We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

4,500
Open access books available

119,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Application of Microsatellite Markers in Grapevine and Olives

Jernej Jakše, Nataša Štajner, Lidija Tomić and
Branka Javornik

Additional information is available at the end of the chapter

1. Microsatellite markers

Since their discovery in the 80s, microsatellites have become a popular molecular marker for studying plant genomes and are still the marker system of choice for various applications, such as genetic diversity and genetic structure studies, fingerprinting of individuals, parentage analyses and mapping studies. Although they have been used as a PCR marker system for more than 20 years now [1, 2], the numerous recent publications on their use confirm their durability and relevance. This is mainly due to their intrinsic properties (associated high polymorphisms) and a constant evolution of the technical methodology in terms of high throughput, ease of use and price. The starting methodology was based on radioactive labelled amplified microsatellite alleles separated on polyacrilamide gels. Nowadays, highly multiplexed fluorescently labelled microsatellites are commonly genotyped in capillary based automatic systems.

1.1. Microsatellite specifications, nomenclature and definitions

Microsatellites are part of tandemly repeated sequences of the genome, where a specific core motif is repeated several times. The term microsatellite is coined from the term "satellite", which originates from DNA buoyant density gradient centrifugation experiments, in which DNA fragments with different base composition were separated from the main genomic DNA and formed a so-called "satellite" band. It was found that these satellite bands contain tandem arrays of repetitive sequences [3]. Based on the length of the core repeat unit, the repetitive DNA is classified as satellite, minisatellite or microsatellite DNA. While the repeat units in satellite and minisatellite DNA can be from 100 kb to over Mb and from 10 to 80 bp long, respectively, the core repeat unit of microsatellites is the shortest and in a range from 2

to 8 bp [4]. Some researchers also consider mononucleotide tracts (e.g., $(A)_n$) to be part of the microsatellite DNA [5], although they are less suitable for marker development and genotyping purposes, due to their properties. In some classifications, only repeats up to 5 bp are considered to be part of microsatellite DNA [6]. Nevertheless, the commonest targets for marker development are di-, tri- and tetranucleotide microsatellites.

In addition to microsatellites, several synonymous terms are used to describe the smallest class of tandem repeats. The term microsatellite was initially used to describe the most frequent human dinucleotide repeat $(CA)_n/(GT)_n$ [2] and various terms were used for other types. Synonyms are also often used for describing microsatellite sequences, such as "simple sequence repeats" (SSR), "short tandem repeats" (STR), and "variable number of tandem repeat" (VNTR). The VNTR term is particularly suitable for describing both microsatellite and minisatellite sequences and for bridging the gap between these two types [7]. Hancock [8] proposed that only the term microsatellite should be used, to avoid confusion. Based on the repeat type and its composition, the following nomenclature and classes of microsatellites have been proposed [7]:

a. a pure or perfect microsatellite consists of only one type of microsatellite repeat, e.g., $(AG)_{14}$, $(ACA)_9$,

b. a compound microsatellite consists of at least two different types of microsatellite repeats, e.g., $(CT)_{10}(AT)_{12}$,

c. an interrupted microsatellite (often also listed as imperfect) has a core sequence repeat interrupted by a short insertion of bases not following the repeat type, e.g., $(AG)_8CCC(AG)_{10}$; they can be of pure or compound type,

d. authors also use the term complex microsatellite, in which short arrays of repeats are interrupted by sequences that are themselves short repeats.

Another phrase that describes microsatellite-like sequences and is useful for proper annotation of such sequence arrays is cryptic simplicity [8]. Such regions resemble microsatellite repeats but are interrupted many times with irregularities. The authors suggested that these sequences are an intermediate stage during the birth or death of the microsatellite.

1.2. Microsatellite frequencies and distributions in plant genomes

Numerous publications deal with analysis of the frequencies and distributions of microsatellites. Citing all of them is beyond the scope of this chapter. We will highlight the first published papers related to database searches of plant sequences, and data on two model plants - rice and *Arabidopsis* - as representatives of monocot and dicot kingdoms. In grapevine, the genome sequence is available and positions of microsatellite sequences are known. In olive, however, the amount of sequence data is still scarce. Microsatellites were at first considered to be part of the "junk" part of the genome but there is planty of evidence today that they are also abundant in genes as part of promoters, UTRs, introns or even coding sequences.

The first surveys of publicly available sequence data of higher plants for the presence, abundance and ubiquity of di- and trinucleotide repeats were conducted in 1993 [9, 10]. They found

that the most frequent dinucleotide repeats were $(AT)_n$ tracts with 74%, followed by $(AG)_n$/ $(TC)_n$ with 24% and $(AC)_n$/$(TG)_n$ with 1%. These were the first publications to indicate the different frequencies of microsatellite repeats in plants compared to animals and humans, in which $(AC)_n$/$(TG)_n$ repeats are by far the most frequent class and the $(AT)_n$ type quite rare. The most abundant trinucleotide repeats were $(TAT)_n$ and $(TCT)_n$ microsatellites, accounting for 27.5% and 25%, respectively. Based on the volume of data they searched, they estimated that the average distance between microsatellites would be about 50 kb. With respect to the coding sequences, they found that 22% of dinucleotide types of repeats can be associated with the 5′ or 3′ UTR regions and introns, whereas trinucleotides can also be found in coding sequences. This is because the change in the repeat length of trinucleotide microsatellites does not disrupt the reading frame. A study by Lagercrantz et al. [9] augmented database search with Southern blot analyses of the microsatellite repeats. A study by Wang et al. [11] searched for microsatellite presence in organellar (1.2 Mb) and genomic (3 Mb) plant DNA sequences. They found a low frequency of organelle specific microsatellites, while in general confirming data found by Morgante and Olivieri [10]. Numerous publications followed, analyzing ever larger volumes of plant sequences or even whole genome data. The results mostly narrowed down the average distance between microsatellite loci, correcting the frequency distributions of specific repeats and highlighting species specific details.

A species specific search was conducted on a large set of rice sequences, with an emphasis on express sequence tags (ESTs) to develop markers for mapping [12]. The most abundant dinucleotides were $(GA)_n$ repeats, while among trinucleotides, GC rich repeats of $(CGG)_n$ and $(GAG)_n$ types were most common. The latter may be due to the higher GC content of Poales genomes [13] or the specific poly amino acid tracts present in certain coding sequences. The next rice study searched over 58 Mb of rice DNA sequences [14], which confirmed GC rich trinucleotides to be the most abundant microsatellites in the rice genome. The authors also noted the association of $(AT)_n$ microsatellites with miniature inverted-repeat transposable elements, which make them unusable for marker development. With the availability of whole genome sequences of rice [15], a complete genome survey of rice microsatellites was possible and a list was published of 18,828 perfect microsatellite repeats in a length > 20 bp, which behave as hypervariable loci. The whole genome scan confirmed previous reports that $(AT)_n$ and $(CCG)_n$ repeats are the most common ones in rice (> 35% and ~ 10%).

A study by Cardle et al. [16] investigated the expanding quantity of sequencing data in public databases and compared *Arabidopsis* genomic DNA sequences > 10 kbp and EST data searches with data of certain other plants. The results showed a lower frequency of microsatellites in EST data, with an average distance between microsatellite loci in genomic data being 6.04 kb and 14 kb for ESTs. In genomic data, the frequency of di- and trinucleotides was comparable, while in EST data trinucleotides were more than 2 times more abundant than dinucleotide repeats. Although the amount of genomic sequences from other plants was lower than with *Arabidopsis*, the average distance between microsatellite loci was comparable with *Arabidopsis*, being 7.4 kb in barley and 6.4 kb in potato. Finally, the *Arabidopsis* genome was the first sequenced plant genome to become available, at the end of 2000 [17]. A study by Morgante et al. [18], in which genome and EST sequences of *Arabidopsis* and 4 ma-

jor crops were used to estimate microsatellite densities, showed that overall microsatellite frequency is related to the investigated genome size and the amount of its repetitive DNA, but the proportion of microsatellite sequences in the transcribed part of the genome remained constant. The authors concluded that plant microsatellites reside in the low-copy part of the genome, which predates known expansions that have occurred in many species.

Due to its economic and cultural importance and relatively small genome size, the genome sequence of the grapevine (highly selfed Pinot Noir and Pinot Noir) is available [19, 20] and the microsatellite content and distribution has been analyzed [20]. The authors reported on 73,853 microsatellite loci (2-8 bp core repeat unit length) totalling up to 1.8 Mb of the grapevine genome.

Olive is a rather neglected species in terms of the availability of sequences compared to other crops or fruit species. The largest available set of olive EST data was obtained by next generation sequencing methodology (454), by which several thousand microsatellites were detected in raw sequencing data [21]. The analyzed data are accessible through WWW available Olea EST db in which 13,636 unique sequences contain microsatellites (including mononucleotide tracts), representing 5.2% of total sequences.

1.3. Searching for microsatellites

Due to their high polymorphism, which is reflected in multi-allelic patterns at a particular locus, microsatellites are ideal targets for the development of molecular markers. Several strategies have been developed for this purpose, the most ideal of which is locus specific amplification of a microsatellite site by PCR [10]. For this purpose, the DNA sequences surrounding the microsatellite need to be known, so sequence data is required as the first step. Where species specific sequence information is not available, therefore, genomic libraries need to be developed and screened for the presence of microsatellites. These isolation methods can be classified as traditional and specific ones, implementing enrichment strategies and, recently, also next generation sequencing (NGS) approaches.

The traditional microsatellite isolation method makes use of a classical genomic library and Southern screening of such a library with a microsatellite sequence [22]. A problem of such an approach is screening several thousand bacterial clones to obtain only a few microsatellite sequences, due to the low frequency of microsatellite containing clones. This approach was used in the first studies of isolating grapevine microsatellites of VVS and VVMD sets, in which reports on 5 [23] and 4 [24] developed markers was published. The authors reported 0.5% and 1.2% of colonies being positive for two different dinucleotide microsatellites [24] and 0.6% of positive ones for one type of dinucleotide repeat [23]. The first microsatellite markers published for olive of ssrOeUA set were also isolated using the classical approach [25].

Because the traditional approach was very labour intensive, various enrichment strategies were adopted to increase the number of microsatellites in genomic libraries. Such strategies were based on different approaches, e.g., using a *dut/ung* bacterial selection [26] or hybridization capture using either biotylinated microsatellite probes and magnetic particles [27, 28] or microsatellite probes attached to small pieces of nylon membrane [4, 29, 30]. These proce-

dures substantially increased the proportion of microsatellite sequences in libraries up to 95%, which in some cases enabled skipping the tedious Southern screening of the library. Such approaches were used in the discovery and development of additional microsatellite markers for olive [31-33] and *Vitis* species [34].

The emergence of NGS enabled a quantum leap in microsatellite discovery, since massive sequencing enabled the production of a huge amount of sequencing data for several species at the same time [35, 36]. The Southern screening step is no longer needed with the NGS approach.

Where larger amounts of species specific DNA sequences are available, they can be mined for microsatellite repeats using devoted software tools, omitting the costly step of library development. A comprehensive overview of mining tools with specific characteristics and their limitations is available [37]. The database mining approach has been used extensively for mining new microsatellite markers in grapevine, for which public DNA sequences were already available [38, 39].

1.4. Genotyping methodology

Several advances in genotyping methodology enable studies partially to automate the process, populate data in real time and to compare and store the genotyping data easily and efficiently. Inter-laboratory comparison of the genotyping data has become easy. All advances have sought to achieve two goals to make genotyping faster and cheaper. Microsatellite genotyping has basically followed the advances of Sanger sequencing, since the same equipment and methodology is used – separating the fragment within a resolution of 1 bp.

Initially, thin denaturating polyacrilamide gels were used and fragments visualized either by means of radioactive nucleotides [2] or radioactively labelled primers [1] or, in laboratories without "hot rooms", silver staining procedures were adopted [40].

Automated laser induced fluorescence sequencing revolutionized DNA sequencing and the first fluorescent dyes were introduced, which were also successfully adopted for genotyping purposes. Equipment still relied on polyacrilamide gel electrophoresis but was able to acquire the data in real-time and no post gel handling was required. Gel based systems were later replaced by capillary ones, whereby a substantial breakthrough in automated sample handling was achieved. These systems are nowadays widely used in microsatellite genotyping applications.

Another achievement that can speed up analysis and reduce the costs is multiplexing – a procedure by which several microsatellite loci are co-amplified together in a single tube. The procedure relies on non-overlapping allele sizes of the loci used and on using different fluorescent fluorophores. Up to five different fluorescent dyes can nowadays be used simultaneously in genotyping applications. Multiplexing requires careful development of primers and precise determination of optimal reaction conditions to achieve co-amplification of several loci, since interactions during PCR are more likely to occur when several loci are amplified together. A multiplexing approach has been developed for grapevine [41]. An easier approach that is often used is post-PCR multiplexing, in which single loci amplifications are

pooled together after PCR and separated in a single lane [42]. A problem associated with the use of fluorescently labelled primers is the high price of the dye. An economic labelling method, based on the elongation of one primer for a common sequence and using a third labelled primer in a PCR reaction, has been developed [43] and is now widely used, especially when a new set of markers is in the developing and optimisation phase.

2. Application of microsatellite markers in grapevine

2.1. Microsatellite marker development

Methods that enable analysis at the level of cultivar genotype have been developed because identification of grapevine cultivars based on morphological differences between plants may be incorrect due to the influence of ecological factors. In the last twenty years, various techniques for the characterization of cultivars at the level of DNA (RFLP, RAPD, AFLP, SCAR and SSR markers) and isoenzymes have been established, of which the most appropriate for genotyping are those using microsatellite markers. Microsatellites, in addition to some basic applications, allow the identification and determination of genetic relationships and the origin of varieties and grapevines preserved in collections or found only in vineyards, where they are usually grown only to a minor extent. Many grapevine varieties have several synonyms, meaning that they have different names, although they carry an identical genotype, which can be proved by analysis of microsatellite loci. In some cases, there are also groups or pairs of varieties that have the same or a very similar name but a different genetic background; such varieties are called homonyms.

Microsatellites or simple sequence repeats (SSRs) have proved to be the most effective markers for grapevine genotyping [24, 44-50]. Many microsatellites are highly variable both within and between species. The polymorphism between individuals is mainly accounted for changes in the number of repetitions of the basic motif [51]. The great variability of microsatellites is associated with the fact that from 10^4 to 10^5 microsatellite loci are randomized in the genome of eukaryotes, which means a large number of polymorphic sites that can be used for genetic markers. Because of the high mutation rate of microsatellite sequences, they are highly informative molecular markers, with a maximum value of polymorphism information and as such have been established for the identification of grapevine cultivars.

Thomas et al. [24] first used microsatellites for the identification of grapevine cultivars and demonstrated that microsatellite sequences are often represented in the grapevine genome and are very informative for the identification of *V. vinifera* cultivars. Detection of microsatellite polymorphism by the PCR technique is fast, easy and efficient, even with a very low quantity of DNA, which means that in the case of grapevine, products such as must and wine can be used for DNA analysis instead of plant tissue [52, 53]. Because of these characteristics, microsatellites have proved to be very effective as molecular markers for genotyping, identification studies, for solving dilemmas of synonyms, homonyms or the origin of varieties, relatedness studies, for population genetic studies, for the identification of clones and for marker assisted selection.

2.2. Comparison of developed markers

Microsatellites are known to have different mutation rates between loci [54] and there are several potential factors that contribute to the diverse dynamics of the development of microsatellite sequences: the number of repetitions, type of repeat sequence motif, the length of repeat units, interruptions in microsatellite, flanking regions, recombination rate etc.

Hundreds of microsatellite markers for grapevines have been developed and most of them are publicly available [23, 38, 41, 55-60], large set also by the Vitis Microsatellite Consortium by the company Agrogene (France). The extraordinary potential of some of them and their usefulness in determining grapevine cultivars and rootstocks has been demonstrated in many studies and they have been used for identification in most European winegrowing regions. A set of six (VVS2, VVMD5, VVMD7, VVMD27, VrZag62, VrZAG79) or nine (+ VVMD32, VVMD36, VVMD25) microsatellite markers has mostly been used in grapevine gentyping studies, which are highly polymorphic and most appropriate for determining genetic variability among European grapevine cultivars [61, 62]. Microsatellite markers are evaluated on the basis of various parameters of variability: observed heterozygosity (Ho) is the proportion of heterozygous individuals in the analyzed sample; expected heterozygosity (He) or genetic diversity shows the percentage of the population that would be heterozygous if an accidental cross occurs between individuals; the polymorphic information content (PIC) includes both the number of alleles detected at each locus, as well as the frequency of each allele and is the rate at which a marker unambiguously determines the genetic identity of an individual; the probability of identity (PI) is the likelihood of two randomly chosen individuals having two identical alleles at any locus; the power of discrimination (PD) is the probability that two randomly sampled accessions in the studied population can be differentiated by their allelic profile at a given locus. Higher PI values or lower PD values show a low discrimination power of the locus, which is usually the consequence of a small number of alleles or the high frequency of one allele.

On average, the number of amplified alleles per locus has been similar among different studies [46, 57, 63, 64] but the variability mostly depends on the size and heterogeneity of the sample. In contrast, the discriminative power of loci can vary significantly; for example,, in Slovenian grapevines SsrVrZAG79 proved to be the most informative locus, with a PD value of 0.928 [65] but in Portuguese grape varieties [63], this locus was considered to be least informative. The comparison confirmed the findings of Sefc et al. [46] that the discrimination power of each marker depends on the set of analyzed samples, which is related to the fact that different alleles are dominant in different regions the vines are growing.

Locus VVMD5 also proved to have high discriminative power in analysis of Slovenian grapevines (0.925) [65], Castilian – Spain grapevines (0.934) [48] and also in the analysis of grapevines collected in Balkan countries (0.932) [66]. In the last study, the maximum power associated with high PD values (0.96, 0.94) was evidenced separately for loci VVMD28 and Vchr8b. Locus Vchr8b is one of the 'new' microsatellite markers, containing tri-, tetra- and penta-nucleotide repeats selected from a total of 26,962 perfect microsatellites in the genome sequence of grapevine PN40024 [38]. In the study by Cipriani et al. [49], based on the genotyping of 1005 grapevine accessions with a 'new' set of 34 SSR markers with a long core re-

peat optimized for grape genotyping [38], the loci with the highest power of discrimination were Vchr3a and Vchr8b. However, from later results it can be concluded that locus Vchr8b is highly discriminative but also shows a high estimated frequency of null alleles (>0.20), which may indicate an excess of homozygotes, expected to some extent in grape or a mutation at the priming site of the locus. The presence of null alleles for the loci, as for example Vchr8b and VVMD36 was observed in different studies [49, 65, 67, 68] and usually loci with null alleles resulted in no PCR amplification for samples representing the homozygous genotypes and lead to greater number of missing data in the study.

The comprehensive ranking of 'new' and 'old' SSR markers was facilitated in the study of Tomić [68] where all potentially good markers were evaluated together and according to their power of discrimination (only for loci with PD>0.9) ranked as follows: VVMD28, VChr8b, VVMD5, VrZAG79, VVMD32, VChr3a.

Based on high values for power of discrimination (PD), it can be said that alleles are uniformly distributed among the analyzed samples and that loci are very informative. A low PD value despite a large number of amplified alleles at a specific locus is sometimes due to the uneven distribution of allele frequencies in the analyzed sample, as for example at locus VVMD7 [65], where the frequencies of three out of ten alleles added up to 85%. Locus Vchr8b amplified 21 alleles in two studies [49, 68] but only 6 alleles were shown to be effective and two alleles prevailed, with frequencies over 20% [49].

A study by Laucou et al. [50] comprises the largest analysis of genetic diversity in grape ever, with an estimate of the usefulness of 20 SSR markers scattered throughout the genome in a set of 4,370 accessions [3,727 Vitis vinifera subsp. sativa accessions, 80 Vitis vinifera subsp. sylvestris individuals, 364 interspecific Vitis hybrid accessions used for fruit production and 199 Vitis rootstocks). Of these markers, 11 were from previous studies [61] and 9 from a genetic map [59], chosen according to their position and ease of genotyping. When arranged according to PI, a set of eight markers (VVIp31, VVMD28, VVMD5, VVS2, VVIv37, VMC1b11, VVMD27 and VVMD32) was determined as sufficient for identification of all the cultivars. The highest observed PD calculated from 2,739 single accessions was obtained for VVIp31 and VVMD28 markers [0.982 and 0.981, respectively) and five out of the eight most discriminative markers belong to a previously reported set of 'old' markers. Based on criteria such as multiplexing and easy-scoring, Laucou et al. [50] defined another minimum set of nine SSR markers (VVMD5, VVMD27, VVMD7, VVMD25, VVIh54, VVIp60, VVIn16, VVIb01, VVIq52) and proposed them for the routine analysis of European grapevines.

However, there are some limitations even with SSR markers, such as when the PCR amplification gives instead of one or two expected fragments (alleles), a group of fragments that differ by only 2 bp. Additional fragments, also called secondary fragments (stutter bands), are usually caused by slippage during amplification with *Taq* polymerase and the determination of allele lengths can therefore be difficult, especially if the two alleles differ only by two bp and it is necessary to distinguish homo-and heterozygous form. In reviewing for stutter bands the set of nine di-nucleotide markers currently in use, locus VVS2 has by far the strongest stutter bands,VVMD32 has two or three stutters, but not distracting because the "main" peak is well established, VVMD5, VVMD7, VVMD27 and ZAG62 all have one

stutter and VVMD25, ZAG79, VVMD28 have no stutter bands. Tri-, tetra- and penta-nucleo-tide SSR markers are less prone to stuttering and the space between adjacent alleles is larger than in di-nucleotide SSRs, which enable a clear distinction between true alleles and stutter bands and minimize miscalling of the true allele. To overcome these limitations, Cipriani et al. [38] developed 'new' tri-, tetra- and penta-nucleotide repeated markers, which have proved to be very efficient [68].

2.3. Effectiveness of microsatellite markers in different applications

2.3.1. Chimerism

Microsatellite markers have often been used to differentiate grapes at a cultivar level and have been less interesting and less effective for the study of clonal variation [46]. Many cases have been recently described in which clones of grape varieties can be distinguished with microsatellite markers, such as 'Pinot Noir' 'Pinot gris', 'Pinot blanc' [69], 'Pinot Meunier' [70] 'Chardonnay' [71], synonyms of variety 'Black Currant' and 'Mavri Corinthiaki' [72], 'Pikolit' [73], etc. Laucou et al. [50] tested whether SSR markers could easily identify cultivars and clones when applied to a very large set of grape samples. Five percent of differentiated clones revealed between 1 and 3 differences (and only one mutant with four differences). Differences were sometimes of a homozygote versus heterozygote type or size shifts in 1 al-lele. It was demonstrated that cultivars showed at least four allelic differences, while clones showed fewer than four allelic differences but can also be distinguished. Studies of microsa-tellites have also demonstrated that the main type of mutation that leads to clonal variation is the development of chimeric growing tips. A chimera is a specific type of genetic mosaic, which is usually the result of mutation in one cell of the shoot apical meristem, spread by replication and cell division. The presence of a third allele suggests that the plant is a pericli-nal chimera, in which a mutant allele is present only in the L1 layer, as described by Riaz et al. [71]. Most chimeric cultivars do not exhibit special phenotypes, although some of them do so, such as Pinot Meunier (trichomes) and pinot gris (berry colour). Chimerism is usually detected at various loci, such as in studies by Tomić [68] and Stenkamp et al. [70], in which it was detected at five (VVMD7, VVMD32, VChr8a, VChr8b and VChr9a) and four (VMC9a3.1, VVS5, VVMD7, and VrZag79) different loci, respectively. Triallelic profiles at lo-ci VVS2 and VVS5 were detected for Pinot Meunier clone [70], and for Primitivo di Gioia at locus VVS19 [74]. In the review of research, we found that a three-allelic profile has ap-peared several times at locus VVMD7 [68, 69, 75] and, in the last study [68] VVMD7 was shown to carry three alleles in 12 different cultivars out of 16 cultivars showing chimerism.

2.3.2. Cultivar identity/synonyms detection

There are around 10,000 grapevine cultivars held in germplasm collections worldwide [48] but, based on DNA analyses, the number of grapevine varieties is estimated at approx. 5,000 [76]. This proves the need for identifying synonyms and homonyms in collections to remove redundant accessions and improve management.

Local winegrowers in the past cultivated primarily less known varieties, but with the intensive renovation of vineyards these have almost disappeared and been replaced by new varieties grown elsewhere in Europe. Most winegrowing countries have initiated a global campaign of collecting, preserving and evaluating old cultivars and clones, as well as organizing collections. Some of these indigenous or local varieties are particularly promising in terms of high quality but, in the past, for various reasons, have not been adequately exploited. Wine produced today from native varieties provides a new niche in the competitive market. The descriptions of some of these varieties and associated data available are incomplete and it is necessary to identify them or resolve their description. In addition, the populations of *Vitis vinifera* L. are often very heterogeneous and vines of each clone can be very different, which also hampers identification at the morphological level [50]. Diverse historical development and multilingual areas have contributed to differences in the naming of local varieties, which have resulted in the high number of synonyms and homonyms [76].

Laucou et al. [50] recently published data that, among 4,370 accessions maintained in the IN-RA germplasm collection, 1,050 cases of questionable synonyms were discovered or confirmed. Santana et al. [48] found 300 synonymic samples among 421 Spanish grapevines and Cipriani et al. [49] 260 out of 1005 international, national and local grapevine accessions. Tomić [68] reported 58 synonyms out of 196 samples included in SSR analysis, discovering 20 groups of synonyms and 12 groups of homonyms associated with the wrong description due to local denominations. In the latest study [68] cultivar identification was performed also by comparing the set of 138 unique profiles (without redundant genotypes) with approximately 2000 other grape genotypes grown in Europe (personal communication with Vouillamoz, Jose) and 15 groups of synonyms and 3 groups of homonyms were found. Comparison of Slovenian genotypes [65] with 161 European varieties described by Sefc et al. [46] helped to identify 3 new pairs of synonyms: Volovnik = Vela Pergolla (Croatia), Pregarc = Garnache Tintorera (Spanish) and Kanarjola = Trebbiano Toscano (Italian).

Microsatellite similarity analysis of Slovenian genotypes [65, 75] confirmed some suspicions of identical varieties made on the basis of morphological characters; such as the variety Ferjanščkova, which was shown to be a synonym for Merlot and Grganc a synonym for Rebula, which means that the ancient names have apparently been preserved in some areas in Slovenian Istria. A group of five varieties (Glera = Prosecco = Briška Glera = Števerjana = Beli teran) is among synonyms that were *de novo* obtained by analysis of microsatellites; some of them have been previously described based on morphological similarity, while, for example, Števerjana is a new synonym of these varieties. A high diversity of microsatellite loci was detected between the varieties Briška Glera and White Glera, which could be explained by the fact that the Glera name was often used in the past for a variety of white grapevine varieties grown in the sub-Mediterranean part of Slovenia. Another group of homonyms represent varieties called Ribolla (Rebula, Old Rebula and Rebula-100 years) also revealing high polymorphism among them. A comparison of genotypes of Slovenian [65, 75] and Croatian varieties [77], performed on the basis of 7 microsatellite loci, also revealed synonyms between Muscat Ruža Porečki (Croatia) and Cipro (Slovenia) and between Ranfol bijeli (Croatia) and Belina Pleterje (Slovenia). Homonymy was detected between the Croatian variety Plavina described by Calo et al. [78] and the Slovenian variety with the same name, although their similarity based

on SSR analysis was only 20% [65]. Varieties called Pagadebiti from Slovenia [65], Croatia [78] and Italy [79] also revealed very different SSR-allelic profiles.

2.3.3. Genetic relatedness, structure and parentage

Additional important applications of genotyping are analyses of genetic variability, genetic structure and parentage. When the data are comparable between different studies or within a larger group of cultivars, it is possible to identify the origin and relationships of cultivars. For example, a comparison of Slovenian genotypes with 161 genotypes [46] from eight European winegrowing regions showed that they are more related to Croatian and Greek varieties than to those from the adjacent Italian peninsula, but most genetically distant from French varieties, which may be a result of maritime trade across the Mediterranean Sea or along commercial routes through the Balkans [65].

Genetic clustering of varieties from the Castilian Plateau of Spain revealed three differentiated grups: Muscat-type accessions and interspecific Vitis hybrids, accessions from France and the western Castilian Plateau, and accessions from the central Castilian Plateau together with local table grapes. The close relatedness of accessions from the western plateau among each other and to French varieties suggested the introduction of the latter along the pilgrimage route to Santiago de Compostela [48].

Analysis of genetic relatedness of Balkan genotypes [68] showed that genotypes from Serbia, Bosnia and Slovenia are genetically fairly similar to each other, while genotypes from Macedonia and Montenegro are genetically more distant from the rest.

Microsatellite analysis and grouping of 1005 international, national and local grapevine accessions resulted in a weak correlation with their geographical origin and/or current area of cultivation, showing a large admixture of local varieties with those most widely cultivated, as a result of ancient commerce and population flows [49].

2.4. *Vitis* microsatellite databases

The main purpose of assembling data in databases with open access is to enlarge the number of varieties available for comparison and to facilitate the identification of genotypes. The largest international Vitis Microsatellite Collection is currently available within the European Vitis Database, which was constructed within the context of the European projects Genres081, GrapeGen06 and maintains SSR-marker data of 4364 accessions evaluated at 9 SSR loci [80]. High priority in these projects was given to the trueness-to-type of valuable and unique genotypes and a prerequisite for true-to-type identification is analysis of identity based on microsatellites. SSR-marker data within this database can be retrieved in two ways; search by cultivars or search by allele lengths. The database also includes SSR-marker data of 46 reference varieties, which enables comparison of data from different laboratories. The database has open access to partners providing SSR-marker data.

Some minor databases also exist, such as the publicly available Swiss microsatellite database (SVMD) [81], which includes 170 domestic and foreign genotypes growing in the given area and their SSR data for six microsatellite loci (VVMD5, VVMD7, VVMD27, VVS2, VrZAG62

and VrZAG79). A Greek collection (Greek Vitis Microsatellite Database) includes all possible information about grapevines that grow in Greece and is a combination of two older ampelographic databases, supplemented by microsatellite data (298 varieties and rootstocks) [82]. The Italian database (GMC - Grape Microsatellite Collection) provides a complete overview of microsatellite analysis of grapevine performed in different laboratories/countries and also includes information on authors and methods of work [83].

The reference varieties presented in the database are prerequisite for the comparison of data revealed from different systems/laboratories. Due to different electrophoresis systems, a difference between the lengths of alleles (shift of relative allele length) can be detected and data needs to be standardized. The length of alleles can be changed or standardized, so that analysis of genotyping includes some of the reference samples on which to compare 'unknown' samples and their allele lengths can be adjusted. Differences in allele lengths are the same within each locus, so reference samples included in the analysis can be used as a base to standardize all 'unknown' samples [61]. Information on allele lengths obtained in different laboratories can thus be compared and combined into a common database. An alternative for grapevine genotyping, where the complete genome sequence is available, is identification of thousands of single nucleotide polymorphisms (SNP), which can be very useful for genotyping purposes, since they can be multiplexed and need no standardization of results with additional reference cultivars. Because SNP markers are bi-allelic, genotypes obtained with different equipment and by different laboratories are always fully comparable [84].

3. Application of microsatellite markers in olives

Microsatellite markers or SSR (simple sequence repeats) have found wide applications in genetic studies of olives, including cultivar identification, assessment of genetic diversity in different sets of genotypes, evaluation of relationships among olive cultivars and among cultivated and wild olives, designation of geographic origin, genetic mapping, construction of core collections and similar studies.

This contribution presents a short review of SSR marker application in olives.

3.1. Microsatellite marker development

In view of their characteristics (high abundance and random in genome, high polymorphism, co-dominant inheritance, locus specific) SSR are desirable markers in plant genetic studies, although considerable input is required for initial marker development. The main features of SSR marker development in olive is summarized in Table1. The first SSR markers in olives were developed in 2000 by two groups. Sefc et al. [25] constructed a genomic library using the DNA of three Portuguese olive cultivars for the identification of SSR loci. The genomic library was probed by $(GA)_n$ and $(CA)_n$ repeats and 28 microsatellite containing sequences suitable for primer development were found and 15 SSR loci gave specific amplifications under optimized PCR conditions. These markers were designated ssrOeUA-DCA, followed by a two digit number, in short, a DCA series. Markers were tested on 48

Iberian and Italian olive trees for the number of amplified alleles (on average 8.3 alleles per primer pair), and observed and expected heterozygosity (Ho and He) showed the character‐ istics of each SSR marker (Table 1). The second group [33] developed 5 SSR markers out of 13 microsatellite loci obtained from a GA-enriched genomic library. The 5 SSR markers were tested on a set of 46 olive cultivars for their characteristics, giving an average of 5.2 alleles per marker. They were were designated IAS-olio, followed by a two digit number. Three new series of SSR markers for olives followed in 2002. Carriero et al. [31] developed 20 SSR markers out of a highly $(GA)_n$ enriched genomic library and 10 markers were further charac‐ terized on twenty olive cultivars, amplifying 5.7 alleles per marker. These markers are desig‐ nated GAPU, followed by a three digit number. Six SSR markers (EMO, followed by two digits) derived from a $(GA)_n$ and $(CA)_n$ enriched genomic library and one marker (EMOL) developed from a gene sequence containing a $(GA)_n$ microsatellite motif, were tested on 23 olive cultivars, giving an amplification of 6.1 alleles per primer pair [32]. Three of these markers also amplified microsatellite alleles in other species of Oleaceae, showing their transferability. Cipriani et al. [85] published 30 SSR markers designated UDO99-, followed by three digits but they are usually designated UDO-two digits. These markers were tested on a small set of 12 olive cultivars, amplifying 1-7 (average 3.6) alleles per primer pair and five markers gave an allelic profile of duplicated loci. A Spanish group undertook the devel‐ opment of a second set of IAS-olio SSR markers [86]. Primer pairs were designed for 24 mi‐ crosatellite containing sequences, of which 12 loci gave an amplification product of expected profile; 10 markers gave a single locus amplification and 2 markers duplicated loci, con‐ firmed by segregation analysis. Markers were characterized on a set of 51 olive cultivars, giving on average 5.6 alleles per locus. The most recent 12 SSR markers for olives were de‐ veloped by Gil et al. [87], which were characterized on 33 olive cultivars giving an average of 6.75 alleles per locus. These SSR markers were designated ssrOeIGP, followed by digits.

3.2. Comparison of developed markers

The developed olive SSR markers, particularly those from 2000 and 2002, have been used by various research groups working on olives. The choice of markers from the literature was mainly based on the researchers' selection based on their own experimental results, usually testing the SSR markers on a small set of genotypes and then selecting the markers with the best performance in terms of single locus amplification, stutter of bands, weak amplification of longer alleles, stability of repeats and number of alleles per marker [88] or by the SSR marker characteristics (number of alleles, Ho and He, polymorphic information content (PIC) or discrimination power (DP) provided in the literature. The citation index (Table 1) gives an idea of the most frequently used SSR markers in olives.

However, comparison of the allelic profiles of olive cultivars across different studies has been hindered by the use of different sets of markers and experimental conditions, resulting in discrepancies in allele size assignment. Bandelj et al. [89] carried out one of the first identifi‐ cations of 19 olive cultivars by SSR markers, using a sequencing gel for allele separation and silver staining. The allele sizes were determined by 10 bp size ladder and sequencing reac‐ tion. Sarri et al. [90] used the same nine SSR markers in an analysis of 118 olive cultivars,

separating the alleles with sequencing apparatus and sizing them with computer software. A comparison of allele sizes at the same loci between those two works shows discrepancies of 1-2bp per allele, making it difficult to decide whether an allele is 238 bp, 239 bp or 240 bp long. Allele size discrepancy is also reflected in the genotypes of a particular cultivar analysed in different laboratories. For example, the cultivar Arbequina was genotyped at eight common SSR loci by Bandelj et al. [89] and Doveri et al. [91] but showed no match at any loci.

A first attempt to provide some common SSR markers for olive cultivar identification and discrimination was reported by Doveri et al. [91]. Four partner laboratories tested eight SSR markers from the DCA series, on seventeen selected cultivars using ABI and LICOR systems for fragment analysis and allele sizing. The allele sizes of each marker from the different laboratories were harmonized by comparison and by the use of three cultivars with standard alleles for each loci. Markers DCA3, DCA8, DCA11, DCA13, DCA14 and DCA15 were assessed as the most reproducible among the four laboratories, stressing that reproducibility depends on the use of the same source of plant material, the same reference cultivars and standardization of analytical conditions. Baldoni et al. [92] later published the most comprehensive evaluation of available SSR markers and produced a consensus list of 11 SSR markers for olive genotyping. Thirty-seven SSR markers were tested for reproducibility (low stutter, strong peak signal, single loci amplification and no null alleles) on a set of 21 cultivars, among four laboratories, three using a capillary sequencer (two labs MegaBACE 1000, one lab ABI3130) and one a 2100 Bioanalyzer Agilent. Up to 5 bp discrepancies in allele size were observed among the labs, mainly due to the use of different sequencers and internal allele references. They selected 11 SSR markers for which an allelic ladder at each locus is provided. Alleles were further sequenced to estimate the true size and to characterize the repeat motifs and mapped such that only unlinked loci were selected. The selected markers, ranked by their information value UDO-043, DCA9, GAPU103A, DCA18, DCA16, GAPU101, DCA3, GAPU71B, DCA5, DCA14 and EMO-90, were further tested on a larger set of 77 cultivars to calculate their genetic parameters. This consensus list of SSR markers, together with allelic references, provides a solid platform for olive genotyping by different labs, enabling inter-lab comparison and the construction of an SSR database of olive genotypes, which would be of great help for true-to-type cultivar identification and management of olive germplasm banks.

3.3. Application of SSR markers

Olive trees have been grown for oil and table olive production in the Mediterranean basin since ancient times. The genetic diversity of cultivated olives is abundant and is characterized by a numerous local cultivars vegatatively propagated by farmers. Bartolini et al. [93] collected information on more than 1,208 cultivars from 52 countries, conserved in 94 collections. The number of cultivars is probably much higher, bearing in mind the lack of information on minor cultivars in different olive growing regions. Cultivar surveys have been initiated in many olive growing countries in order to describe existing cultivars, thus obtaining information for germplasm preservation, description of cultivars of specific growing regions and for breeding purposes. For the description and management of the existing

genetic diversity in olives, molecular markers have been found to be particularly valuable because of such characteristics as high genetic informativeness, environmental independence, relatively easy use and the possibility of accumulating a large amount of data. SSR markers, in particular, have been extensively used in olives for cultivar identification, assessment of genetic diversity and other genetic studies.

Cultivar identification in olives is important to confirm true-to-type denominated cultivars, solve problems relating to synonyms, homonyms and mislabelled planting material. One of the first cultivar identifications by SSR markers was done on a small set of Slovene and Italian [19] cultivars using the DCA series [89]. The work was extended to practical application of SSR markers for confirmation of true-to type denomination of 13 olive samples from nursery using two DCA markers [94]. Comparison of the genotyped samples with the genotypes of reference cultivars obtained from three collections enabled confirmation of the correct denomination of six samples, 5 samples were mislabelled and no reference cultivar was available for two samples. Thirty-five Spanish and Italian olive cultivars of commercial interest were then genotyped by UDO series [95]. Olive cultivars were further genotyped for identification purposes or for assessment of genetic diversity on international (world germplasm collections), national (Spain, Italy, Tunis, Morocco, Turkey, Greece, Croatia, Slovenia, Portugal, Lebanon, Alger) and regional scales (olive growing region with characteristic variety structure). There have been numerous publications from these studies and we present here only a few examples. Sarri et al. [90] genotyped 118 olive cultivars from several Mediterranean countries by use of twelve SSR markers (10 DCA series, GAPU89 and UDO12) showing high discrimination power. A combination of only three markers distinguished almost all analysed cultivars and a selection of six markers was sufficient to assign cultivars to their geographic origin, divided into eastern, central and western Mediterranean. Geographic structuring of diversity was also found in a set of 211 autochthonous cultivars in six southern Italian olive growing regions [96]. The cultivars were analysed by 11 SSR loci (DCA, GAPU and UDO), which discriminated 199 unique genotypes and identified ten pairs of synonyms, four cases of homonyms and a possible parent-offspring relationship. Poljuha et al. [97] analysed 27 olive accessions from an olive growing region in Croatia and Slovenia (Istria), using 12 SSR markers (DCA) and finding a distinction between native and introduced cultivars, as well as some cases of synonyms and homonyms.

Khadari et al. [98] analysed 215 olive trees sampled in all Moroccan traditional growing regions. Using 15 SSR (4 DCA, 3 GAPU and 8 UDO) they, identified 60 SSR profiles among which 52 genotypes belonged to cultivated trees with no denomination, demonstrating high genetic diversity in Moroccan olive germplasm. However, a single Moroccan cultivar, belonging to a different gene pool to local cultivars, which were probably derived from local domestication, was predominant in all growing regions. Local olive domestication in two out of three sampled olive growing regions in Spain was also suggested by Belaj et al. [99] in a study of the relationship between wild and cultivated olives using eight SSR markers (4 DCA, 3 UDO and EMO). A low level of local olive domestication was found in a study of Sardinian wild (21), local (22) and ancient cultivars (35) [100], using 6 DCA, 4 UDO and 3 GAPU SSR markers, however most of the Sardinian local cultivars were also very closely related to ancient cultivars analysed. The relationship between ancient olive trees and cultivars in Southern Spain

is slightly different, since only 9.6% of 106 ancient trees matched olive cultivars, as revealed by analysis using 14 SSR markers (7 DCA, 2 GAPU and 5 UDO) [101].

Several cases of synonyms, homonyms and mislabelled samples, as well as high diversity were revealed in a survey of 84 accessions from a Tunisian germplasm collection, using eight SSR markers of series DCA (5), GAPU(2) and UDO. On the basis of the SSR analysis, an improved classification of accessions was proposed for better management of the germplasm collection [102].

Proper management of germplasm collections in terms of evaluation, documentation, regeneration and effective use of available genetic diversity present in a collection is hindered by the large sizes of collections, redundancy and lack of accession information. In order to overcome these problems, core collections have been established that contain a limited number of accessions, capturing maximum allelic diversity. There are two world olive germplasm collections, one in Cordoba, Spain (C1) and the other in Marrakech (M), Morocco, which have in common 153 accessions and both core collections have been established using SSR markers for measuring genetic diversity [103, 104]. In the Marrakech collection, 561 accessions were analysed by 12 SSR markers (8 DCA, 2 GAPU, 1 UDO, 1 EMO) and the estimated core collection comprises 67 accessions; a slightly lower number [56] of accessions to represent the total allelic diversity was estimated in the Cordoba collection on the basis of analysing 378 accessions with 14 SSR markers (6 DCA, 4 GAPU, 4 UDO, 1 EMO). The Cordoba (C2) collection of 361 accessions was additionally assessed with 23 SSR markers (5 DCA, 6GAPU, 8 UDO, 1 EMO, 3 GP) as well as DaRT, SNP and morphological markers and their estimate for a core collection adequate for conservation of genetic diversity was 68 accessions [105]. Enormous work was carried out in genotyping all these accessions. However, the set of SSR markers used were unfortunately selected arbitrarily. Seven SSR markers were the same in M and C1 collections but only three and one SSR markers were in common with the C2 collection, respectively. In comparison with the Baldoni et al. [92] recommended list of SSR markers, the C1 collection had in common 8 markers, the M collection 6 and the C2 collection only 2 SSR markers. The advantages of SSR markers, which enable inter-laboratory comparison, in these cases not really fully exploited, since not only the same markers but also harmonized protocols are needed for reliable comparison of analysed genotypes.

In conclusion, SSR markers have been proven through numerous applications to be a very powerful tool in studies of olive genetic structure, domestication processes, genetic relationships among different cultivars, wild and cultivated olives, in the management of germplasm collection etc.

Some sort of agreement on the use of SSR markers and protocols should be reached in the future, which would allow inter-laboratory comparisons and, most importantly, the establishment of an international olive microsatellite database.

SSRs series	Source	Screening	No. SSR loci	No. cvs. tested	No. alleles	No. alleles/ locus	H_o range	H_e range	Reference	Cited (SCI)
DAC	genomic library; DNA cvs: Porto Martins, Terceira, Açores	(GA)n, (CA)n	15	47	124	8.3	0.28-0.98	0.36-0.86	[25]	131
IAS-oli	Enriched genomic library, DNA cv: Arbequina	(GA)n	5 (13)*	46	26	5.2		0.46-0.71	[33]	120
GAPU	Enriched genomic library, DNA cvs: 6 different cultivars	(GA)n	10(20)*	20	57	5.7			[31]	104
EMO	Enriched genomic library, DNA cv: Picual	(GA)n, (CA)n	7	23	43	6.1	0.39-0.91	0.62-0.81	[32]	71
UDO	Enriched genomic library, DNA cv: Frantoio	(AC)n, (AG)n	29 (30)*	12	103	3.6		0.44-0.77	[85]	139
IAS-oli	Enriched genomic library, DNA cv: Arbequina	(GA)n, (GT)n, (ACT)n	10(2)*	51	68	5.6	1-0.82	1-0.94	[86]	26
ssrOeIGP	Enriched genomic library, DNA cv: Lezzo	?	12 (19)*	33	60	6.7	0.42-0.89	0.19-0.81	[87]	12

Table 1. SSR markers development in olives * single locus amplification (monomorphic, two or multiple loci amplification)

Author details

Jernej Jakše[1*], Nataša Štajner[1], Lidija Tomić[2] and Branka Javornik[2*]

*Address all correspondence to: branka.javornik@bf.uni-lj.si

1 University of Ljubljana, Biotechnical Faculty, Slovenia

2 University of Banja Luka Faculty of Agriculture, Bosnia and Herzegovina

References

[1] Litt, M., & Luty, J. A. (1989). A hypervariable microsatellite revealed by invitro amplification of a dinucleotide repeat within the cardiac-muscle actin gene. *American Journal of Human Genetics.*, Mar, 44(3), 397-401.

[2] Weber, J. L., & May, P. E. (1989). Abundant class of human dna polymorphisms which can be typed using the polymerase chain-reaction. *American Journal of Human Genetics.*, Mar, 44(3), 388-96.

[3] Miklos, G. L., & John, B. (1979). Heterochromatin and satellite DNA in man: properties and prospects. *Am J Hum Genet.*, May, 31(3), 264-80.

[4] Armour, J. A. L., Neumann, R., Gobert, S., & Jeffreys, A. J. (1994). Isolation of human simple repeat loci by hybridization selection. *Human Molecular Genetics*, Apr, 3(4), 599-605.

[5] Goldstein, D. B., & Pollock, D. D. (1997). Launching microsatellites: A review of mutation processes and methods of phylogenetic inference. *Journal of Heredity.*, Sep-Oct, 88(5), 335-42.

[6] Schlotterer, C., & Tautz, D. (1992). Slippage synthesis of simple sequence dna. *Nucleic Acids Research*, Jan, 20(2), 211-5.

[7] Chambers, G. K., & Mac, Avoy. E. S. (2000). Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology.*, Aug, 126(4), 455-76.

[8] Hancock, J. M., Goldstein, D. B., & Schlötterer, C. (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms. Microsatellites: Evolution and Applications. *Oxford: Oxford University Press*, 1-9.

[9] Lagercrantz, U., Ellegren, H., & Andersson, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research*, Mar, 21(5), 1111-5.

[10] Morgante, M., & Olivieri, A. M. (1993). PCR-amplified microsatellites as markers in plant genetics. *The Plant journal : for cell and molecular biology*, Jan, 3(1), 175-82.

[11] Wang, Z., Weber, J. L., Zhong, G., & Tanksley, S. D. (1994). Survey of plant short tandem dna repeats. *Theoretical and Applied Genetics*, Apr, 88(1), 1-6.

[12] Akagi, H., Yokozeki, Y., Inagaki, A., & Fujimura, T. (1996). Microsatellite DNA markers for rice chromosomes. *Theoretical and Applied Genetics*, Nov, 93(7), 1071-7.

[13] Kuhl, J. C., Cheung, F., Yuan, Q. P., Martin, W., Zewdie, Y., Mc Callum, J., et al. (2004). A unique set of 11,008 onion expressed sequence tags reveals expressed sequence and genomic differences between the monocot orders Asparagales and Poales. *Plant Cell.*, Jan, 16(1), 114-25.

[14] Temnykh, S., De Clerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., & Mc Couch, S. (2001). Computational and experimental analysis of microsatellites in rice (Oryza sativa L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Research.*, Aug, 11(8), 1441-52.

[15] Matsumoto, T., Wu, J. Z., Kanamori, H., Katayose, Y., Fujisawa, M., Namiki, N., et al. (2005). The map-based sequence of the rice genome. *Nature.*, Aug, 436(7052), 793-800.

[16] Cardle, L., Ramsay, L., Milbourne, D., Macaulay, M., Marshall, D., & Waugh, R. (2000). Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics.*, Oct, 156(2), 847-54.

[17] Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., et al. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature.*, Dec, 408(6814), 796-815.

[18] Morgante, M., Hanafey, M., & Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics*, Feb, 30(2), 194-200.

[19] Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature.*, Sep, 449(7161), 463-U5.

[20] Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D. A., Cestaro, A., Pruss, D., et al. (2007). A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety. *Plos One.*, Dec, 2(12).

[21] Alagna, F., D'Agostino, N., Torchia, L., Servili, M., Rao, R., Pietrella, M., et al. (2009). Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development. *Bmc Genomics.*, Aug, 10.

[22] Maniatis, T., Sambrook, J., Fritsch, E. F., Cold, Spring., & Harbor, L. (1982). Molecular cloning : a laboratory manual / T. Maniatis, E.F. Fritsch, J. Sambrook. *Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory*.

[23] Bowers, J. E., Dangl, G. S., Vignani, R., & Meredith, CP. (1996). Isolation and characterization of new polymorphic simple sequence repeat loci in grape (Vitis vinifera L). *Genome.*, Aug, 39(4), 628-33.

[24] Thomas, M. R., & Scott, N. S. (1993). Microsatellite repeats in grapevine reveal dna polymorphisms when analyzed as sequence-tagged sites (stss). *Theoretical and Applied Genetics.*, Sep, 86(8), 985-90.

[25] Sefc, K. M., Lopes, S., Mendonca, D., Dos, Santos. M. R., Machado, M. L. D., & Machado, A. D. (2000). Identification of microsatellite loci in olive (Olea europaea) and their characterization in Italian and Iberian olive trees. *Molecular Ecology*, Aug, 9(8), 1171-3.

[26] Ostrander, E. A., Jong, P. M., Rine, J., & Duyk, G. (1992). Construction of small-insert genomic dna libraries highly enriched for microsatellite repeat sequences. *Proceedings of the National Academy of Sciences of the United States of America*, Apr, 89(8), 3419-23.

[27] Hamilton, M. B., Pincus, E. L., Di Fiore, A., & Fleischer, R. C. (1999). Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques*, Sep, 27(3), 500-507.

[28] Kijas, J. M. H., Fowler, J. C. S., Garbett, C. A., & Thomas, M. R. (1994). Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Biotechniques.*, Apr, 16(4), 656-662.

[29] Edwards, K. J., Barker, J. H. A., Daly, A., Jones, C., & Karp, A. (1996). Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques*, May, 20(5), 758-760.

[30] Karagyozov, L., Kalcheva, I. D., & Chapman, V. M. (1993). Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. *Nucleic Acids Research*, Aug, 21(16), 3911-2.

[31] Carriero, F., Fontanazza, G., Cellini, F., & Giorio, G. (2002). Identification of simple sequence repeats (SSRs) in olive (Olea europaea L.). *Theoretical and Applied Genetics*, Feb, 104(2-3), 301-307.

[32] De La Rosa, R., James, C. M., & Tobutt, K. R. (2002). Isolation and characterization of polymorphic microsatellites in olive (Olea europaea L.) and their transferability to other genera in the Oleaceae. *Molecular Ecology Notes*, Sep, 2(3), 265-7.

[33] Rallo, P., Dorado, G., & Martin, A. (2000). Development of simple sequence repeats (SSRs) in olive tree (Olea europaea L.). *Theoretical and Applied Genetics*, Oct, 101(5-6), 984-989.

[34] Lefort, F., Kyvelos, C. J., Zervou, M., Edwards, K. J., & Roubelakis-Angelakis, K. A. (2002). Characterization of new microsatellite loci from Vitis vinifera and their conservation in some Vitis species and hybrids. *Molecular Ecology Notes*, Mar, 2(1), 20-1.

[35] Malausa, T., Gilles, A., Meglecz, E., Blanquart, H., Duthoy, S., Costedoat, C., et al. (2011). High-throughput microsatellite isolation through 454 GS-FLX Titanium pyro-sequencing of enriched DNA libraries. *Molecular Ecology Resources*, Jul, 11(4), 638-44.

[36] Santana, Q. C., Coetzee, M. P. A., Steenkamp, E. T., Mlonyeni, O. X., Hammond, G. N. A., Wingfield, M. J., et al. (2009). Microsatellite discovery by deep sequencing of enriched genomic libraries. *Biotechniques.*, Mar, 46(3), 217-23.

[37] Sharma, P. C., Grover, A., & Kahl, G. (2007). Mining microsatellites in eukaryotic genomes. *Trends in Biotechnology*, Nov, 25(11), 490-8.

[38] Cipriani, G., Marrazzo, M. T., Di Gaspero, G., Pfeiffer, A., Morgante, M., & Testolin, R. (2008). A set of microsatellite markers with long core repeat optimized for grape (Vitis spp.) genotyping. *Bmc Plant Biology.*, Dec, 8.

[39] Scott, K. D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E. M., Lee, L. S., et al. (2000). Analysis of SSRs derived from grape ESTs. *Theoretical and Applied Genetics.*, Mar, 100(5), 723-6.

[40] Echt, C. S., May, Marquardt. P., Hseih, M., & Zahorchak, R. (1996). Characterization of microsatellite markers in eastern white pine. *Genome.*, Dec, 39(6), 1102-8.

[41] 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. *Molecular Breeding*, May, 15(4), 349-66.

[42] Hall, J. M., Le Duc, C. A., Watson, A. R., & Roter, A. H. (1996). An approach to high-throughput genotyping. *Genome Research.*, Sep, 6(9), 781-90.

[43] Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, Feb, 18(2), 233-4.

[44] Cipriani, G., Frazza, G., Peterlunger, E., & Testolin, R. (1994). Grapevine fingerprinting using microsatellite repeats. *Vitis.*, Dec, 33(4), 211-5.

[45] Sefc, K. M., Regner, F., Glossl, J., & Steinkellner, H. (1998). Genotyping of grapevine and rootstock cultivars using microsatellite markers. *Vitis.*, Mar, 37(1), 15-20.

[46] Sefc, K. M., Lopes, M. S., Lefort, F., Botta, R., Roubelakis-Angelakis, K. A., Ibanez, J., et al. (2000). Microsatellite variability in grapevine cultivars from different European regions and evaluation of assignment testing to assess the geographic origin of cultivars. *Theoretical and Applied Genetics.*, Feb, 100(3-4), 498-505.

[47] Sanchez-Escribano, E. M., Martin, J. R., Carreno, J., & Cenis, J. L. (1999). Use of sequence-tagged microsatellite site markers for characterizing table grape cultivars. *Genome*, Feb, 42(1), 87-93.

[48] Santana, J. C., Heuertz, M., Arranz, C., Rubio, J. A., Martinez-Zapater, J. M., & Hidalgo, E. (2010). Genetic Structure, Origins, and Relationships of Grapevine Cultivars

from the Castilian Plateau of Spain. *American Journal of Enology and Viticulture*, 61(2), 214-24.

[49] Cipriani, G., Spadotto, A., Jurman, I., Di Gaspero, G., Crespan, M., Meneghetti, S., et al. (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. Nov. *Theoretical and Applied Genetics*, 121(8), 1569-85.

[50] Laucou, V., Lacombe, T., Dechesne, F., Siret, R., Bruno, J. P., Dessup, M., et al. (2011). High throughput analysis of grape genetic diversity as a tool for germplasm collection management. *Theoretical and Applied Genetics*, Apr, 122(6), 1233-45.

[51] Eisen, J. A. (1999). Mechanistic basis of microsatellite instability. *Goldstein DB, Schlotterer C, editors. Microsatellites: Evolution and Applications. Oxford: Oxford University Press*.

[52] Faria, M. A., Magalhaes, R., Ferreira, M. A., Meredith, C. P., & Monteiro, F. F. (2000). Vitis vinifera must varietal authentication using microsatellite DNA analysis (SSR). *Journal of Agricultural and Food Chemistry.*, Apr, 48(4), 1096-100.

[53] Zulini, L., Russo, M., & Peterlunger, E. (2002). Genotyping wine and table grape cultivars from Apulia (Southern Italy) using microsatellite markers. *Vitis*, 41(4), 183-7.

[54] Di Rienzo, A., Donnelly, P., Toomajian, C., Sisk, B., Hill, A., Petzl-Erler, M. L., et al. (1998). Heterogeneity of microsatellite mutations within and between loci, and implications for human demographic histories. *Genetics.*, Mar, 148(3), 1269-84.

[55] Bowers, J. E., Dangl, G. S., & Meredith, C. P. (1999). Development and characterization of additional microsatellite DNA markers for grape. *American Journal of Enology and Viticulture*, 50(3), 243-6.

[56] Sefc, K. M., Regner, F., Turetschek, E., Glossl, J., & Steinkellner, H. (1999). Identification of microsatellite sequences in Vitis riparia and their applicability for genotyping of different Vitis species. *Genome.*, Jun, 42(3), 367-73.

[57] Lefort, F., & Roubelakis-Angelakis, K. K. A. (2001). Genetic comparison of Greek cultivars of Vitis vinifera L. by nuclear microsatellite profiling. *American Journal of Enology and Viticulture*, 52(2), 101-8.

[58] Arroyo-Garcia, R., & Martinez-Zapater, J. M. (2004). Development and characterization of new microsatellite markers for grape. *Vitis*, 43(4), 175-8.

[59] 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. *Theoretical and Applied Genetics.*, Sep, 109(5), 1017-27.

[60] Di Gaspero, G., Cipriani, G., Marrazzo, M. T., Andreetta, D., Castro, M. J. P., Peterlunger, E., et al. (2005). Isolation of (AC)n-microsatellites in Vitis vinifera L. and anal-

ysis of genetic background in grapevines under marker assisted selection. *Molecular Breeding.*, Jan, 15(1), 11-20.

[61] This, P., Jung, A., Boccacci, P., Borrego, J., Botta, R., Costantini, L., et al. (2004). Development of a standard set of microsatellite reference alleles for identification of grape cultivars. *Theoretical and Applied Genetics.*, Nov, 109(7), 1448-58.

[62] Sefc, K. M., Lefort, F., Grando, Scott. K., Steinkellner, H., & Thomas, M. (2001). Microsatellite markers for grapevine: A state of the art. *Roubelakis-Angelakis KA, editor. Amsterdam: Kluwer Publishers*, 407-438.

[63] Lopes, M. S., Sefc, K. M., Dias, E. E., Steinkellner, H., Machado, M. L. D., & Machado, A. D. (1999). The use of microsatellites for germplasm management in a Portuguese grapevine collection. *Theoretical and Applied Genetics.*, Aug, 99(3-4), 733-9.

[64] Ibanez, J., de Andres, M. T., Molino, A., & Borrego, J. (2003). Genetic study of key Spanish grapevine varieties using microsatellite analysis. *American Journal of Enology and Viticulture*, 54(1), 22-30.

[65] Stajner, N., Rusjan, D., Korosec-Koruza, Z., & Javornik, B. (2011). Genetic Characterization of Old Slovenian Grapevine Varieties of Vitis vinifera L. by Microsatellite Genotyping. *American Journal of Enology and Viticulture*, 62(2), 250-5.

[66] Tomić, L. (2012). Molecular characterization and analysis of the genetic relatedness of old grapevine (Vitis vinifera L.) cultivars from the Western Balkan. *Ljubljana.*

[67] Vouillamoz, J. F., Maigre, D., & Meredith, CP. (2004). Identity and parentage of two alpine grape cultivars from Switzerland (Vitis vinifera L. Lafnetscha and Himbertscha). *Vitis*, 43(2), 81-7.

[68] Tomić, L. (2012). Molecular characterization and analysis of the genetic relatedness of old grapevine (Vitis vinifera L.) cultivars from the Western Balkan [doctoral dissertation]. *Ljubljana: University of Ljubljana.*

[69] Hocquigny, S., Pelsy, F., Dumas, V., Kindt, S., Heloir, M. C., & Merdinoglu, D. (2004). Diversification within grapevine cultivars goes through chimeric states. *Genome.*, Jun, 47(3), 579-89.

[70] Stenkamp, S. H. G., Becker, M. S., Hill, B. H. E., Blaich, R., & Forneck, A. (2009). Clonal variation and stability assay of chimeric Pinot Meunier (Vitis vinifera L.) and descending sports. *Euphytica.*, Jan, 165(1), 197-209.

[71] Riaz, S., Garrison, K. E., Dangl, G. S., Boursiquot, J. M., & Meredith, CP. (2002). Genetic divergence and chimerism within ancient asexually propagated winegrape cultivars. *Journal of the American Society for Horticultural Science*, Jul, 127(4), 508-14.

[72] Ibanez, J., de Andres, M. T., & Borrego, J. (2000). Allelic variation observed at one microsatellite locus between the two synonym grape cultivars Black Currant and Mavri Corinthiaki. *Vitis*, Dec, 39(4), 173-4.

[73] Zulini, L., Fabro, E., & Peterlunger, E. (2005). Characterisation of the grapevine culti-var Picolit by means of morphological descriptors and molecular markers. *Vitis*, 44(1), 35-8.

[74] Franks, T., Botta, R., & Thomas, M. R. (2002). Chimerism in grapevines: implications for cultivar identity, ancestry and genetic improvement. *Theoretical and Applied Genet-ics.*, Feb, 104(2-3), 192-9.

[75] Stajner, N., Korosec-Koruza, Z., Rusian, D., & Javornik, B. (2008). Microsatellite geno-typing of old Slovenian grapevine varieties (Vitis vinifera L.) of the Primorje (coastal) winegrowing region. *Vitis*, 47(4), 201-4.

[76] This, P., Lacombe, T., & Thomas, M. R. (2006). Historical origins and genetic diversity of wine grapes. *Trends in Genetics*, Sep, 22(9), 511-9.

[77] Maletic, E., Sefc, K. M., Steinkellner, H., Kontic, J. K., & Pejic, I. (1999). Genetic char-acterization of Croatian grapevine cultivars and detection of synonymous cultivars in neighboring regions. *Vitis.*, Jun, 38(2), 79-83.

[78] Calo, A., Costacurta, A., Maras, V., Meneghetti, S., & Crespan, M. (2008). Molecular correlation of Zinfandel (Primitivo) with Austrian, Croatian, and Hungarian culti-vars and Kratosija, an additional synonym. *American Journal of Enology and Viticul-ture*, 59(2), 205-9.

[79] Muganu, M., Dangl, G., Aradhya, M., Frediani, M., Scossa, A., & Stover, E. (2009). Ampelographic and DNA Characterization of Local Grapevine Accessions of the Tuscia Area (Latium, Italy). *American Journal of Enology and Viticulture*, 60(1), 110-5.

[80] The European Vitis Database [Internet]. Available from:, http://www.eu-vitis.de/index.php.

[81] Swiss Vitis microsatellite database [Internet]. Available from:, http://www1.unine.ch/svmd/.

[82] Greek Vitis Database [Internet]. Available from:, http://gvd.biology.uoc.gr/gvd/contents/databases/index.htm.

[83] Grape Microsatellite Collection [Internet]. Available from:, http://meteo.iasma.it/genetica/gmc.html.

[84] Cabezas, J. A., Ibanez, J., Lijavetzky, D., Velez, D., Bravo, G., Rodriguez, V., et al. (2011). A 48 SNP set for grapevine cultivar identification. *Bmc Plant Biology*, Nov, 11.

[85] Cipriani, G., Marrazzo, M. T., Marconi, R., Cimato, A., & Testolin, R. (2002). Microsa-tellite markers isolated in olive (Olea europaea L.) are suitable for individual finger-printing and reveal polymorphism within ancient cultivars. *Theoretical and Applied Genetics.*, Feb, 104(2-3), 223-8.

[86] Diaz, A., De la Rosa, R., Martin, A., & Rallo, P. (2006). Development, characterization and inheritance of new microsatellites in olive (Olea europaea L.) and evaluation of

their usefulness in cultivar identification and genetic relationship studies. *Tree Genetics & Genomes.*, Jul, 2(3), 165-75.

[87] Gil, F. S., Busconi, M., Machado, A. D., & Fogher, C. (2006). Development and characterization of microsatellite loci from Olea europaea. *Molecular Ecology Notes*, Dec, 6(4), 1275-7.

[88] Bandelj, D., Jakse, J., & Javornik, B. (2004). Assessment of genetic variability of olive varieties by microsatellite and AFLP markers. *Euphytica*, 136(1), 93-102.

[89] Bandelj, D., Jakse, J., & Javornik, B. (2002). DNA fingerprinting of olive varieties by microsatellite markers. *Food Technology and Biotechnology.*, Jul-Sep, 40(3), 185-90.

[90] Sarri, V., Baldoni, L., Porceddu, A., Cultrera, N. G. M., Contento, A., Frediani, M., et al. (2006). Microsatellite markers are powerful tools for discriminating among olive cultivars and assigning them to geographically defined populations. *Genome.*, Dec, 49(12), 1606-15.

[91] Doveri, S., Gil, F. S., Diaz, A., Reale, S., Busconi, M., Machado, A. D., et al. (2008). Standardization of a set of microsatellite markers for use in cultivar identification studies in olive (Olea europaea L.). *Scientia Horticulturae*, May, 116(4), 367-73.

[92] Baldoni, L., Cultrera, N. G., Mariotti, R., Ricciolini, C., Arcioni, S., Vendramin, G. G., et al. (2009). A consensus list of microsatellite markers for olive genotyping. *Molecular Breeding.*, Oct, 24(3), 213-31.

[93] Bartolini, G., Prevost, G., Messeri, C., & Carignani, C. (2005). Olive germplasm: cultivars and world-wide collections Rome: FAO/Plant Production and Protection. [cited 2012, 19 Aug 2012]. Available from:, www.oleadb.it.

[94] Bandelj, D., & Javornik, B. (2007). Microsatellites as a powerfull tool for identification of olive (Olea europaea L.) planting material in nurseries. *Annales, Series Historia Naturalis*, 17, 133-8.

[95] Belaj, A., Cipriani, G., Testolin, R., Rallo, L., & Trujillo, I. (2004). Characterization and identification of the main Spanish and Italian olive cultivars by simple-sequence-repeat markers. *Hortscience.*, Dec, 39(7), 1557-61.

[96] Muzzalupo, I., Stefanizzi, F., & Perri, E. (2009). Evaluation of Olives Cultivated in Southern Italy by Simple Sequence Repeat Markers. *Hortscience*, Jun, 44(3), 582-8.

[97] Poljuha, D., Sladonja, B., Setic, E., Milotic, A., Bandelj, D., Jakse, J., et al. (2008). DNA fingerprinting of olive varieties in Istria (Croatia) by microsatellite markers. *Scientia Horticulturae*, Feb, 115(3), 223-30.

[98] Khadari, B., Charafi, J., Moukhli, A., & Ater, M. (2008). Substantial genetic diversity in cultivated Moroccan olive despite a single major cultivar: a paradoxical situation evidenced by the use of SSR loci. *Tree Genetics & Genomes*, Apr, 4(2), 213-21.

[99]   Belaj, A., Munoz-Diez, C., Baldoni, L., Satovic, Z., & Barranco, D. (2010). Genetic diversity and relationships of wild and cultivated olives at regional level in Spain. *Scientia Horticulturae*, Apr, 124(3), 323-30.

[100] Erre, P., Chessa, I., Munoz-Diez, C., Belaj, A., Rallo, L., & Trujillo, I. (2010). Genetic diversity and relationships between wild and cultivated olives (Olea europaea L.) in Sardinia as assessed by SSR markers. *Genetic Resources and Crop Evolution*, Jan, 57(1), 41-54.

[101] Diez, C. M., Trujillo, I., Barrio, E., Belaj, A., Barranco, D., & Rallo, L. (2011). Centennial olive trees as a reservoir of genetic diversity. *Annals of Botany*, Oct, 108(5), 797-807.

[102] Fendri, M., Trujillo, I., Trigui, A., Rodriguez-Garcia, M. I., & Ramirez, J. D. A. (2010). Simple Sequence Repeat Identification and Endocarp Characterization of Olive Tree Accessions in a Tunisian Germplasm Collection. *Hortscience.*, Oct, 45(10), 1429-36.

[103] Haouane, H., El Bakkali, A., Moukhli, A., Tollon, C., Santoni, S., Oukabli, A., et al. (2011). Genetic structure and core collection of the World Olive Germplasm Bank of Marrakech: towards the optimised management and use of Mediterranean olive genetic resources. *Genetica.*, Sep, 139(9), 1083-94.

[104] Diez, C. M., Imperato, A., Rallo, L., Barranco, D., & Trujillo, I. (2012). Worldwide Core Collection of Olive Cultivars Based on Simple Sequence Repeat and Morphological Markers. *Crop Science*, Jan, 52(1), 211-21.

[105] Belaj, A., Dominguez-Garcia, M. D., Atienza, S. G., Urdiroz, N. M., De la Rosa, R., Satovic, Z., et al. (2012). Developing a core collection of olive (Olea europaea L.) based on molecular markers (DArTs, SSRs, SNPs) and agronomic traits. *Tree Genetics & Genomes.*, Apr, 8(2), 365-78.