) are reported to be mutations of Sugraone (<http://www.vivc.de/>), and therefore a high level of similarity between these cultivars was expected. While SSRs have a very high discriminatory power, they only sample the genome at the location of the marker. If a mutation event does not influence the binding of the primers or the amplicon size, it goes unnoticed by SSR genotyping, as was shown for the white and red clones of Keshmeshi (Doulati Baneh *et al.*, 2009). Other studies similarly failed to differentiate clones of a cultivar using SSRs (Dangl *et al.*, 2001; Imazio *et al.*, 2002). Prime and Early Sweet also produced identical

fingerprints and, with a lack of official information on their origin, it is likely that these two cultivars are clones.

Unexpectedly, all 20 markers in the final DNA fingerprint (and indeed all of the original 35 tested) displayed identical genotypes for Honeybunch and Muscat Seedless. Honeybunch was developed using embryos rescue from open-pollinated Muscat Seedless (Table S1, Online Resource 1). Muscat Seedless is heterozygous for 13 of the 20 markers used, and the probability of Honeybunch having this genotype as a result of the self-fertilisation of Muscat Seedless is only 0.0004. It is thus more likely that Honeybunch resulted from the somatic embryogenesis of Muscat Seedless tissue rather than from a zygotic embryo. As expected, the full sibs Bonheur and La Rochelle displayed different genotypes for the majority of the markers in the fingerprint. In contrast, the sibs Scarlet Dew and accessions G4-568 and G4-655, as well as G1-1074 and G1-1075, shared identical fingerprints (Table S3, Online Resource 4). In these cases, both parental plants are heterozygous for most of the markers used, and thus the probability of obtaining identical sibs through individual fertilisations is low. They most likely originated through spontaneous poly-embryogenesis, which often occurs in tissue culture during the embryo-rescue process. Poly-embryos may arise through different mechanisms, but they often originate through adventive embryogenesis from the zygote. Plants developed from them will thus be identical (Durham *et al.*, 1989).

SSR allele scoring can be complicated, depending on the marker profile. For dinucleotide repeat-based markers, stutter peaks are frequently encountered, resulting in multiple peaks for each allele. Furthermore, complications like allelic dropout (some alleles in the population amplify less effectively than others), or the occurrence of null alleles (some alleles do not amplify at all), can lead to errors in data typing and interpretation (Dakin & Avise, 2004). Despite efforts to eliminate unreliable markers, VMC6G1 proved to be problematic. VMC6G1 was the only marker discriminating between G4-1437 (184/184) and G4-1478 (184/205). G4-1437 and G4-1478 are full sibs and no phenotypic differences between them have been reported. Closer inspection of the data for samples G4-1478 and G4-1437 revealed a peak at 205 bp, but the intensity of this peak is significantly lower than the peak for the 184 bp allele. Allele VMC6G1-205 was not seen in either of the parents, namely Muscat Seedless and Lady Ann. Since Lady Ann is typed as a 184/184, we believe that allele VMC6G1-205 suffers from severe allelic dropout, which in some cases can lead to the allele not being detected at all. This was the only marker that did not support the parentage of G4-1478.

Unexpectedly, one VMC6G1-allele discriminated Red Emperor (195/201) from the derived seedless mutant, Emperor Seedless (195/199) (<http://www.vivc.de/>). Potential association between this allele difference and the seedless trait can only be determined by using Emperor Seedless as one of the parents in a mapping study.

It is best to avoid markers displaying allelic dropout or null alleles in cultivar and parentage analysis. Null alleles in particular pose a problem, since null alleles can only be detected as a higher than expected occurrence of homozygotes in a population, or as inconsistencies in the

inheritance of alleles in known families. Six of the 35 markers had predicted null allele frequencies greater than 0.05, while a seventh had an observed null allele frequency higher than 0.05 (Table 1). This was not unexpected, given the high frequency of sequence variation in grapevine. Variants have been reported with a frequency as high as one variant every 64 bp (Lijavetzky *et al.*, 2007). This was illustrated by the pedigree of Datal (Waltham Cross \times Muscat of Alexandria). Manual inspection revealed that the VVIB63 genotypes for Waltham Cross (129/129) and Muscat of Alexandria (108/123) were not compatible with the genotype of Datal (108/108) (Table S3, Online Resource 4). If the genotype rather includes a null allele for Waltham Cross (129/null), and thus a genotype of 108/null for Datal, the stated pedigree is supported. Similarly, if the null alleles in Waltham Cross and Datal are accepted, the pedigrees of some of the cultivars derived from them, namely Dauphine (Ohanes \times Waltham Cross; 135/135) and Victoria (Cardinal \times Waltham Cross; 131/131), can be explained. In the case of VVIB63, neither the predicted nor the observed null allele frequency was higher than 0.05, and therefore the marker was included in the final analysis. Allelic dropout and null alleles are therefore a valid concern, but the interpretation of a DNA fingerprint profile can be significantly enhanced by including core families in the analysis.

Two half-sibs, namely Regal Seedless (Datal \times Centennial Seedless) and Sundance Seedless (Datal \times Dawn Seedless), had fingerprints that excluded Datal as a parent. Five of the 20 marker genotypes for Regal Seedless (Table S3, Online Resource 4) did not support the reported pedigree. Similarly, marker profiles generated with five markers were also incompatible with the reported pedigree for Sundance Seedless. In total, seven markers yielded genotypes incompatible with the pedigrees. The genotype of Datal was confirmed by its agreement with its reported pedigree (Waltham Cross \times Muscat of Alexandria). The Lynch and Ritland mean relatedness index (Lynch & Ritland, 1999) indicated that Sundance Seedless is related to Dauphine. Closer inspection showed that only one marker, VVIB63, did not support Dauphine as a parent of Sundance Seedless instead of Datal. If the presence of a null allele at VVIB63 for Dauphine is accepted (as discussed above), then it is possible that the genotype for Sundance Seedless at VVIB63 contains a null allele, which was passed from Waltham Cross to its offspring Datal, Dauphine and Victoria. This null allele was then passed from Dauphine to Sundance Seedless. The pedigree analysis with Identity V4 was done with the set of 20 markers, and then repeated with marker VVIB63 excluded. The result confirmed that Sundance Seedless could be the offspring of Dauphine and Dawn Seedless. It also confirmed the pedigree of both Datal and Victoria, and suggested that White Gem could be the offspring of Centennial Seedless and Datal (Table S1, Online Resource 1). No alternative pedigree for Regal Seedless could be identified in the dataset.

Data from seven markers did not confirm the reported pedigree for Ebony Star (Redglobe \times Bonheur) (Fig. S1, Online Resource 3), and the presence of null alleles could not explain this conflict. However, all 20 markers supported an alternative pedigree, with Redglobe and La Rochelle as

Ebony Star's parents. This implies that either the Bonheur and La Rochelle accessions in the Nietvoorbij collection were switched, or the pedigree was recorded incorrectly. Data in Prins *et al.* (2009) reported identical profiles for multiple clones of Redglobe, La Rochelle and Ebony Star, which makes it more likely that the pedigree of Ebony Star was originally recorded incorrectly.

The two accessions of Isabella did not have identical genotypes. One accession of Isabella (Isabella-V131) shared at least one allele per marker with Pirobella (Isabella \times Pirovano 15) and Muska (Isabella \times Pirovano 15), which is consistent with the pedigrees, while the alternate profile obtained for Isabella-V82 was not compatible with the pedigrees. Furthermore, the genotype of Pirovano 15, the other reported parent of Pirobella and Muska, excluded it as a parent of both. It is most likely that the cultivar Pirovano 15 in this study is incorrectly named (mistaken identity), or that the parentage was recorded incorrectly. The large number of differences between the two accessions of Isabella (Table S3, Online Resource 4) can only be attributed to mistaken identity.

Two *Rpv3*-linked markers, UDV108 and VVIN16-cjvh (~15 cM interval; Van Heerden *et al.*, 2014), revealed a single common haplotype among eight of the nine cultivars that are assumed to carry the *Rpv3* downy mildew resistance locus, based on pedigree data (Fig. 1). The ninth cultivar (Ajvaz) retained only one resistance-linked allele (VVIN16-cjvh - 245) of the haplotype. Furthermore, this haplotype was unique to these eight cultivars in the complete dataset. In contrast, the randomly selected marker pairs VrZAG79 and VVIT68 (~6 cM; chromosome 5) and VVMD7 and VrZAG62 (3 cM; chromosome 7), which are not linked to *Rpv3*, did not display a common haplotype across these cultivars. Seven of these nine cultivars also shared an allele (146 bp) for the *Ren3* powdery mildew resistance-associated marker UDV116 (~3 cM, chromosome 15; Van Heerden *et al.*, 2014). In all cases, the allele frequencies of the alleles shared by these cultivars were relatively low. Despite the smaller interval between the markers on chromosomes 5 and 7, recombination was observed between these markers in the offspring of Seibel 6468. However, the *Rpv3* locus retained a single haplotype, revealing the effect of selection. A report on the genetic diversity of the downy mildew resistance locus on chromosome 18 (*Rpv3*) yielded similar results (Di Gaspero *et al.*, 2012). Isabella, the downy mildew-resistant offspring of an interspecies cross between *V. vinifera* and *V. labrusca*, does not share any *Rpv3*-linked alleles with the nine cultivars assumed to carry the *Rpv3* locus. While there is no evidence supporting common ancestry with these cultivars (Fig. 1), it is possible that Isabella carries an *Rpv3* locus originating from a different source. Di Gaspero *et al.* (2012) found seven different haplotypes for the *Rpv3* locus, which could be traced back to *V. rupestris*, *V. lincedumii*, *V. riparia* and *V. labrusca*.

CONCLUSIONS

Here we have reported the use of a selected set of 20 SSRs with an acceptable ability to generate a DNA fingerprint for 111 grape cultivars. The predicted and observed null allele frequency for these 20 markers were below 0.05, and the

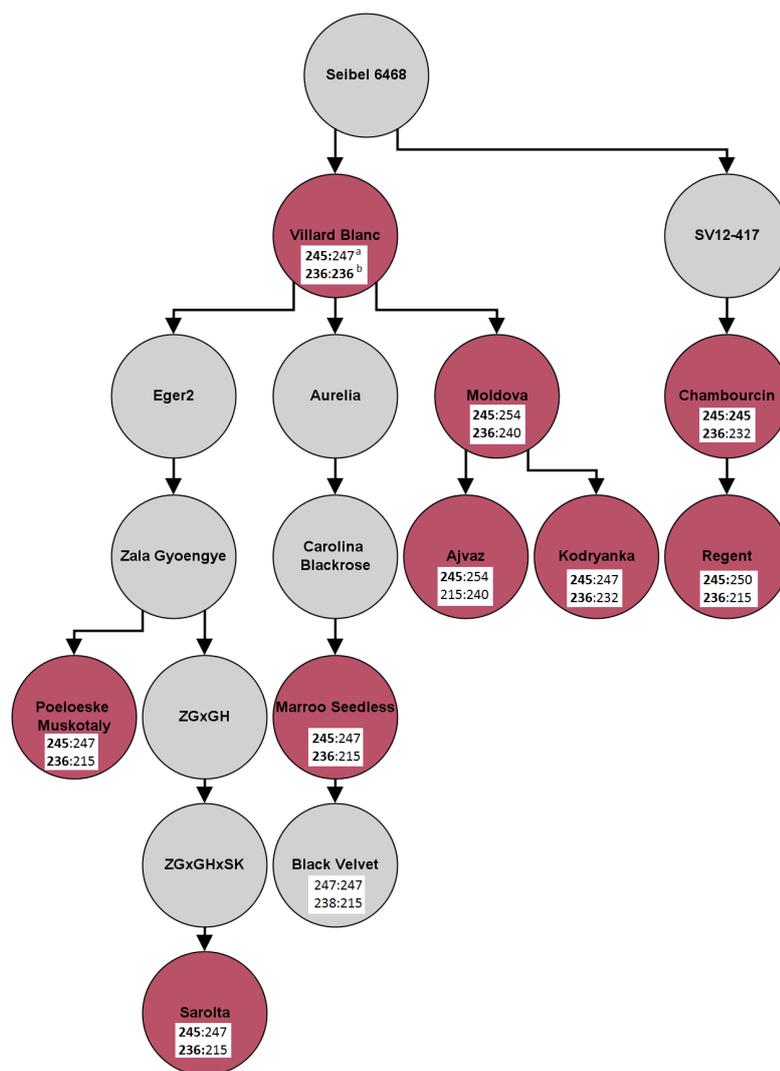


FIGURE 1

Pedigree of the descendants of Seibel 6468. The purple-shaded individuals are known to be resistant to downy mildew. The allele calls for the *Rpv3* flanking markers (^a) VVIN16-cjvh and (^b) UDV108 are indicated. For each marker, the bold allele indicates the allele common to all the downy mildew-resistant phenotypes.

genotype frequencies did not deviate significantly from the expected Hardy-Weinberg equilibrium. We assigned unique DNA fingerprints to 98 cultivars, while the remaining 13 were grouped according to seven different DNA fingerprints. The use of these DNA fingerprints in parentage analyses revealed several cases where either the reported pedigree was incorrect or the identities of the cultivars included in the study were incorrect. The high number of SSRs that had to be discarded because they were likely to yield unreliable results highlighted the potential for errors when SSR markers are used without proper care. We also illustrated the usefulness of including the known pedigrees of individuals in identifying null alleles. Many different alleles were found for most markers, of which a small number occurred at a higher than expected frequency. This most likely results from relatedness, as illustrated in the case of the large number of descendants from either Thompson Seedless or Muscat of Alexandria (Fig. S1, Online Resource 3). It is

possible that these common alleles were retained over many generations due to their linkage to advantageous traits. The gene-associated marker screen provided useful information for breeding to design informative crosses for MAS efforts to pyramid the reported fungal resistance genes with other resistance sources. The conserved *Rpv3* haplotype reported here (UDV108 to VVIN16-cjvh) spans the smaller region (UDV305 to UDV737) used to identify the origin of the *Rpv3* locus (Di Gaspero *et al.*, 2012; Venuti *et al.*, 2013). This large linkage block was retained in eight of the nine resistant cultivars, despite many opportunities for recombination that could have occurred during their development (Fig 1). As the resistant variety, Ajvaz, derived from Moldova, retained only the VVIN16-cjvh - 245 allele, it could be a useful parent in crosses aimed at reducing potential linkage drag at the *Rpv3* locus. The results support the importance of utilising molecular data in the management of plant breeding material.

LITERATURE CITED

- Adam-Blondon, A.-F., Roux, C., Claux, D., Butterlin, G., Merdinoglu, D. & This, P., 2004. Mapping 245 SSR markers on the *Vitis vinifera* genome: A tool for grape genetics. *Theor. Appl. Genet.* 109, 1017-1027.
- Akkak, A., Boccacci, P. & Botta R., 2007. 'Cardinal' grape parentage: a case of a breeding mistake. *Genome* 50, 325-328.
- Alleweldt, G. & Possingham, J.V., 1988. Progress in grapevine breeding. *Theor. Appl. Genet.* 75, 669-673.
- Bautista, J., Dangl, G.S., Yang, J., Reisch, B. & Stover E., 2008. Use of genetic markers to assess pedigrees of grape cultivars and breeding program selections. *Am. J. Enol. Vitic.* 59, 248-254.
- Bellin, D., Peressotti, E., Merdinoglu, D., Wiedemann-Merdinoglu, S., Adam-Blondon, A.-F., Cipriani, G., Morgante M., Testolin, R. & Di Gaspero, G., 2009. Resistance to *Plasmopara viticola* in grapevine 'Bianca' is controlled by a major dominant gene causing localized necrosis at the infection site. *Theor. Appl. Genet.* 120, 163-176.
- Bowers, J.E., Dangl, G.S. & Meredith, C.P., 1999. Development and characterization of additional microsatellite DNA markers for grape. *Am. J. Enol. Vitic.* 50, 243-246.
- Bowers, J.E., Dangl, G.S., Vignani, R. & Meredith, C.P., 1996. Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). *Genome* 39, 628-633.
- Cadle-Davidson, L., 2008. Variation within and between *Vitis* spp. for foliar resistance to the downy mildew pathogen *Plasmopara viticola*. *Plant Dis.* 92, 1577-1584.
- Cipriani, G., Spadotto, A., Jurman, I., Di Gaspero, G., Crespan, M., Meneghetti, S., Frare, E., Vignani, R., Cresti, M., Morgante, M., Pezzotti, M., Pe, E., Policriti, A. & Testolin, R., 2010. The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. *Theor. Appl. Genet.* 121, 1569-1585.
- Clingeffer, P.R. & Possingham J.V., 1988. Marroo Seedless – a new table grape variety. *Agri. Sci. (Aust.)* 1, 18-19.
- Dakin, E. & Avise, J., 2004. Microsatellite null alleles in parentage analysis. *Heredity* 93, 504-509.
- Dalbó, M.A., Ye, G.N., Weeden, N.F., Steinkellner, H., Sefc, K.M. & Reisch, B.I., 2000. A gene controlling sex in grapevines placed on a molecular marker-based genetic map. *Genome* 43, 333-430.
- Dangl, G.S., Mendum, M.L., Prins, B.H., Walker, M.A., Meredith, C.P. & Simon, C.J., 2001. Simple sequence repeat analysis of a clonally propagated species: A tool for managing a grape germplasm collection. *Genome* 44, 432-438.
- Di Gaspero, G., Copetti, D., Coleman, C., Castellarin, S., Eibach, R., Kozma, P., Lacombe, T., Gambetta, G., Zvyagin, A., Cindrić, P., Kovács, L., Morgante, M. & Testolin, R., 2012. Selective sweep at the *Rpv3* locus during grapevine breeding for downy mildew resistance. *Theor. Appl. Genet.* 124, 277-286.
- Di Gaspero, G., Cipriani, G., Marrazzo, T.M., Andretta, D., Castro, M.J.P., Peterlunger, E. & Testolin, R., 2005. Isolation of (AC)*n*-microsatellites in *Vitis vinifera* L. and analysis of genetic background in grapevines under marker assisted selection. *Mol. Breed.* 15, 11-20.
- Doulati Baneh, H., Mohammadi, S.A., Mahmoudzadeh, H., De Mattia, F. & Labra, M., 2009. The use of AFLP and morphological markers to study Iranian grapevine germplasm to avoid genetic erosion. *J. Hortic. Sci. Biotechnol.* 82, 745-752.
- Durham, R.E., Moore, G.A., Gray, D.J. & Mortensen, J.A., 1989. The use of leaf GPI and IDH isozymes to examine the origin of polyembryony in cultured ovules of seedless grape. *Plant Cell Rep.* 7, 669-672.
- Eibach, R., 1994. Investigations about the genetic resources of grapes with regard to the resistance characteristics to powdery mildew (*Oidium tuckeri*). *Vitis* 33, 143-150.
- Eibach R. & Töpfer R., 2003. Success in resistance breeding: 'Regent' and its steps into the market. *Acta Hort.* 603, 687-691.
- Emanuelli, F., Battilana, J., Costantini, L., Le Cunff, L., Boursiquot, J.M., This, P. & Grando, M.S., 2010. A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.). *BMC Plant Biol.* 10, 241.
- Emanuelli, F., Lorenzi, S., Grzeskowiak, L., Catalano, V., Stefanini, M., Troggo, M., Myles, S., Martinez-Zapater, J.M., Zyprian, E., Moreira, F.M. & Grando, M.S., 2013. Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. *BMC Plant Biol.* 13, 39.
- Fischer, B.M., Salakhutdinov, I., Akkurt, M., Eibach, R., Edwards, K.J., Topfer, R. & Zyprian, E.M., 2004. Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. *Theor. Appl. Genet.* 108, 501-515.
- Hoffmann, S., Gaspero, G., Kovács, L., Howard, S., Kiss, E., Galbács, Z., Testolin, R. & Kozma, P., 2008. Resistance to *Erysiphe necator* in the grapevine "Kishmish vatkana" is controlled by a single locus through restriction of hyphal growth. *Theor. Appl. Genet.* 116, 427-438.
- Ibáñez, J., Vargas, A.M., Palancar, M., Borrego, J. & de Andrés, T., 2009. Genetic relationships among table-grape varieties. *Am. J. Enol. Vitic.* 60, 35-42.
- Imazio, S., Labra, M., Grassi, F., Winfield, M., Bardini, M. & Scienza, A., 2002. Molecular tools for clone identification: the case of the grapevine cultivar 'Traminer'. *Plant Breed.* 121, 531-535.
- Laucou, V., Lacombe, T., Dechesne, F., Siret, R., Bruno, J.-P., Dessup, M., Dessup, T., Ortigosa, P., Parra, P., Roux, C., Santoni, S., Varès, D., Péros, J.-P., Boursiquot, J.-M. & This, P., 2011. High throughput analysis of grape genetic diversity as a tool for germplasm collection management. *Theor. Appl. Genet.* 122, 1233-1245.
- Lijavetzky, D., Cabezas, J., Ibáñez, A., Rodríguez, V. & Martínez-Zapater, J., 2007. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. *BMC Genomics* 8, 1-11.
- Lopes, M.S., Rodrigues dos Santos, M., Eiras Dias, J.E., Mendonça, D. & da Câmara Machado, A., 2006. Discrimination of Portuguese grapevines based on microsatellite markers. *J. Biotech.* 127, 34-44.
- Lynch, M. & Ritland, K., 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* 152, 1753-1766.
- Martínez, L.E., Cavagnaro, P.F., Masuelli, R.W. & Zúñiga, M., 2006. SSR-based assessment of genetic diversity in South American *Vitis vinifera* varieties. *Plant Science* 170, 1036-1044.
- Merdinoglu, D., Butterlin, G., Bevilacqua, L., Chiquet, V., Adam-Blondon, A.-F. & Decroocq, S., 2005. Development and characterization of a large set of microsatellite markers in grapevine (*Vitis vinifera* L.) suitable for multiplex PCR. *Mol. Breed.* 15, 349-366.
- Nybom, H., Weising, K. & Rotter, B., 2014. DNA fingerprinting in botany: Past, present, future. *Investig. Genet.* 5, 1-35.
- Pavloušek, P., 2006. Evaluation of resistance to downy mildew in grapevine genetic resources. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* 5, 99-106.
- Pavloušek, P., 2007. Evaluation of resistance to powdery mildew in grapevine genetic resources. *J. Cent. Eur. Agr.* 8, 105-113.
- Peakall, R. & Smouse, P.E., 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics* 28, 2537-2539.

- Pelsy, F., Hocquigny, S., Moncada, X., Barbeau, G., Forget, D., Hinrichsen, P. & Merdinoglu, D., 2010. An extensive study of the genetic diversity within seven French wine grape variety collections. *Theor. Appl. Genet.* 120, 1219-1231.
- Pernes, G., 2004. New resistant table grape cultivars bred in Hungary. *Acta Hort.* 652, 321-327.
- Prins, R., Van Heerden, C.J., Burger, A.L., Burger, P. & Smit, W.A., 2009. DNA fingerprinting of table grape cultivars. *Acta Hort.* 872, 269-274.
- Riechert, S.E. & Roeloffs, R.M., 1993. Evidence for and consequences of inbreeding in the cooperative spiders. In: Thornhill, N.W. (ed). *The natural history of inbreeding and outbreeding: Theoretical and empirical perspectives.* pp. 283 – 303.
- Sefc, K.M., Regner, F., Turetschek, E., Glössl, J. & Steinkellner, H., 1999. Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. *Genome* 42, 367-373.
- Shaw, P.D., Kennedy, J., Graham, M., Milne, I. & Marshall, D.F., 2014. Helium: Visualization of large scale plant pedigrees. *BMC Bioinformatics* 15, 259.
- Staudt, G. & Kassemeyer, H.H., 1995. Evaluation of downy mildew resistance in various accessions of wild *Vitis* species. *Vitis* 34, 225-228.
- Telfer, E., Graham, N., Stanbra, L., Manley, T. & Wilcox, P., 2013. Extraction of high purity genomic DNA from pine for use in a high-throughput genotyping platform. *N. Z. J. For. Sci.* 43, 1-8. doi:10.1186/1179-5395-43-3
- This, P., Jung, A., Boccacci, P., Borrego, J., Botta, R., Costantini, L., Crespan, M., Dangel, G.S., Eisenheld, C., Ferreira-Monteiro, F., Grando, S., Ibáñez, J., Lacombe, T., Laucou, V., Magalhães, R., Meredith, C.P., Milani, N., Peterlunger, E., Regner, F., Zulini, L. & Maul, E., 2004. Development of a standard set of microsatellite reference alleles for identification of grape cultivars. *Theor. Appl. Genet.* 109, 1448-1458.
- Troggio, M., Malacarne, G., Coppola, G., Segala, C., Cartwright, D.A., Pindo, M., Stefanini, M., Mank, R., Moroldo, M., Morgante, M., Grando, M.S. & Velasco, R., 2007. A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (*Vitis vinifera* L.) anchoring 'Pinot Noir' bacterial artificial chromosome contigs. *Genetics* 176, 2637-2650.
- Ulea, E., Irimia, N. & Lipşa, F.D., 2012. Incidence of major fungal diseases in ampelographic collection of USAMV Iaşi, Romania. *Research Journal of Agricultural Science* 44, 112-117.
- Van Heerden, C.J., Burger, P., Vermeulen, A. & Prins, R., 2014. Detection of downy and powdery mildew resistance QTL in a 'Regent' × 'RedGlobe' population. *Euphytica* 200, 281-295.
- Van Treuren, R., 1998. Estimating null allele frequencies at a microsatellite locus in the oystercatcher (*Haematopus ostralegus*). *Mol. Ecol.* 7, 1413-1417.
- Venuti, S., Copetti, D., Foria, S., Falginella, L., Hoffmann, S., Cindrić, P., Kozma, P., Scalabrin, S., Morgante, M., Testolin, R. & Di Gasparo, G., 2013. Historical introgression of the downy mildew resistance gene *Rpv12* from the Asian species *Vitis amurensis* into grapevine varieties. *PLoS ONE* 8(4), e61228. doi:10.1371/journal.pone.0061228
- Vignani, R., Bowers, J.E. & Meredith, C.P., 1996. Microsatellite DNA polymorphism analysis of clones of *Vitis vinifera* 'Sangiovese'. *Sci. Hortic.* 65, 163-169.
- Vignani, R., Scali, M., Masi, E. & Cresti, M., 2002. Genomic variability in *Vitis vinifera* L. Sangiovese assessed by microsatellite and non-radioactive AFLP test. *Electron. J. Biotechnol.* 5, 1-11.
- Vouillamoz, J.F. & Grando, M.S., 2006. Genealogy of wine grape cultivars: 'Pinot' is related to 'Syrah'. *Heredity* 97, 102-110.
- Vouillamoz, J.F., Schneider, A. & Grando, M.S., 2007. Microsatellite analysis of Alpine grape cultivars (*Vitis vinifera* L.): Alleged descendants of Pliny the Elder's Raetica are genetically related. *Genet. Resour. Crop. Ev.* 54, 1095-1104.
- Wagner, H.W. & Sefc, K.M., 1999. IDENTITY 1.0. Centre for Applied Genetics, University of Agricultural Sciences Vienna. Available from <http://www.uni-graz.at/~sefc>
- Zyprian, E., Ochßner, I., Schwander, F., Šimon, S., Hausmann, L., Bonow-Rex, M., Moreno-Sanz, P., Grando, M.S., Wiedemann-Merdinoglu, S., Merdinoglu, D., Eibach, R. & Töpfer, R., 2016. Quantitative trait loci affecting pathogen resistance and ripening of grapevines. *Mol. Genet. Genomics* 291, 1573-1594. doi:10.1007/s00438-016-1200-5