Chapter 1

Characterization of Grapevines by the Use of Genetic Markers

Lidija Tomić, Nataša Štajner and Branka Javornik

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/52833

1. Introduction

Grapevine (*Vitis vinifera* L.), used worldwide for producing wine, table grapes and dried fruits, is an important horticultural species; the total number of grapevine cultivars in ampelographic collections worldwide is estimated to be 10,000 [1]. Grapevine cultivars have traditionally been characterized and identified by standard ampelographic descriptors. In order to establish comparable evaluation of grapevines, a unique system for cultivar description was introduced. In 1873, the International Ampelography Committee was established in Vienna, which prepared the first international standards for the classification of grapevines based on morphological traits. Ampelography is based on visual observation of certain traits, while ampelometry developed as a method that relies on precise measurement of the phenotypic characteristics of grapevines, mainly based on leaf traits. Today, the ampelographic description of cultivars includes 150 descriptors. The Office International de la Vigne et du Vin (OIV), the Union International pour la Protection des Obtentions Végétales (UPOV) and the International Board for Plant Genetic Resources (IBPGR) agreed to establish a common methodology for the ampelographic description of cultivars, which is used for the characterization and evaluation of cultivars in order to identify them, characterize their traits, to protect authors’ rights and for the needs of gene banks. Ampelographic descriptions enable the identification of cultivars taking into account the development stage of the plants, their health status and environmental conditions [2]. Standard ampelographic methods can sometimes result in misunderstandings because the expression of morphological characters depends on the developmental phase of the plant (sample), health status of the sample and environmental conditions. At the same time, the vast number of different established cultivars makes it hard to differentiate them all by morphological characteristics [3]. In parallel, genetic erosion in grapevine germplasm has been observable, due to the world-
wide predominance of few successful cultivars in all major wine producing regions. There is a significant shift in varietal structure in favour of modern cultivars and thus a decrease or even disappearance of regionally typical or local cultivars. Accurate identification is needed for numerous such cultivars, as well as systematic characterization of identified cultivars in terms of their sustainable use and breeding for future needs and conservation. Modern viticulture must be innovative and of high quality but, at the same time, must also take environmental protection into consideration. Grape growers and wine producers need to have access to a variety of grape genetic resources, in order to be able to create new varieties and new wine tastes. Growers also need to be able to certify their products, so the accurate names of local, potentially valuable grapevine varieties, and their genetic and geographic origins, need to be available. Biochemical characterization of grapevines was developed as a supplementary method to ampelographic characterization but issues associated with enzyme extraction, the general lack of a discriminating enzyme system and inconsistency in assaying enzymes have hindered the wider application of this method. Characterization of grapevines has today been complemented by the use of molecular markers, providing a different set of data, which enables more accurate identification and extended characterization. The introduction of molecular markers has allowed more accurate identification, since the results are independent of environmental factors. DNA based markers have enabled a new approach to genetic characterization and to the assessment of diversity within an analyzed set of samples, which is important for evaluation of the range and distribution of genetic variability. In grapevines, diverse marker techniques, such as RFLP or PCR based RAPD, SSR or AFLP and, recently, SNP have been widely used during recent decades. Among them microsatellites, or SSR (simple sequence repeat) markers, have become molecular markers of choice, since they offer some advantages over other molecular markers, including their co-dominant inheritance, hyper-variability and, once they are developed, they are easy to use and the data can be readily compared among laboratories. Microsatellites have also become favoured molecular markers for identifying grapevine cultivars, and their properties enable a wide range of applications, since they are ubiquitous, abundant and highly dispersed in genomes, with high variability at most loci. In *Vitis*, a large number of markers have been developed by individual groups and these markers have been very successfully applied for genetic studies. The suitability of *Vitis* SSR markers for assessing genetic origin and diversity in germplasm collections, cultivar identification, parentage analysis and for genetic mapping is well documented.

2. Biochemical methods

Isoenzyme analysis was an important tool in the characterization of grapevines during the nineties, thus preceding the wide use of molecular marker technologies. Biochemical characterization of grapevines was developed as a supplementary method to ampelographic characterization. The biochemical approach includes analysis of isoenzymes, phenolic and aromatic compounds, as well as serological analysis of pollen proteins.
During the nineties, various studies applied isoenzymes in the characterization of grapevines. Bachmann [4] developed simplified and improved isolation of active cytoplasmatic enzymes in grapevines. The polymorphism of peroxidase isoenzyme activity in phloem and dormant canes in 313 cultivars and species in *Vitis* has been evaluated. Single polymorphic isoenzyme peroxidase was sufficient to group cultivars and to discriminate between two samples. Royo *et al.* [5] characterized eight Spanish grapevine varieties and their clones by analysis of the polymorphism of isozyme activities carried out for esterases, peroxidises, catechol oxidase, glutamate oxalacetate transaminase and acid phosphatase. In the analyses, the zymograms varied in relation to the time of sampling, phenophase and origin of the plant tissues. In this case, it was concluded that two or more repetitions of sampling and isoenzyme analysis are needed for the generation of repetitive zymogram patterns. Isoenzyme analyses were also used to assess differentiation among table grapevine cultivars. A combination of four isoenzyme zymograms (peroxidises, catechol oxidase, glutamate oxalacetate transaminase and superoxide dismutase) allowed differentiation of 31 cultivars out of 43. The catechol oxidase system showed the highest level of polymorphism. This methodology was recommended for the differentiation of grapevine cultivars by Sanchez-Escribano *et al.* [6]. Analysis of isoenzymes of catechol-oxidase and acid phosphatase also allowed differentiation of the additional cultivars Kéknyelű and Picolit, considered to be synonymic [7]. Cultivars have been reported as synonyms in the *Vitis* International Variety Catalogue, despite differences in leaf morphology and type of wine produced. Cabernet Sauvignon and Chardonnay were used as reference cultivars for isoenzyme analysis, in which the same zymograms were obtained as with previous studies while Kéknyelű and Picolit differed in both studied enzyme systems.

Isoenzymes have mostly been used in biochemical characterisation for differentiation between cultivars but issues related to the success of enzyme extraction, lack of zymogram repeatability between repeated reactions, as well as the lack of a general discriminating enzyme have hindered wider application of this method [2].

3. Molecular methods

Ampelographic and biochemical methods for genotype characterization have been shown to be dependent on environmental conditions and sample status (developmental stage of plant and health status), resulting in a lack of repeatability and reproducibility in the analyzed set of parameters. In recent decades, classical methodologies have been supplemented by molecular techniques using various marker systems for the detection of DNA polymorphism.

4. Restriction Fragment Length Polymorphism (RFLP)

Restriction Fragment Length Polymorphism (RFLP) was the first widely used marker technique for molecular characterization of grapevines. Digestion of genomic DNA by restric-
tion enzymes results in the production of numerous DNA fragments, and RFLP markers are detected by the hybridization of known probes to these fragments. Point mutations, insertions and deletions that occur within or between restriction sites can result in an altered length of RFLP fragments, revealing polymorphism among the analyzed genotypes. The main advantage of RFLP markers is their co-dominance and high reproducibility but they require a high amount of relatively pure DNA and a high labour input.

RFLP markers in grapevines have been used to differentiate between genotypes and for cultivar or rootstock identification, as well as for studying polymorphism within an analyzed set of cultivars and for verifying known relationships.

Bourquin et al. [8] used RFLP markers for the identification of grapevine rootstocks. Sixteen *Vitis* rootstocks were differentiated by means of RFLP analysis by the combination of the *Hinfl* restriction endonuclease and probes obtained from DNA sequences of cv. Chardonnay. Additionally, 5 clones of SO 4 (*V. berlandieri × V. riparia*) and 3 clones of 41 B Mgt (*V. berlandieri × V. vinifera*) were analysed however no difference within clones of a same hybrid were found, since no polymorphism appeared using different probes. These analyses were a successful continuation of the study by Bourquin et al. [9], in which rootstocks of cultivars were differentiated by RFLP analysis with the restriction enzymes *Alu*I and *Hinf*I, using 9 different *Pst*I inserts from *E. coli* recombinant clones derived from cv. Chardonnay as probes. Bourquin et al. [10] analyzed 46 grapevine cultivars by RFLP markers and detected significant polymorphism among all of them. As with rootstocks, RFLP markers could not identify cultivars belonging to the Pinot, Gewuerztraminer and Gamay group of cultivars. Forty six cultivars could be defined as belonging to six taxonomic groups, which were partially in accordance with relationships assessed from ampelographic data.

The RFLP technique showed high reproducibility but it is very demanding in terms of labour. Bourquin et al. [11] therefore reported PCR primers developed from four cloned *Pst*I DNA fragments of the cultivar Chardonnay, which had been shown to be the most informative RFLP probes from previous studies. PCR products were then digested by *Ddel*, *Hinfl* and *Alu*I. This method was shown to be suitable for rapid differentiation among the majority of commercialized rootstocks (22 rootstocks), either by direct amplification or by RFLP analysis of the amplified products but they were not able to discriminate between clones of the same hybrid (rootstock 3309 C).

Versatile techniques have been developed based on polymerase chain reaction (PCR), which is more sensitive for germplasm characterization in terms of the ability for fast generation of a huge number of markers. PCR based techniques are less laborious than RFLP and require small amounts of DNA. Randomly Amplified Polymorphic DNA (RAPD), microsatellites (SSR, simple sequence repeats) and Amplified Fragment Length Polymorphism (AFLP) have proved to be most useful for grapevine germplasm analysis.
5. Random Amplified Polymorphic DNA (RAPD)

The RAPD technique is based on a PCR reaction and the use of short primers of an arbitrary nucleotide sequence, which results in amplification of an anonymous fragment (RAPD markers) of genomic DNA. The most important advantages of the RAPD technique are its technical simplicity and the fact that there is no need for advance knowledge of the DNA sequence. RAPD reproducibility among different laboratories and the requirement for strict experimental conditions are hard to achieve, which are the main disadvantages of this technique [12]. This technically least demanding method (RAPD) became popular during the nineties and due to its ease of application, it is also used nowadays.

Collins and Symons [13] used a sensitive and reproducible RAPD technique to establish a unique fingerprint of grapevine cultivars and for assessing polymorphism within the cultivars analyzed. They demonstrated that distinguishing between cultivars is already possible using single primer or by a mixture of two primers. Jean-Jaques et al. [14] confirmed this possibility by using RAPD markers in identity analysis of eight cultivars. Among 50 RAPD primers that were used in the analysis, reliable identification of analyzed cultivar was found by comparison between the RAPD patterns obtained by at least two primers (OPA 01 and OPA 18). Grando et al. [15] used 44 RAPD primers in order to assess the genetic diversity existing between wild and cultivated grapevines. The amplification patterns of the primers used did not differentiate between cultivated and wild grapevines but this RAPD approach enabled the analysis of genetic relationships within V. vinifera L. species.

Stavrakakis et al. [16] analyzed 8 grapevine cultivars grown on the island of Crete with the use of 15 RAPD decamer primers. Each grape cultivar showed a unique banding pattern for 5 or more primers used. Genetic similarity was calculated and a dendrogram of the 8 cultivars was constructed. The obtained results demonstrated that RAPD is a reliable method for the identification, discrimination and genomic analysis of grapevine cultivars. RAPD analysis of genetic diversity has been performed for cultivars from the Carpathian Basin [17], Turkish grape cultivars [18], Indian cultivars [19], and many others.

RAPD markers have also been shown to be very efficient in distinguishing between grapevine rootstocks. This et al. [20] demonstrated a high level of polymorphism among 30 grapevine rootstock cultivars by the use of 21 decamer primers. Using three primers (OPA09, OPA20 and OPP17), it was possible to identify each of the analyzed rootstock.

RAPD marker analysis has been shown to be advantageous since it is cheaper and easier to perform than RFLP analysis or isoenzyme characterization.

RAPD markers have been successfully applied in genetic mapping. Lodhi et al. [21] constructed one of the first genetic linkage maps using population derived from a cross between Cayuga White and Aurore. The map was based on 422 RAPD markers and also included some RFLP and isozyme markers. The seedlessness of grapevines, defined through various traits (mean fresh weight of one seed, total fresh weight of seeds per berry, perception of seed content, seed size categories evaluated visually, degree of hardness of the seed coat, degree of development of the endosperm and degree of development of the embryo)
were assessed in 82 offsprings from of a cross between Early Muscat and Flame Seedless [22]. One hundred and sixty RAPD decamer primers were used, among which 12 molecular markers were identified that correlated with the seven traits of seedlessness. Identified markers can be used in a marker assisted selection to exclude seeded offsprings at an early stage breeding process. Luo et al. [23] used 280 RAPD primers to construct linkage map and found marker tightly linked to a major gene for resistance to downy mildew (Plasmopara viticola) (RPv-1). Similarly, Merdinoglu et al. [24] used 151 RAPD primers for linkage analysis related to downy mildew resistance.

6. Amplified fragment length polymorphism (AFLP)

The AFLP technique is the selective amplification of DNA fragments generated by restriction enzyme digestion. The AFLP approach enables simultaneous analysis of a large number of loci in a single assay, providing stable and reproducible marker patterns. AFLP, just as RAPD, are dominant markers, so are not suitable for parentage analysis. In grapevine germplasm analysis, the AFLP technique has mainly been used to assess genetic similarities among different varieties and to study genetic relationships among grapevines. Fanizza et al. [25] studied genetic relationships among aromatic grapevines varieties by the use of AFLP markers. The result of cluster analysis showed a separation between Moscato and Malvasia varieties but no grouping of V. vinifera varieties into aromatic and non-aromatic grapevines could be made, as had been done by some ampelographers in the past. AFLP markers were used for the characterization of a collection of 35 table grapevine varieties [26]. They detected that genetic similarity among them varied between 0.65 and 0.90, while sibling varieties derived from the same cross showed a genetic similarity over 0.80. AFLP analysis enabled distinction of all 35 analyzed cultivars and can be a powerful technique in identifying variety specific polymorphic fragments for distinguishing table grapevine cultivars.

AFLP markers have also been applied for assessing intra-varietal variability and for differentiation between clones of the same variety. The variety Flame Seedless, characterized by earlier bud burst, was differentiated from its parental genotype by analysis of 64 AFLP primer combinations. Two markers were identified, which were unique either only to the mutant or only to the parental line [27]. Cervera et al. [28] analyzed the intra-varietal diversity of 31 accessions called Tempranillo or described as a synonym of this Spanish cultivar. Two AFLP primer combinations generated 95 markers, indicating that the cultivar Tempranillo consists of various clones, with a genetic similarity over 0.97. Tomić [29] analyzed 56 samples from 5 locations of the Bosnian and Herzegovinan cultivar Žilavka by AFLP markers in order to assess intra-cultivar heterogeneity in the Herzegovina region. No clustering of Žilavka samples in relation to the location or names of the samples was detected. AFLP results showed high intra-varietal variability of cultivar Žilavka, expressing average polymorphism above 50.

AFLP have been used together with microsatellite markers in various studies in order to analyze genetic diversity within a single cultivar [30,31]; to evaluate genetic relatedness [32,33] or to identify and characterize grapevine rootstocks [34].
AFLP markers have also been used a great deal for the construction of genetic linkage
[35-40], primarily aimed at mapping markers closely linked to important grapevine traits. For example, resistance to powdery mildew is controlled by single locus Run derived from *M. rotundifolia*. Pauquet et al. [41] identified 13 AFLP markers linked to Run1 and constructed a local map around the gene. Three markers out of 13 were shown to be always present in all resistant genotypes (absent in susceptible), which makes them a good diagnostic tool for selection for resistance.

7. Short sequence repeats (SSRs) – microsatellites

Microsatellites have become widely used genetic markers for the characterization of grapevine germplasm. Microsatellites are short (1-5 bp), tandemly repeated DNA sequences that are ubiquitous, abundant and highly dispersed in genomes. The variability of length of microsatellites is caused by changes in the number of repeats units, which can be easily detected by PCR, thus providing highly informative markers. The advantage of microsatellite markers is their co-dominant inheritance, as well as high polymorphism in terms of size due to the variable number of tandem repeats. Reproducibility and standardization of the SSR technique is easy to achieve but this marker system requires prior knowledge of primer binding, which increases the cost inputs for markers development. SSR markers are used for the identification of cultivars, revealing synonyms and homonyms, pedigree reconstruction and genetic relatedness, as well as population genetic studies, genome mapping and for marker assisted selection [3,12].

Large microsatellite sets of data in grapevines have been generated by numerous studies worldwide. Many of them are available in published papers and various on-line databases. The public availability of microsatellite genetic profiles of genotyped grapevine cultivars enables comparison of the obtained data, thus allowing even wider characterisation by confirmation of trueness-to-typeness and elimination of duplicates.

*Vitis* microsatellites markers have been developed within various laboratories [42-49]. Microsatellite primer sequences from these studies are available in the literature. Thomas and Scott [42] identified 26 grapevine cultivars, 6 *Vitis* species and *Muscadinia rotundifolia* L. by means of microsatellites. They established five microsatellite loci (VVS1, VVS2, VVS3, VVS4 and VVS5) from the genomic library of *V. vinifera* L. cultivar Sultana, of which VVS2 and VVS5 were shown to be the most polymorphic ones. Thomas et al. [43] and Cipriani et al. [2] used the same microsatellites for accurate and reliable identification of 80 and 16 grapevine cultivars, respectively. Bowers et al. [44] developed four new microsatellite loci (VVMD5, VVMD6, VVMD7 and VVMD8) from the genomic library of *V. vinifera* L. cultivar Pinot Noir. Seventy-seven cultivars of *V. vinifera* L. were analyzed and all four loci showed high polymorphism, with PIC values over 75%. Bowers et al. [45] developed an additional 22 VVMD loci for CT repeat motifs, initially cloned from the genomic library of Pinot Noir and Cabernet Sauvignon. They analyzed 51 to 347 cultivars, respectively, and twelve markers out of 22 proved to be polymorphic (VVMD6, VVMD8, VVMD17, VVMD21, VVMD24, VVMD25, VVMD27, and VVMD28).
VVMD26, VVMD27, VVMD28, VVMD31, VVMD32 and VVMD36). An Austrian research group developed 15 markers from *Vitis riparia* [46, 50]. Two out of 15 loci did not amplify in *V. vinifera*, while the remaining 13 (ssrVrZAG7, ssrVrZAG15, ssrVrZAG21, ssrVrZAG25, ssrVrZAG29, ssrVrZAG30, ssrVrZAG47, ssrVrZAG62, ssrVrZAG64, ssrVrZAG67, ssrVrZAG79, ssrVrZAG83 and ssrVrZAG112) were successively analyzed in 120 cultivars. Four to fifteen alleles per locus were detected and expected heterozygosity ranged between 0.37 and 0.88. The highest information content was provided by locus ssrVrZAG79 (PI 0.05) because of the even distribution of the frequencies of the 13 alleles found. The remaining most informative markers were ssrVrZAG47, ssrVrZAG62, ssrVrZAG64 and ssrVrZAG67. Microsatellite loci from previous research with the highest values of polymorphic content are mainly used in microsatellite studies of grapevines. Loci VVS2 [42], VVMD5 and VVMD7 [44], VVMD27 [45], ssrVrZAG62 and ssrVrZAG79 [46] were chosen as a standard set of alleles for cultivar identification and distinction among cultivars [51], while loci VVMD25, VVMD28 and VVMD32 [45] have recently been used as additional microsatellite DNA markers for grapevines. Once microsatellite markers have been developed, they can be used for the analysis of different genotypes within a species and transferred between two different species within the same genus. Lefort et al. [52] designed primers for seven microsatellite loci (ssrVvUCH2, ssrVvUCH11, ssrVvUCH12, ssrVvUCH19, ssrVvUCH29, ssrVvUCH35 and ssrVvUCH40) from a microsatellite enriched genomic DNA library from the grapevine cultivar Syrah. These loci proved to be highly polymorphic for genotyping analysis of various *Vitis* species and hybrids used as rootstocks. These seven markers display high heterozygosity, all of them having a high number of amplified alleles, which makes them useful for genotype identification. Goto-Yamamoto et al. [49] also used cv. Syrah for development of new microsatellite markers. They developed 9 microsatellite primer pairs which have been successfully used for analysis of oriental and occidental cultivars, as well as for characterization of non-vinifera species (*V. labrusca*, *V. riparia* and *V. rotundifolia*).

Microsatellite studies of grapevines have many practical implications. The generation of unique cultivar profiles and assessment of true identity enables the genetic fidelity of planting material to be tested and offer solution to errors occurring through a long period of vegetative propagation. Identification and characterization of genetic material helps the selection of parents in breeding programmes and the sustainable management of germplasm collections. Microsatellite data obtained for a single genotype provide the microsatellite profile of that cultivar [3]. Since microsatellites have been shown to be a reliable tool for genotype identification, many research groups have adopted the technology and sets of microsatellite profiles have been increasing rapidly. This has enabled comparison of newly studied cultivars with those already genotyped. Comparison of genotypes of cultivars has revealed unique profiles of cultivars, as well as many cases of synonyms and homonyms. Microsatellites have been used for the identification of Portuguese cultivars [53], Greek cultivars [54], Spanish autochthonous grapevine varieties [55], Albanian [56] and Turkish varieties [57], old Slovenian varieties [58, 59]; Macedonian autochthonous varieties [60]; Algerian grapevine cultivars [61], Bulgarian cultivars [62], Romanian cultivars [63] and Bosnia and Herzegovina cultivars [64]. Microsatellites have proved to be reliable tools for identification and differentiation of grapevine rootstock [34, 50, 65].
In terms of the identification of grapevine cultivars, the question has been raised of the minimum sufficient number of loci required for accurate analysis of identity. In theory, five unlinked markers, each with five equally frequent alleles, could produce over 700,000 different genotypes [44]. In practice, this is not always easy to achieve and so the markers that are most informative should be selected for reliable discrimination [3]. Calculation of different genetic parameters has been used for assessing the informativeness of specific microsatellite loci. Counting alleles can overestimate the value of a given microsatellite locus due to the unequal distribution of alleles. Calculations that are based on allele frequencies are a more reliable measure of the informativeness of a locus. Two measures that are based on allele frequencies and genotype frequencies are probability of identity (probability of identical genotypes) (PI) and discrimination power (D) [3]. They describe the probability that two unrelated cultivars can be differentiated by a particular marker.

Discovering parentage and kinship analysis in grapevines is important for revealing the origin of particular cultivars. Selection of grapevines started almost seven centuries ago but reconstruction of the events that have led to the creation of specific cultivars is difficult. Many ancestors that could have provided evidence of the origin of grapevine cultivars have probably already become extinct [66]. Microsatellites have proved to be a reliable tool for parentage analysis, allowing the reconstruction of crosses. The origins of the widespread and best known grapevine cultivars from northeastern France were discovered by microsatellite analysis of 300 cultivars by 32 markers showing that Chardonnay, Gamay noir, Aligoté and Melon are the progeny of a single pair of parents, Pinot and Gouais blanc, dating from the Middle Ages [45]. Using 25 polymorphic microsatellite markers, Piljac et al. [67] analyzed possible parent progeny relationships within fourteen Croatian cultivars. Crespan [68] confirmed that the cultivar Muscat of Hamburg, which is a fine black table grape variety with a muscat flavour, is the progeny of Schiava Grossa × Muscat of Alexandria, which had been previously assumed in the literature. In this case, parentage was determined by analysis of chloroplast microsatellite loci. Since cytoplasm is inherited from the maternal side, it is possible to deduce the female parent. Microsatellites have been used to determine parent-offspring relationships among many grapevine cultivars. The cultivar Vitouska, which is grown in north-eastern Italy and western Slovenia, was shown to be the progeny of Prosecco and Malvasia Bianca Lunga, with one allele derived from each parent at 37 microsatellite loci [69]. The Italian important cultivar Sangiovese was shown to be the progeny of Ciliegio-lo and Calabrese di Montenuovo confirmed by the high likelihood value [70]. Cardinal is one of the most successful table grapes and, after many decades, has remained the most used table grapevine variety grown worldwide, accounting for 20% of total production. This cultivar is a Californian grapevine created by E. Snyder and F. Harmon in 1939 and should have been derived from the cross between Flame Tokay and Alphonse Lavaleé, however microsatellite analyses did not confirm Flame Tokay as a maternal parent [71]. Cipriani et al. [72] analyzed a set of grapevines consisting of 1005 international, Italian national and local varieties. Altogether, 211 putative trios (2 parents and their offspring) were determined, of which 94 were designated with high confidence (95%), 19 with relaxed confidence (80%) and the remainder with an assigned confidence level. The assigned confidence level was due to an inability to select one parent of the pair, amongst a number of candidates with equal
probability. Finally, 74 complete pedigrees were found, some of which were already known and some newly revealed. Recently, a total of 138 grapevine cultivars collected in five countries from the Balkan Peninsula were analyzed using 22 microsatellite loci. Kinship analysis resulted in various trios. Some were false trios because the apparent parent-offspring relationship was a result of near synonyms (clones or siblings). In the set of 138 samples, one unknown parentage [Furmint (Knipperlé, Ortlieber) = Pinot Noir × Rebula Stara] was revealed and one pedigree related to Serbian cultivars already reported in the literature (Župljanka = Pinot Noir × Prokupac) was confirmed. The microsatellite analysis also gave the first evidence of the origin of cv. Žilavka, most widespread autochthonous cultivar in Bosnia and Herzegovina. However, the pedigree of Serbian cultivar Petra was found to be false as the origin of cv. Godominka [73].

Microsatellites can be also used for determining the parentage of grapevine rootstock. For example, microsatellite analysis confirmed that the rootstock Fercal, which is important due to its high tolerance to limestone chlorosis, is the progeny of B.C.n°1B and 31 Richter [74]. Pedigree analysis should usually be confirmed by ampelographic observations, since misnaming and mislabeling of samples cannot be entirely excluded. Successful reconstruction of many pedigrees depends on the availability of ancient cultivars and pedigree data of cultivars.

The first genetic map based on microsatellite markers was developed by Riaz et al. [75]. The mapping population was represented by 153 progeny plants from a cross of Riesling and Cabernet Sauvignon and 152 microsatellite markers were mapped to the 20 linkage groups (LG), with an average distance between markers of 11.0 cM. Adam Blondon et al. [76] developed a second microsatellite reference map, consisting of 245 SSR markers, which was derived from the progeny of Syrah and Grenache. This map was more saturated, with 6.5 new markers per linkage group. These reference microsatellite genetic linkage maps have been further used for the fine mapping and QTL analysis.

Resistance locus Run1 was located by the microsatellite marker VMC4f3.1 [77], placed within LG12. A single dominant allele, designated Ren1, represents another source of resistance to powdery mildew (resistance to *Erysiphe necator*) [1]. Hofmann et al. [78] deduced that the closest markers to the Ren1 locus were microsatellite loci VMC9H4-2, VMCNG4E10-1 and UDV-020, assigned to LG13. Downy mildew resistance is inferred by the unique major gene Rpv1 and was found to be closely linked to Run1. Microsatellite loci that were mapped on the same linkage group have been shown to have a high correlation with the Rpv1 [24]. In relation to the presence of different flower types in grapevines (female, hermaphroditic and male), a cross between male and hermaphroditic plants was performed. The segregating ratio was 1:1 of these two types, assuming a single-locus hypothesis. The microsatellite locus VVS3 was shown to be close to the sex locus, which was mapped on LG2 [35]. Fernandez et al. [79] discovered the microsatellite locus linked to the fleshless berry mutation (flb locus) on LG18 (VMC2A3), while a seed development inhibitor, the Sdl locus, related to seedlessness, was also mapped on LG18, close to microsatellite VMC7F2 [39, 40]. Microsatellite maps have also been used for QTL mapping as for example, microsatellite markers VVS2 and VMC6G1 showed tight linkage to the magnesium deficiency QTL [80].
8. Single nucleotide polymorphism (SNP)

Advanced sequencing technologies have made available ever more sequence data, which can be used for marker development, particularly single nucleotide polymorphism (SNP). SNPs are sites in genomes where mutations naturally occur as a single nucleotide exchange (base substitutions), as a consequence of either transition or transversion events [12]. One locus of an SNP can comprise two, three or four alleles [12] but SNPs are rather biallelic markers, representing two alleles that may differ in a given nucleotide position in a diploid genome. SNPs are highly abundant, their density depends on the genome region and they differ among organisms. They are usually categorized according to their position in the genome and their effect on coding or regulatory sequences. Exonic SNPs that do not cause a change in the amino acid composition in the coded protein are synonymous SNPs, while SNPs causing a change in amino acid are non-synonymous SNPs. Non-synonymous SNPs that affect the protein function, thus influencing the phenotype, are called diagnostic SNPs. Diagnostic SNPs may be linked to specific important traits and their detection is one of the most important aims of discovering and developing SNPs.

A number of methods for SNP discovery and genotyping are available, although not all of them are equally useful nor it is clear which is the most suitable and most efficient [81]. The discovery of SNPs can usually be done by either a database search or an experimental approach. Most SNPs are extracted from expressed sequence tag (EST) databases [12]. In the experimental approach, candidate genes or genome regions are screened for the presence of SNPs by a series of techniques, such as microchip hybridization, direct sequencing or electrophoresis of PCR fragments containing candidate sequences on DNA single strand conformation polymorphism (SSCP) or denaturing gradient (DGGE) gels [12, 81]. SNP genotyping techniques can be classified into various groups: direct sequencing, cleaved amplified polymorphic sequences (CAPS), allele-specific PCR, allele specific primer extension, allele specific oligonucleotide hybridization etc. [12].

In Vitis, the identification and detection of single nucleotide polymorphisms for the development of molecular marker systems have recently dramatically increased with the publication of whole genome sequences [82, 83]. Previously, Salmaso et al. [84] scanned grapevine genes (sugar metabolism, cell signalling, anthocyanin and defence related pathways) to explore the possibility of developing an SNP marker system. Seven V. vinifera L. cultivars, the American species V. riparia L. and one complex hybrid were analysed for the distribution of SNPs along the gene fragments in order to assess the frequency and type of SNPs, nucleotide diversity, haplotypes and polymorphic information content using SSCP on none-denaturing gel electrophoresis and DNA re-sequencing of PCR amplicons. They discovered 247 SNPs among analysed genotypes which present useful markers for genetic analysis in grapevine.

Troggio et al. [81] also successfully used SSCP methodology and mini-sequencing for the development of SNP markers in grapevines, showing this to be an affordable mid-throughput methodology, which could be used for medium sized marker assisted selection projects.

Dong et al. [85] developed 21 primer pairs from grapevine EST sequences, generating 144 sequences by PCR amplification which revealed 154 SNPs. A phylogenetic tree was con-
constructed from these data, which discriminated well among the analyzed 16 cultivars (11 Eurasian and 5 Euramerica cultivars), proving SNPs to be effective for grapevine genotype identification.

Lijavetzky et al. [86] employed high throughput SNP discovery approach for analysing 230 gene fragments of eleven genotypes. The approach enabled the discovery of 1573 SNPs of which 96 were submitted to high throughput genotyping technology for marker development. 80 SNPs were successfully genotyped in 360 grapevine genotypes, with a success rate of 93.5% within a sample.

At the start of large-scale development of SNP markers, low and mid throughput methods were available for SNP detection and identification of grapevines. Pindo et al. [87] provided a high throughput SNP genotyping method (SNPlex genotyping system), which correlated with the completion of the sequencing of the heterozygous genome of Pinot Noir [83]. About 950 candidates SNP from non-repetitive contigs of the assembled genome of Pinot Noir, were tested on 90 progeny of a Syrah × Pinot Noir cross. They obtained 563 new eSNPs and mapped them according to their quality values. This methodology was shown to be accurate and reproducible, and the high level of throughput enabled analysis of several hundred SNP in a hundred samples at the same time. Myles et al. [88] identified 469,470 SNPs from reduced representation libraries from 17 grapevine samples (10 V. vinifera L. cultivars and 7 wild species), which were sequenced using sequencing-by-synthesis technology. A subset consisting of 8898 SNPs were validated which are referred to as a Vitis9KSNP genotyping array. This 9K array demonstrated the power to distinguish between V. vinifera L. cultivars, hybrids and wild species, resolving the genetic relationships among diverse cultivars.

Cultivar identification is one of the many applications of the various marker systems. In relation to the greatly used microsatellites, it has been proved that six SSR loci are enough for genetic identification of most cultivars, with a cumulative probability of identity of 4.3 × 10^-9 [51]. Lijavetzky et al. [86] found that SNP markers generated a lower PIC than microsatellites, thus requiring a higher numbers of markers to achieve similar PI values. It has been estimated that 20 SNPs with a minor allele frequency above 0.30 are needed to achieve a similar PI as when six SSR loci are used. The advantage of SNPs is reflected in their bi-allelic nature, since there are still frequent problems of microsatellite allele identification among different labs using different techniques for allele separation.

A set of 48 SNPs was proposed as a standard set for grapevine genotyping [89]. For successful genotyping, these 48 SNPs were chosen from an initial set of 332 SNPs, and are showing high information content, small minor allele frequency and are equally distributed across 17 chromosomes of grapevine (2-3 SNPs per chromosome). They have similar discrimination power to a set of 15 microsatellite markers.

SNPs markers have been shown to be efficient in parentage/offspring and kinship analysis. Zinelabidine et al. [90] used SNP markers to assess the role of the cultivar Cayetana Blanca in terms of its genetic relationships with other Iberian and Mediterranean cultivars. A total of 427 cultivars were analyzed as possible parent candidates, using 243 SNPs. It was discovered that Cayetana Blanca is a putative parent of several other Iberian varieties. Cayetana
Blanca and Alfrocheiro Preto gave rise to 5 cultivars used in Portugal and found in this study to be sibling cultivars. Cayetana Blanca parents remain unknown but the analysis indicated that this cultivar is the progenitor of several cultivars that are grown on the Iberian Peninsula, thus also being of Iberian origin.

SNP markers are useful in genetic mapping studies particularly in search of trait-linked markers. SNP markers highly associated with berry weight variability in grapevines have been identified. While searching for SNP markers linked to the fleshless berry mutation, 554 SNPs were identified along the \textit{flb} region (assumed to comprise four genes involved in berry weight variation). This nucleotide diversity demonstrated by the discovered SNPs could be further used for developing a genotyping chip useful for fine mapping of the \textit{flb} gene and analysis of genetic diversity [91]. Emanuelli et al. [92] confirmed the role of the candidate gene \textit{VvDXS} in determining the muscat flavour in grapevines. This study revealed three SNPs that are significantly associated with muscat flavoured varieties, while an SNP in the coding region of \textit{VvDXS} has been suggested as the causal gain of function mutation. Polymorphisms in the nucleotide sequence of \textit{VvDXS} could be applied in marker assisted selection for rapid screening of seedlings for their potential to express muscat flavour.

Single nucleotide polymorphisms represent a new generation marker system that is nowadays compared favourably to the greatly used microsatellite markers in grapevines. The major advantage of SNPs is their higher abundance within a genome, and they are more present in coding regions with a high possibility of being trait linked in genome mapping. Since the assessment of the grapevine genome sequence of a highly homozygous genotype [82] and heterozygous clone of Pinot Noir [83], high throughput methodologies for SNP detection and identification have become available, with the results easily transferable between different laboratories. This transferability is also reflected in the bi-allelic nature of SNPs as opposed to the allele binning related to microsatellites, and no use of reference cultivars is needed. The allele binning issue in microsatellites has been partially overcome with the discovery of 3 to 5 core repeats and microsatellites still remain markers with higher PIC values than SNPs.

Author details

Lidija Tomić\textsuperscript{1,*}, Nataša Štajner\textsuperscript{2} and Branka Javornik\textsuperscript{2}

*Address all correspondence to: lidija.tomic@agrofabl.org

1 University of Banjaluka Faculty of Agriculture, Bosnia and Herzegovina

2 University of Ljubljana Biotechnical Faculty, Slovenia
References


