

Article

# Genetic Diversity, Population Structure, and Parentage Analysis of Croatian Grapevine Germplasm

Maja Žulj Mihaljević <sup>1,\*</sup> , Edi Maletić <sup>2,3</sup>, Darko Preiner <sup>2,3</sup> , Goran Zdunić <sup>4</sup>,  
Marijan Bubola <sup>5</sup> , Eva Zyprian <sup>6</sup>  and Ivan Pejić <sup>1,3</sup>

<sup>1</sup> Department of Plant Breeding, Genetics and Biometrics, Faculty of Agriculture, University of Zagreb, Svetošimunska 25, 10000 Zagreb, Croatia; ipejic@agr.hr

<sup>2</sup> Department of Viticulture and Enology, Faculty of Agriculture, University of Zagreb, Svetošimunska 25, 10000 Zagreb, Croatia; emaletic@agr.hr (E.M.); dpreiner@agr.hr (D.P.)

<sup>3</sup> Centre of Excellence for Biodiversity and Molecular Plant Breeding (CoE CroP-BioDiv), Faculty of Agriculture, University of Zagreb, Svetošimunska 25, 10000 Zagreb, Croatia

<sup>4</sup> Institute for Adriatic Crops and Karst Reclamation, Put Duilova 11, 21000 Split, Croatia; goran.zdunic@krs.hr

<sup>5</sup> Institute of Agriculture and Tourism, Karla Huguesa 8, 52440 Poreč, Croatia; marijan@iptpo.hr

<sup>6</sup> JKI Institute for Grapevine Breeding Geilweilerhof, 76833 Siebeldingen, Germany; eva.zyprian@julius-kuehn.de

\* Correspondence: mzulj@agr.hr

Received: 29 May 2020; Accepted: 30 June 2020; Published: 2 July 2020



**Abstract:** Croatian viticulture was most extensive at the beginning of the 20th century, when about 400 varieties were in use. Autochthonous varieties are the result of spontaneous hybridization from the pre-phylloxera era and are still cultivated today on about 35 % of vineyard area, while some exist only in repositories. We present what is the most comprehensive genetic analysis of all major Croatian national repositories, with a large number of microsatellite, or simple sequence repeat (SSR) markers, and it is also the first study to apply single nucleotide polymorphism (SNP) markers. After 212 accessions were fingerprinted, 95 were classified as unique to Croatian germplasm. Genetic diversity of Croatian germplasm is rather high considering its size. SNP markers proved useful for fingerprinting but less informative and practical than SSRs. Analysis of the genetic structure showed that Croatian germplasm is predominantly part of the Balkan grape gene pool. A high number of admixed varieties and synonyms is a consequence of complex pedigrees and migrations. Parentage analysis confirmed 24 full parentages, as well as 113 half-kinships. Unexpectedly, several key genitors could not be detected within the present Croatian germplasm. The low number of reconstructed parentages (19%) points to severe genetic erosion and stresses the importance of germplasm repositories.

**Keywords:** grapevine; Croatia; genetic diversity; genetic structure; parentage; synonyms; SSR; SNP

## 1. Introduction

Domesticated grapevine (*Vitis vinifera* L. subsp. *sativa*) is one of the major fruit crops in temperate regions [1]. Most of the world's vineyards are planted with varieties that have been perpetuated for centuries by vegetative propagation [2]. Until the beginning of the 20th century, the vast majority of these cultivars arose from seedlings of mainly unknown origin, and most of today's well-known cultivars are their clonal offspring preserved throughout history [3]. Preference for a limited number of elite cultivars recognized worldwide for wine production and globalization of wine and table grape production, as well as market forces, have led to uniformity and loss of grape biodiversity. The structure

of germplasm has been changing drastically; in the year 2000, half of the world's vineyards were planted to 21 varieties, while, in 2010, this number decreased to only 15 varieties [4].

Estimation of genetic diversity and determination of the genetic relationships among the germplasm accessions, including parentage analysis, are the first steps in dissection of a crop's genetics. Among the molecular markers used for this purpose, microsatellites, or simple sequence repeats (SSRs), have been the most utilized due to their well-documented advantages and are even today the most common system used for discrimination of grape germplasm, and even between individuals in human forensic investigations [5]. Recommendation of the common SSR set for cultivar identification [6,7], as well as integration of the results via a public European *Vitis* database [8], were key steps in the integration of the results obtained by various research groups [7,9–11].

The world's leading grapevine collections have been extensively characterized during the past ~15 years, resolving issues of duplicates, homonyms, synonyms, establishing genetic relationships, and analyzing genetic structure of collected varieties. The largest collection is the National Institute of Agricultural Research (INRA) Vassal Collection in France [12–16]; the second largest is the Spanish Encin, IMIDRA (Instituto Madrileño de Investigación y Desarrollo Rural, Agrario y Alimentario) collection [17,18], followed by the Italian Cra-Vit in Conegliano [19], the German JKI collection at Geilweilerhof [7], the Italian FEM (Fondazione Edmund Mach) collection [20], and the USDA (United States Department of Agriculture) collection [21,22]. Other smaller winegrowing countries or regions have also genotyped their germplasm, primarily for identification of autochthonous varieties, like Austria [23], Portugal [24,25], Greece [26], Hungary [27], Transcaucasia [28], Slovenia [29,30], Bulgaria [31], Serbia [32], Bosnia, and Herzegovina [33], among others. The same has been done through international cooperation, e.g., for seven European countries [34], countries of the Iberian Peninsula [35], Southeast European countries [36], Western Balkan countries [37], and Eastern Europe [9].

The first alternative to SSR markers for fingerprinting was a set of 48 single nucleotide polymorphism (SNP) markers [38]. Although this set has been used by other researchers through direct comparison with IMIDRA's database [39,40], the SNP profiles obtained are not publicly accessible, even though they have been proposed for routine fingerprinting. Large grapevine collections have also genotyped their germplasm with SNP markers [20,41,42], and the most comprehensive SNP analysis of grapevine germplasm [43] is result of a large GrapeReSeq consortium that used the Vitis18KSNP chip and, consequently, suggested the use of 14 most informative SNPs for cultivar distinction.

Assignment of a variety to a certain gene pool helps to define its origin. In order to analyze it, it is necessary to include as much grapevine genetic diversity as possible, with as many informative markers as possible. This has primarily been achieved by genotyping of a larger number of representative cultivars [34,44] or large world collections [15,19,20,22,41,45]. The latest analysis [43] confirmed previously reported main divisions based on geography (Eastern-Western gradient) and human utilization (table vs. wine grape) and refined the structure through eight clusters.

It is very interesting to understand the events from the genetic point of view that resulted in today's assortment. Grapevine varieties are part of each country's cultural heritage and are often economically important. This knowledge is also essential for genome-wide association studies. Scientists have been interested in grape parentage reconstructions and kinship analysis since the late 1990s [46], and, in the last 20 years, full parentage validation or at least half kinship has been suggested in the analysis of large collections [14,19,41,43], also with the help of SNPs. In addition, parentages have been reconstructed in smaller national collections [10,47–52]. The use of chloroplast DNA microsatellites (chloroplast (cp)SSRs) facilitates the investigation of domestication and phylogeny of grape species [53], as well as in defining maternal genotypes in parentage reconstruction.

Viticulture on Croatian territory dates back to Illyrian times [54], prior to the ancient Greek colonization that occurred in 4th century B.C. At the time when Croatian viticulture was most extensive (the late 19th and early 20th century), there were around 200,000 hectares and about 400 varieties in use, with more than 200 varieties in Dalmatia only [55]. Today, more than half of Croatian autochthonous varieties, listed a century ago, have been irreversibly lost due to various (historical)

reasons. Conservation of grapevine genetic resources in Croatia started about 25 years ago and six germplasm collections were established [56]. Biodiversity of Croatian grapevine germplasm has been investigated by SSRs, to a certain extent, in terms of genetic identification and variability and clarification of synonyms and homonyms [36,57–61], genetic similarity to other germplasm [62], partial reconstruction of chlorotypes [53] and parentage reconstruction [63–65], and also partially indirectly as part of worlds' largest collections [14,15].

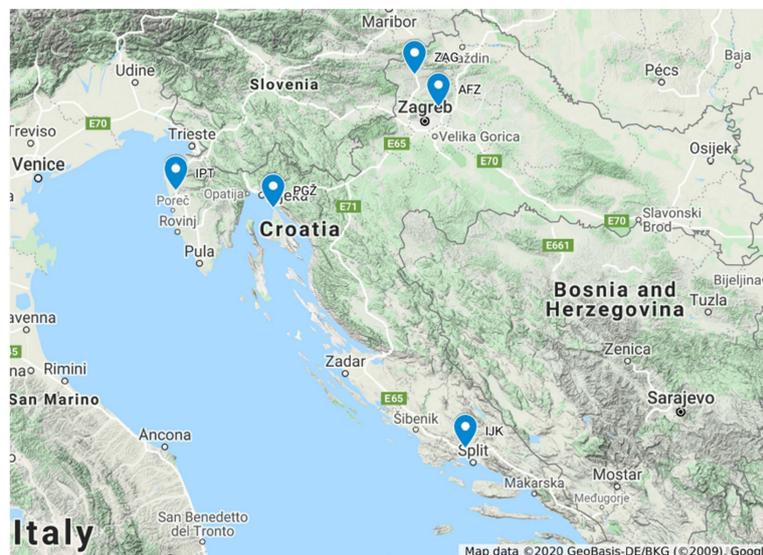
Croatia never had its own domestic grape breeding program, nor is there any documentation of individual intentional crosses made. Croatian autochthonous varieties, which are still in cultivation or exist only in repositories, originate from the pre-phylloxera era. The assumption has been, therefore, that these varieties are mostly the result of spontaneous hybridization between pre-existing and/or introduced varieties in the area of present-day Croatia.

The aim of this study was: (i) to fingerprint all available accessions, clarify synonymy, homonymy, and misnaming within Croatian grape repositories; (ii) to analyze genetic diversity and structure of Croatian cultivars; (iii) to determine and analyze chlorotypes distribution; (iv) and to infer the unknown and confirm the already known pedigrees.

## 2. Materials and Methods

### 2.1. Plant Material

The research included a total of 212 grapevine accessions, collected from all five national and regional grapevine collections, which are managed and maintained by public scientific institutions, with the exception of 7 accessions collected in situ, thus representing all Croatian grape germplasm. Geographical location of the collections within Croatia and their basic facts are presented in Figure 1. A list of grapevine accessions, including their codes, presumed varietal names, skin color, grape utilization, sampling collection, main cultivation region, conservation status, and types of analyses conducted, is given in Table S1.



**Figure 1.** Geographical location and basic information on Croatian grapevine collections: AFZ—National Collection of Autochthonous Varieties, University of Zagreb, Faculty of Agriculture; Experimental station Jazbina, Zagreb; GPS: 45.857860, 16.003847; IJK—Collection of Dalmatian varieties, Institute for Adriatic Crops and Karst Reclamation; Split; Experimental station Duilovo, Split; GPS: 43.505778, 16.497970; IPT—Collection of Istrian varieties, Institute of Agriculture and Tourism Poreč; Poreč; GPS: 45.222686, 13.603673; PGŽ—Collection of Primorje-Gorski Kotar County; Risika; GPS: 45.097715, 14.640855; ZAG—Hrvatsko Zagorje grapevine collection; Donja Pačetina, Škaričevo; GPS: 46.111417, 15.869363.

An additional set of 47 genotypes (*Vitis* SSR allele reference set), obtained from the Vassal collection, National Institute of Agricultural Research (INRA), France, was genotyped in order to ensure proper allele sizing and adjustment with external data sources (Table S2). The genotype 'Bombino bianco', a variety which in previous studies [14] was shown to be related to the Croatian gene pool, was obtained from The Research and Innovation Centre, Fondazione Edmund Mach, S. Michele all'Adige, Italy. DNA extraction was performed using the peqGOLD Plant DNA Mini Kit (PEQLAB Biotechnologie GmbH, Erlangen, Germany) according to the manufacturer's instructions.

## 2.2. SSR Genotyping

The selection of SSR loci was based on the level of their polymorphism, genome coverage, good amplification assessed through preliminary analyses, and compatibility with the previously published and available genetic profiles [66–75]. All selected loci were divided into four categories (cat.) according to their purpose (Table S3). Primers from cat.1 (which included nine SSRs proposed for cultivar identification [7]) were used for initial fingerprinting of all accessions and reference genotypes, cat.1 and cat.2 combined were used for analyses of genetic diversity and genetic structure, as well as parentage assignment, and primer pairs from cat.3 for analysis of genetic structure (just Croatian germplasm) and as additional support of reconstructed parentages. Cat.4 included primers needed for chlorotype determination. All SSR loci, except those from cat.1, were amplified using M13 (-21) tailed universal primers [76] in singleplex PCR for which the respective temperature regimes and PCR protocols were optimized. Additional data like primer sequences, associated annealing temperatures, and PCR and capillary electrophoresis arrangement of given loci are also presented in Table S3.

PCR amplifications were carried out in a Veriti™ Thermal Cycler (Applied Biosystems, Foster City, CA, USA). Multiplex PCR reactions were performed in a total volume of 10 µL containing 10 ng of DNA, 0.75 U Taq DNA polymerase (Sigma-Aldrich, St. Louis, MO, USA), 0.2 mM single dNTP, 0.3 µM forward and reverse primers, 2X PCR buffer (Sigma-Aldrich), 4.0 mM MgCl<sub>2</sub>, and 1 M Betaine solution, which served as a PCR enhancer. PCR temperature conditions were as follows: an initial denaturation of 2 min at 94 °C, followed by 35 successive cycles of denaturation (45 s at 94 °C), annealing (45 s at different T<sub>a</sub> ranging from 50 to 60 °C, depending upon primer used), and elongation (60 s at 72 °C), as well as a final elongation at 72 °C for 20 min. After fingerprinting, a non-redundant set was defined and further genotyped at additional SSR and cpSSR loci. A unique M13 PCR protocol was used. Singleplex PCR reactions were performed in a total volume of 9.5 µL containing 10 ng of DNA, 0.75 U Taq DNA polymerase (Sigma-Aldrich), 0.21 mM single dNTP, 0.26 µM universal M13 and reverse primers, 0.068 µM forward primer, 1X PCR buffer (Sigma-Aldrich), 4.0 mM MgCl<sub>2</sub>, and 1.05 M Betaine solution. After an initial denaturation of 3 min at 94 °C, 29 consecutive cycles of denaturation (30 s at 94 °C), followed by 45 s of annealing and 1 min of elongation at 72 °C, were performed and then followed by 10 cycles of 30 s denaturation at 94 °C, 45 s of annealing at 53 °C, and 1 min of elongation at 72 °C, with final elongation at 72 °C for 10 min. Amplified products were separated using an ABI3130 Genetic Analyzer (Applied Biosystems) with GeneScan- 500 LIZ™ size standard. Sizing of the fragments was performed using GeneMapper 4.0 software (Applied Biosystems).

## 2.3. SNP Genotyping

SNP genotyping was performed on a set of non-redundant genotypes that differed to a lesser extent from that used in the SSR analysis due to technical reasons (highlighted in Table S1). The SNP set consisted of 47 SNP loci previously used and proposed for grapevine fingerprinting [38], to which the SNP locus (i.e., cp4527) for chlorotype determination was added [77] to validate the reactions (instead of locus SNP555\_132). SNP loci names and allelic variants expected are listed in the supplement data (Table S4). A set of 47 SNP markers was run on a Biomark system applying the Fluidigm methodology. Primer pairs were used to amplify 47 DNA templates per run and one non-template control. The primers and fluorescently labeled oligo probes were designed and synthesized by Fluidigm (South San Francisco, CA, USA). Allele-specific PCR products were generated after pre-amplification

of the target regions by duplex quantitative real time PCR using Fluidigm's FR48.48 dynamic arrays. Genotypes were assigned based on the relative fluorescence intensities of the two alternative dyes (HEX and FAM) labeling the allele-specific probes targeting each SNP locus. Data analysis was performed in SNP Genotyping Analysis v4, and at least one negative control was included in the analysis for data normalization and auto-calibration ('no-template' control - NTC normalization method). Obtained SNP profiles were compared with available published SNP profiles [40].

#### 2.4. Microsatellite Database – Formation and Harmonization

Microsatellite profiles at nine common SSR loci were downloaded from the European *Vitis* database [8] for all available *Vitis vinifera* L. subsp. *sativa* accessions and adjusted using reference varieties and an Excel Macro Tool that converts actual allele lengths (bp) into encoded alleles, and vice versa, according to the principle previously described [6] and is freely available [78]. This resulted in the creation of a database of 2152 non-redundant genotypes, including 100 Croatian accessions previously genotyped through the GrapeGen06 project [7]. In addition, 1906 genetic profiles from publications containing shared loci [11,14,26,29–33,36,37,60,79–94] have been manually harmonized using common genotypes and added to the database for further analysis. Comparing all alleles across loci using the Excel-Microsatellite Toolkit [95], two conditions were identified: genotypes' uniqueness and/or synonymy/duplication. Accessions were considered as duplicates when they had an identical SSR profile, and assuming the existence of genotyping errors/spontaneous mutations, maximum difference of two alleles was allowed. Redundant accessions were later considered and discussed as known/unknown synonyms but excluded from further analyses. A non-redundant genotype set that included all genotypes of putative Croatian origin was used for allele and chlorotype frequency calculations, analysis of genetic structure, and parentage reconstructions. Apart from the determination of synonyms, published genetic profiles [14] were used to back up analysis of genetic structure and parentage reconstruction. As M13 primers extend the amplicon length, standardization of the data was performed matching the length of alleles recorded in Reference [14].

#### 2.5. Data Analysis

Non-redundant data was analyzed for Polymorphism Information Content [96], null allele probabilities [97], and probability of identity ( $P_{(ID)}$ ), under the two assumptions: that the individuals within the test set are non-related ( $P_{(ID)unrelated}$ ), and, in the other case ( $P_{(ID)sib}$ ), that the individuals are related [98] as well as deviation tests from Hardy–Weinberg equilibrium (HW), were all calculated in Cervus [99]. In GenAlEx 6.503 [100,101], the average number of alleles per locus ( $N_a$ ) was calculated, as well as the number of effective alleles per locus ( $N_e$ ), expected heterozygosity, and observed heterozygosity ( $H_O$ ), and fixation index (F). Allelic richness ( $N_{ar}$ ) was calculated in FSTAT [102] using the formula of Reference [103]. The allele frequency calculation for SSRs, as well as SNP, markers (expressed as minor allele frequency) was performed in Microsatellite Toolkit [95]. Each Croatian accession was assigned to one of four regions, according to their cultivation area: Istria and the Croatian Littoral ( $n = 23$ ), northwestern Croatia ( $n = 28$ ), Slavonia and the Danube region ( $n = 1$ ), and Dalmatia ( $n = 75$ ), all detailed in Table S1. The genetic dissimilarity coefficients between each pair of genotypes were calculated as  $-\ln(PSA)$ , where PSA is the proportion of shared alleles distance [104] in the Microsat program [105]. The dissimilarity matrix was further used for clustering of genotypes with UPGMA (Unweighted Pair Group Method with Arithmetic Mean) algorithm (for SSR data). UPGMA construction and dendrogram visualization was done in MEGA software [106].

#### 2.6. Genetic Structure

The model-based Bayesian method in STRUCTURE version 2.3.4 [107] was also used to imply the genetic structure and define the number of groups within the analyzed data set in order to assign Croatian germplasm to a certain ancestral group. The analysis was performed on a set of 771 accessions, which also included 664 individuals (defined as traditional) from Reference [14], together with Croatian

assumed set genotyped at 20 SSR loci. Those that have been defined as modern are the result of deliberate crosses by man and as such interfere with the assessment of the natural structure of the population. Refined Vassal genotypes were categorized into countries/regions of their presumed origin (based on the ampelographic literature), respectively, as done in Reference [15]: Balkans—BALK; Eastern Mediterranean and Caucasus—EMCA; Iberian Peninsula—IBER; Apennine Peninsula—ITAP; Maghreb—MAGH; Middle and Far East—MFEAS; Croatia; Russia and Ukraine—RUUK, Western and Central Europe—WCEUR, and a group of undetermined varieties—“ND”. A total of 10 repetitions were performed for each assumed  $k$  that ranged from 1 to 10. Each repetition consisted of a burn in a period of 150,000 steps followed by 150,000 Monte Carlo Markov Chain iterations under the assumption of the admixture model and the correlated allele frequencies. The choice of the most probable number of groups ( $k$ ) was carried out in the Structure Harvester [108] program by comparing the average probability estimators, calculated as  $\ln [\Pr (X | K)]$  for each  $k$  value, as well as by the  $\Delta K$  calculation based on the change in the probability log of data between successive values [109]. CLUMPP v.1.1.2 [110] was used to align data between independent repetitions for each  $k$  using the “greedy” algorithm with 10,000 repetitions. Structure Plot [111] was used for the graphical representation of the groups. Individuals with  $Q \geq 75\%$  of their genome were assigned to the group of question, while individuals with  $Q$  values  $< 75\%$  were considered to be of admixed origin. In order to confirm the genetic structure estimated by Structure analysis, only individuals with  $Q$  values  $> 0.75$ , were included in the molecular variance analysis (AMOVA). AMOVA was performed in GenAlEx 6.503 on a genetic distance matrix also based on the proportion of shared alleles calculated in Microsat [105].

### 2.7. Parentage Analysis

Parentage analysis was performed in Cervus [99] using a non-redundant set of genetic profiles comprised of the Croatian set ( $n = 127$ ) and Vassal genotypes ( $n = 1095$ ). Based on the results of allele frequency analysis, parent pair (gender unknown) simulation and analysis were performed and, subsequently, determination of “single parent – paternity” was done in Cervus in the same way, thus revealing first-degree relationship (with possible types of relationship—parent, offspring, full-sib, or half-sib). A million of offspring were simulated with a proportion of 0.01 sampled parents with the possibility of self-fertilization and the existence of relatives among potential parents. The critical LOD (logarithm of odds) value, with a higher confidence level (95%), obtained by simulation, was used as the criterion for parentage assignment, as well as allowing for no more than 5% mismatching (one out of 20 SSRs). In the case of successful parentage reconstruction, the maternal type was designated, when possible, using chlorotype data.

## 3. Results

### 3.1. Genetic Characterization of Croatian Grapevine Germplasm Repositories

The status of each studied accession was determined based on the genetic profile of 9 common SSRs compared versus the internal database and classified as follows: (I) unique genotype; (II) synonym/duplicate in Croatia; (III) synonym in neighboring countries; (IV) synonym in distant countries; (V) domesticated variety of known foreign origin; (VI) intruder—intruder found within the true to type vines; (VII) questionable varietal status—accession has a unique genotype, but, due to the lack of ampelographic or literature data, its varietal status is questionable; (VIII) mislabeled accession/incorrect collection—accession is planted under the wrong name in the collection and is not true to type (mostly, these accessions have never been fingerprinted); and IX) incorrect sampling of tissue for DNA analysis. Details for each accession can be found in Table S5.

Among 212 accessions, 19 duplicates were identified between collections (or within a single collection), so only one genetic profile (with positive ampelographic identification) was included in the non-redundant set. The same was done for 26 cases of synonymy among analyzed accessions. All identified synonyms and duplicates are highlighted in Table 1. In addition, one case of homonymy

identified as a variety of the Istrian collection called Plavina istarska (5IPT) shares the same name, but not the genetic profile, with the homonymous variety Plavina (4-C) grown in Dalmatia. After elimination of duplicates, synonyms, and intruders, the unique (non-redundant) set of SSR profiles consisted of 145 accessions. Within the collections, 145 different genotypes were detected, and, for 46 genotypes (32% of them), a synonym was found in one of the foreign database or scientific publications. Considering the status of each of the 145 unique accessions, all those that did not have synonyms, or do have synonyms abroad but their foreign origin is not strictly proven, were assumed as a “Croatian set” of varieties, totaling 127 of such accessions. Their SSR profiles for 36 loci are presented in Table S6.

**Table 1.** List of identified synonyms and duplicates at nine standard simple sequence repeat (SSR) loci.

No.	Accession 1	Accession 2	Accession 3	Accession 4	Matches
1	Plavina istarska (5IPT)	Surina (18IPT)			Maločrn, Plavina Maločrn [30]; Tükörszolo (P14#992), EuVDB *; Piccola nera [14]; Gnjjet [37]
2	Garganja (6IPT)	Garganja (2-H)			Lipolist (B&H) [36]
3	Vela pergola (24IPT)	Popovo oko (20IPT)	Malvazija crvena (21IPT)	Unknown 5 (7IPT)	Prošip Bijela (B&H) [37]
4	Plavinica (8IPT)	Plavina (4-C)	Lun (3-K)		Plavka (B&H) [36]
5	Duranija (9IPT)	Draganela [57]			Duranja [30]
6	Jarbola (10IPT)	Jarbola (11-A)			
7	Borgonja (11IPT)				Frankovka (CZE041-24V0100221), EuVDB; Tamjanika crna [37]
8	Muškat ruža porečki (13IPT)				Rosenmuskateller (DEU098-2000-021), EuVDB; Cipro [30]
9	Hrvatica (14 IPT)	Kamenina (KRK02, KRK14)			Corredera (DEU098-1993-001), EuVDB
10	Teran (15IPT)	Teran (2-B)			Terrano [14]; Lambrusco picol ross [79]; Refošk [93]
11	Malvazija istarska (16IPT)	Malvazija istarska (9-D)	Malvazija istarska (12-A)		
12	Muškat momjanski (17IPT)				Muscat a petit grains (REF09); Muscat d'Alsace rose (FRA139-559Mtp8), EuVDB Moscatel Galego Branco ** [40]
13	Cimaroša (22IPT)				Veltliner rouge (FRA139-284Mtp4), EuVDB
14	Dolcin (23IPT)	Dolčin (2-D)			Vitouska (ITA368#SANOSVALDO/VC_10), EuVDB, [83,90] Vitouska [30]
15	Unknown 7 (25IPT)				Refosco peduncolo rosso (ITA362-1662), EuVDB
16	Pagadebit istarski (26IPT)	Kuč (3-C)	Maraština omiška (26-SK)		Mostosa = Empibotte bianco, [14]; Uva vacca [79]
17	Bašćan (KRK01)	Stara brajda (3-L)	Susac (10-E)	Crna Brajda (8-SK)	Bascina bijela (DEU098-1990-130), EuVDB
18	Plavčina (KRK04)	Plavčina (10-B)			
19	Ošljevin (KRK03)				Zelenika [29]; Rožica [30]
20	Troiščina (KRK08)	Trojščina (11-B)			
21	Žumić (KRK)	Žumić (5-G)			

Table 1. Cont.

No.	Accession 1	Accession 2	Accession 3	Accession 4	Matches
22	Brajdica (KRK)	Brajdica (2-J)			
23	Debejan (KRK)	Debejan (10-C)			
24	Gnjatonja(12-SK)	Dugovrst (5-B/2013)			
25	Svrdlovina (17-SK)	Svrdlovina (27-A)	Gustopupica bibinjska (4-i_novi)		
26	Okatica crna(19-SK)	Okatica vrgorska (VRG08)			
27	Palaruša bijela (22-SK)	Bumba bijela (8-K)	Medna bijela (internal acc.)		Žložder [37]
28	Okatica bijela (23-SK)	Palaruša hvarska (4-D)	Palaruša viška (5-D)		
29	Silbijanac (25-SK)	Silbijanac (3-G)			
30	Divljak (ZAG01)	Ranfol (ZAG05)	Ranfol (ZAG27)	Ranfol (9-E)	
31	Dišeća ranina (ZAG08)	Dišeća ranina (9-K)			
32	Kavčina (ZAG10)				Kövidinka (P14#520), EuVDB
33	Belina Mala (ZAG11)				Biancghera [30]
34	Belina starohrvatska (ZAG16)				Gouais blanc (FRA139-211Mtp1), EuVDB; Heunisch weiss [88]; Branco Valente **[40]
35	Sokol (ZAG21)	Sokol (12-B)			Luglienga bianca (ITA360-501), EuVDB Caramela ** [40]
36	Volovina (ZAG26)				Moscato violetto [14]; Muscat rouge de Madère (ITA025-CO018), EuVDB
37	Modra kosovina (ZAG30)				Zimmettraube weiss (DEU098-1980-381), EuVDB
38	Moslavac (ZAG32)	Moslavac (9-L)			Furmint [14] Šipon [37]
39	Kozjak (ZAG42)				Coarna alba (DEU098-1991-188), EuVDB; Uva picciona [79]
40	Smudna belina (ZAG48)	Svjetljak (12-G)			
41	Beli Debejan (10-D)	Gegjć (3-F)			
42	Verdić (11-F)				Teran bijeli [57]; Glera, Prosecco [29]; Beli Teran [30]; Prosecco tondo [89]
43	Petovka (3-E)	Cetinka (6-F)	Plavina bijela ninska (12-C)		
44	Malvasia dubrovačka (2-A)				Malvasia delle Lipari [94]

Table 1. Cont.

No.	Accession 1	Accession 2	Accession 3	Accession 4	Matches
45	Drnekuša vela (3-D)	Glavanjuša (5-A)			
46	Cipar (3-J)				Grec rouge [14]; Rabigato Franco ** [40]
47	Maraština (4-A)				Malvasia del Chianti [14]; Pavlos [91]
48	Babić (4-B)	Babica plosnata (8-L)			
49	Lasina (4-E)	Vlaški crljenak Brač (5-J)			
50	Ninska crvena (4-K)	Šemperinka (8-G)	Vranac (7-G)	Vranac (15-F)	Vranac [14]
51	Pavčić (5-E)	Oskorušica (8-E)			
52	Bak (5-F)	Siložder (6-L)			
53	Bilan (5-H)				Racuk [29]
54	Beretinjok (5-I)	Topol (3-H_novi)			Bianco d'Alessano [11]
55	Muškat ruža omiški (5-K)				Muscat de Hambourg [14]
56	Muškat bijeli omiški (5-L)				Muscat d'Alexandrie [14]
57	Zlatarica vrgorska (6-C)				Kadarun IV, Surac IV [33]; Kadarun (B&H) [36]; Francavidda [11]
58	Pošip crni (6-D)	Šljiva (26-A)	Razaklija (28-B)		
59	Marinkovića grozje (6-I)				Dattier de Beyrouth = Afuz Ali [14]; Radovača [37]
60	Plavac mali (7-A)	Plavac mali sivi (7-C)			
61	Mijajuša (7-K)				Assouad karech [14]; Xerichi kokkino, EuVDB
62	Tribidrag (8-A)				Primitivo = Zinfandel [14] Kratošija [92]
63	Lelekuša (9-J)	Bratkovina crvena (30-A)			
64	Frmentum (6-H)				Santa Teresa [11]
65	Bratkovina bijela (29-A)	Pošipica (33-B/2013)	Stradunska [36]		Maruggio (Maresco)= Uva del Monaco [11]; Popetre [30]
66	Bljuzgavac (27-B_novi)				Blank blauer [14];
67	Lipovina (24-A/2013)				Harslevelu [14]
68	VRG10				Lisičina [60]
69	Kadarka				Olasz Kadarka [14]
70	Portugizac	Portugieser blau (REF12)			Portugais bleu [14]

\* European *Vitis* Database [8], accessions numbers also presented when available. \*\* matches obtained through single nucleotide polymorphism (SNP) comparison.

### 3.2. Descriptive Molecular Statistics for SSR And SNP Data

SRR alleles and their frequencies for non-redundant set are given in Table S7. The genotypic data were fitted into allelic classes based mostly on a difference of three or two base pairs, but there were some alleles with a consistent difference of just one base pair. In addition, novel alleles not previously reported were detected, e.g., a new second allele at locus VMC4f3 (246 bp) in 'Kraljevina'. However, no new alleles were found for any of the 9 common fingerprinting SSR loci. The genetic diversity parameters for a total of 36 analyzed SSR loci were calculated for the entire set of 127 assumed Croatian varieties (Table 2).

Percentage of missing data was below 5% for most loci with exceptions of VViv37, Vchr4a, Vchr11b, and Vchr2b loci, which had somewhat lower average amplification rate. On average, nine alleles per locus were amplified, with values ranging from two (Vchr17a) to 15 for Vchr8b and VMC4f3, which also had the highest allelic richness (13.59). The effective number of alleles ranged from 1.47 for locus VVIn73 to 8.20 for locus VVMD28, with a mean of 3.9. The same loci were at the end of range for gene diversity (0.32 and 0.88) and observed heterozygosity (0.32 and 0.94). Mean  $H_O$  (0.71) and  $H_e$  (0.70) were almost the same. For Vchr8b and Vchr14b loci, the highest fixation index, deviation from the Hardy–Weinberg equilibrium, and the highest incidence of null alleles was estimated (so they were therefore excluded from further parentage analyses). All other alleles used in the parentage analysis had a satisfactorily low chance of null allele's occurrence. In correlation with previous parameters, informativeness of the SSR loci expressed as polymorphic information content (PIC) ranged from 0.29 for VVIn73 to 0.78 for VVMD28, with the average PIC value of 0.66. Together with VVMD28, four more loci (ssrVrZAG79, ssrVrZAG62, VVMD5, and VVMD32) from the common set of nine SSRs were amongst the eight most informative ones. The single-locus  $P_{(ID) \text{ unrelated}}$  values ranged from 0.027 (VVMD28) to 0.496 (VVIn73), whereas the  $P_{(ID) \text{ sib}}$  values ranged from 0.318 (VVMD28) to 0.715 (VVIn73). The  $P_{(ID) \text{ unrelated}}$  across all 36 loci was  $1.19 \times 10^{-34}$ , whereas  $P_{(ID) \text{ sib}}$  across loci was  $5.29 \times 10^{-14}$ .

SNP analysis was performed for the 124 genotypes (defined in Table S1) across 47 SNP loci (data provided in Table S8). Two loci were excluded from further analysis, one due to being non-polymorphic (SNP579\_187) and the other one due to its poor amplification (SNP1445\_218). Genetic diversity parameters were calculated for all remaining SNP loci (Table 3). The amplification success rate was very high, almost 100% except at the SNP259\_199 locus, where only 0.81% of the data was missing. The lowest number of effective alleles ( $N_e = 1.03$ ) was recorded for the SNP425\_205 locus and the highest (2.0) for the five SNP loci (SNP197\_82, SNP879\_308, SNP945\_88, SNP325\_65, SNP1215\_138), with mean of 1.71. The highest observed heterozygosity ( $H_O = 0.7$ ) was detected for SNP581\_114 and the lowest ( $H_O = 0.03$ ) for SNP425\_205. Significant deviation from the Hardy–Weinberg equilibrium was observed for SNP581\_114 and VVI\_10329, as well as a fixation index that was significantly greater or less than 0. At 9% of loci, the frequency of the minor allele was <0.1. The highest PIC (0.38) was observed for SNP879\_308, SNP945\_88, SNP197\_82, and SNP325\_65, while the lowest (0.03) was for the SNP425\_205. The single-locus  $P_{(ID) \text{ unrelated}}$  values ranged from 0.38 to 0.94 (SNP425\_205), whereas the  $P_{(ID) \text{ sib}}$  values ranged from 0.59 to 0.97, for the same loci, respectively. The  $P_{(ID) \text{ unrelated}}$  across all 45 loci was  $1.19 \times 10^{-34}$ , whereas  $P_{(ID) \text{ sib}}$  across loci was  $5.29 \times 10^{-14}$ . High frequency of null alleles ( $f > 0.05$ ) was detected in 40% of the loci.

**Table 2.** Genetic diversity parameters assessed for 36 SSR loci.

Locus	Allele Range (bp)	N	MD	N <sub>a</sub>	N <sub>e</sub>	N <sub>ar</sub>	H <sub>e</sub>	H <sub>o</sub>	F	HW	F null	PIC	P(ID)unrelated	P(ID)sib
VVS2	123–153	127	0	12	4.65	10.51	0.79	0.78	0.007	NS	0.0063	0.758	0.073	0.376
VVMD7	231–261	127	0	10	3.66	8.70	0.73	0.80	−0.095	NS	−0.0436	0.685	0.116	0.416
VVMD27	171–190	127	0	8	3.49	6.94	0.71	0.80	−0.115	NS	−0.0583	0.67	0.126	0.425
ssrVrZAG62	179–203	127	0	10	5.34	8.97	0.81	0.89	−0.095	NS	−0.0484	0.789	0.058	0.358
ssrVrZAG79	234–258	127	0	11	5.76	10.45	0.83	0.86	−0.039	NS	−0.0205	0.804	0.053	0.35
VVMD5	218–242	127	0	10	5.08	9.36	0.80	0.84	−0.049	NS	−0.0258	0.78	0.062	0.364
VVMD25	237–265	127	0	10	3.77	8.07	0.74	0.74	−0.007	NS	−0.0024	0.691	0.115	0.411
VVMD28	216–276	127	0	12	8.20	11.09	0.88	0.94	−0.067	ND	−0.0329	0.866	0.027	0.318
VVMD32	240–274	127	0	9	5.07	8.47	0.80	0.91	−0.128	NS	−0.0627	0.778	0.064	0.365
VViq52	101–111	126	0.79	5	2.78	4.71	0.64	0.71	−0.103	NS	−0.0544	0.57	0.2	0.48
VVip31	192–216	127	0	12	7.18	11.18	0.86	0.95	−0.098	ND	−0.0476	0.845	0.035	0.328
VVip60	321–348	124	2.36	12	4.74	10.61	0.79	0.82	−0.032	NS	−0.0184	0.76	0.074	0.374
VMC1b11	185–215	125	1.57	10	4.88	9.62	0.80	0.85	−0.066	NS	−0.0362	0.774	0.063	0.368
VMC4f3	185–246	127	0	15	4.27	12.38	0.77	0.76	0.013	NS	0.0066	0.741	0.079	0.387
VVih54	166–205	126	0.79	12	4.11	10.15	0.76	0.81	−0.07	NS	−0.0341	0.732	0.084	0.393
VViv67	354–406	126	0.79	12	5.25	10.28	0.81	0.83	−0.03	NS	−0.0152	0.785	0.061	0.361
VVib01	307–325	127	0	5	2.79	5.00	0.64	0.67	−0.044	NS	−0.0205	0.58	0.19	0.477
VVMD24	224–235	127	0	6	2.55	5.92	0.61	0.61	0.003	NS	0.0046	0.559	0.202	0.496
VVMD21	244–281	126	0.79	7	2.60	5.92	0.62	0.71	−0.148	NS	−0.0712	0.57	0.193	0.491
VVin16	167–177	127	0	5	2.45	4.98	0.59	0.58	0.03	NS	0.0116	0.523	0.235	0.513
VVin73	272–285	127	0	7	1.47	5.14	0.32	0.32	0.011	ND	0.0095	0.287	0.496	0.715
VViv37	153–183	106	16.54	13	3.93	11.33	0.75	0.79	−0.063	NS	−0.0422	0.721	0.09	0.399
Vchr8b	115–169	125	1.57	15	7.41	13.59	0.87	0.69	0.205	**	0.1177	0.852	0.031	0.325
Vchr10b	145–154	126	0.79	3	2.45	3.00	0.59	0.66	−0.112	NS	−0.0525	0.505	0.254	0.517
Vchr14b	188–239	125	1.57	13	4.09	11.63	0.76	0.32	0.576	***	0.4073	0.728	0.087	0.394
Vchr4a	200–224	115	9.45	6	2.67	5.08	0.63	0.64	−0.03	NS	−0.0106	0.552	0.214	0.491
Vchr9a	108–142	126	0.79	8	4.26	7.36	0.77	0.83	−0.079	NS	−0.0355	0.73	0.09	0.39
Vchr16a	118–186	127	0	8	1.98	7.32	0.50	0.49	0.014	NS	0.012	0.457	0.293	0.576
Vchr19a	143–170	118	7.09	8	3.13	7.27	0.68	0.71	−0.047	NS	−0.0305	0.639	0.144	0.446
Vchr17a	197–205	119	6.3	2	1.83	2.00	0.45	0.38	0.168	NS	0.0914	0.351	0.401	0.623
Vchr18a	170–210	122	3.94	10	4.72	8.28	0.79	0.75	0.043	NS	0.0211	0.761	0.072	0.374
Vchr1b	112–128	122	3.94	4	2.85	3.97	0.65	0.61	0.065	NS	0.0277	0.578	0.194	0.474
Vchr7b	189–205	121	4.72	4	3.74	4.00	0.73	0.80	−0.094	NS	−0.0447	0.683	0.122	0.414
Vchr11b	169–181	109	14.17	6	4.05	5.38	0.75	0.78	−0.036	NS	−0.0188	0.708	0.106	0.4
Vchr12a	141–162	121	4.72	6	2.98	5.40	0.67	0.72	−0.082	NS	−0.0417	0.605	0.172	0.461
Vchr2b	130–139	62	51.18	4	1.64	4.00	0.39	0.37	0.047	ND	0.0171	0.348	0.414	0.659
Mean		122	3.718	9	3.9	7.72	0.70	0.71	−0.015		−0.0038	0.66		
Over loci													$1.19 \times 10^{-34}$	$5.29 \times 10^{-14}$

\*\* significant at the 1 % level, \*\*\* significant at the 0.1 % level, NS = not significant, ND = not done. N = number of observed accessions; MD = % missing data; N<sub>a</sub> = number of alleles, N<sub>e</sub> = effective number of alleles, Ar = allelic richness, H<sub>e</sub> = expected heterozygosity, H<sub>o</sub> = observed heterozygosity, F = fixation index, HW = exact test of departure from Hardy–Weinberg equilibrium, F null = estimated frequency of null alleles, PIC = polymorphic information content, P(ID) probability of identity = unrelated and sib .

**Table 3.** Genetic diversity parameters assessed for 45 SNP loci in the subset of 124 unique accessions of the Croatian grape germplasm.

SNP Locus *	N	MD	Na	Ne	Nar	Ho	He	F	HW	MAF	PIC	P(ID) unrelated	P(ID) sib	F null	
SNP1003_336	124	0	2	1.78	2.00	0.42	0.44	0.04	NS	A:	0.323	0.342	0.412	0.635	0.0207
SNP1015_67	124	0	2	1.46	2.00	0.35	0.32	-0.094	ND	C:	0.476	0.267	0.517	0.721	-0.0447
SNP1027_69	124	0	2	1.96	2.00	0.51	0.49	-0.036	NS	G:	0.266	0.37	0.38	0.6	-0.0175
SNP1035_226	124	0	2	1.39	2.00	0.31	0.28	-0.089	ND	T:	0.48	0.242	0.556	0.748	-0.0427
SNP1079_58	124	0	2	1.96	2.00	0.60	0.49	-0.219	NS	A:	0.246	0.37	0.38	0.6	-0.0988
SNP1119_176	124	0	2	1.98	2.00	0.48	0.49	0.02	NS	C:	0.375	0.372	0.378	0.598	0.01
SNP1127_70	124	0	2	1.64	2.00	0.44	0.39	-0.115	NS	A:	0.198	0.314	0.448	0.667	-0.0543
SNP1157_64	124	0	2	1.08	2.00	0.07	0.07	-0.038	ND	G:	0.266	0.068	0.867	0.932	-0.0097
SNP1215_138	124	0	2	2.00	2.00	0.53	0.50	-0.067	NS	C:	0.117	0.374	0.376	0.594	-0.0324
SNP1229_219	124	0	2	1.64	2.00	0.47	0.39	-0.197	NS	C:	0.016	0.314	0.448	0.667	-0.0899
SNP1323_155	124	0	2	1.69	2.00	0.48	0.41	-0.164	NS	G:	0.44	0.325	0.433	0.654	-0.0759
SNP1347_100	124	0	2	1.87	2.00	0.56	0.47	-0.198	NS	A:	0.351	0.357	0.395	0.616	-0.09
SNP1349_174	124	0	2	1.99	2.00	0.56	0.50	-0.116	NS	T:	0.431	0.374	0.376	0.595	-0.055
SNP1399_81	124	0	2	1.06	2.00	0.06	0.06	-0.029	ND	A:	0.286	0.053	0.895	0.946	-0.0062
SNP1411_565	124	0	2	1.32	2.00	0.27	0.24	-0.098	ND	C:	0.351	0.213	0.603	0.78	-0.0466
SNP1453_40	124	0	2	1.64	2.00	0.44	0.39	-0.115	NS	T:	0.379	0.314	0.448	0.667	-0.0543
SNP1471_179	124	0	2	1.26	2.00	0.23	0.21	-0.132	ND	C:	0.395	0.185	0.651	0.809	-0.0569
SNP1513_153	124	0	2	1.84	2.00	0.56	0.46	-0.222	NS	T:	0.302	0.352	0.4	0.622	-0.0998
SNP191_100	124	0	2	1.07	2.00	0.07	0.06	-0.033	ND	C:	0.169	0.06	0.881	0.939	-0.0079
SNP197_82	124	0	2	2.00	2.00	0.51	0.50	-0.016	NS	A:	0.367	0.375	0.375	0.594	-0.008
SNP227_191	124	0	2	1.85	2.00	0.48	0.46	-0.034	NS	T:	0.032	0.354	0.397	0.619	-0.0167
SNP259_199	123	0.81	2	1.76	2.00	0.46	0.43	-0.051	NS	G:	0.488	0.339	0.415	0.637	-0.025
SNP269_308	124	0	2	1.92	2.00	0.45	0.48	0.055	NS	C:	0.302	0.364	0.387	0.608	0.0284
SNP325_65	124	0	2	2.00	2.00	0.40	0.50	0.208	NS	A:	0.427	0.375	0.375	0.594	0.1163
SNP425_205	124	0	2	1.03	2.00	0.03	0.03	-0.016	ND	G:	0.472	0.031	0.938	0.969	-0.0023
SNP447_244	124	0	2	1.89	2.00	0.45	0.47	0.041	NS	A:	0.496	0.36	0.391	0.612	0.0207
SNP581_114	124	0	2	1.97	2.00	0.70	0.49	-0.424	***	A:	0.44	0.371	0.379	0.598	-0.1749
SNP593_149	124	0	2	1.84	2.00	0.51	0.46	-0.115	NS	T:	0.379	0.352	0.4	0.622	-0.0546
SNP613_315	124	0	2	1.34	2.00	0.27	0.25	-0.048	ND	A:	0.306	0.222	0.589	0.77	-0.0236
SNP697_296	124	0	2	1.30	2.00	0.25	0.23	-0.084	ND	A:	0.444	0.204	0.618	0.789	-0.0401
SNP819_210	124	0	2	1.59	2.00	0.44	0.37	-0.196	NS	G:	0.028	0.302	0.465	0.681	-0.0892
SNP829_281	124	0	2	1.97	2.00	0.57	0.49	-0.162	NS	C:	0.359	0.371	0.379	0.598	-0.075
SNP873_244	124	0	2	1.92	2.00	0.50	0.48	-0.046	NS	C:	0.351	0.364	0.387	0.608	-0.0225
SNP879_308	124	0	2	2.00	2.00	0.57	0.50	-0.146	NS	A:	0.488	0.375	0.375	0.594	-0.068
SNP895_382	124	0	2	1.89	2.00	0.52	0.47	-0.096	NS	A:	0.113	0.36	0.391	0.612	-0.046
SNP945_88	124	0	2	2.00	2.00	0.52	0.50	-0.049	NS	T:	0.266	0.375	0.375	0.594	-0.0239
SNP947_288	124	0	2	1.99	2.00	0.63	0.50	-0.263	NS	A:	0.141	0.374	0.376	0.595	-0.1163
VVL_10113	124	0	2	1.89	2.00	0.40	0.47	0.143	NS	A:	0.317	0.36	0.391	0.612	0.0772
VVL_10329	124	0	2	1.88	2.00	0.27	0.47	0.432	***	T:	0.149	0.359	0.392	0.614	0.2757
VVL_10353	124	0	2	1.84	2.00	0.49	0.46	-0.08	NS	A:	0.468	0.352	0.4	0.622	-0.0385
VVL_10992	124	0	2	1.73	2.00	0.41	0.42	0.025	NS	G:	0.387	0.333	0.423	0.645	0.0128
VVL_12882	124	0	2	1.73	2.00	0.35	0.42	0.178	NS	A:	0.036	0.333	0.423	0.645	0.0978
VVL_1617	124	0	2	1.74	2.00	0.52	0.43	-0.214	NS	G:	0.395	0.335	0.421	0.643	-0.0967
VVL_9227	124	0	2	1.25	2.00	0.19	0.20	0.034	ND	G:	0.133	0.18	0.66	0.815	0.0172
VVL_9920	124	0	2	1.90	2.00	0.48	0.48	-0.02	NS	G:	0.379	0.362	0.389	0.61	-0.0098

\* SNP579\_187e is not shown due to being non polymorphic. \*\*\* significant at the 0.1 % level; N = number of observed accessions; MD = % missing data; Na = number of alleles, Ne = effective number of alleles, Nar = allelic richness, Ho = observed heterozygosity, He = expected heterozygosity, F = fixation index, HW = exact test of departure from Hardy–Weinberg equilibrium, MAF = minor allele frequency, PIC = polymorphic information content, P(ID) probability of identity = unrelated and sib; F null = estimated frequency of null alleles.

### 3.3. Comparison of SNP and SSR for Variety Discrimination

A set of nine SSRs, as well as set of 45 SNPs, both proposed for cultivar identification, were able to discriminate all investigated genotypes (data not shown). To compare the performance of nine SSRs and 45 SNP loci in variety distinction, a comparative summary of the genetic variation and polymorphism statistics (Table 4) is given. Amplification success was similar for both types of markers. The total number of alleles per locus ( $A = 92$  for SSRs vs.  $A = 90$  for SNPs) was almost identical. Expected and observed heterozygosity values were twice higher for 9 SSRs (e.g.,  $H_O = 0.84$  for SSRs vs.  $H_O = 0.416$  for SNPs). Average PIC values were more than twice higher for SSR (0.76) compared to SNPs (0.31). However, cumulative  $P_{(ID)sib}$  values were more favorable in both cases for SNP markers (cum  $P_{(ID)unrelated} 4.57 \times 10^{-16}$ , cum  $P_{(ID)sib} 1.00 \times 10^{-08}$ ) versus SSR markers (cum  $P_{(ID)unrelated} 4.07 \times 10^{-11}$ , cum  $P_{(ID)sib} 1.44 \times 10^{-04}$ ). To obtain equal values of cumulative  $P_{(ID)}$  values for SNP markers, the 18–19 most informative SSR markers should be used. Having in mind these differences, nine common SSR markers were 2.5 times more informative than 45 proposed SNP markers.

**Table 4.** Comparison of genetic diversity statistics assessed for SSR and SNP sets recommended for routine fingerprinting.

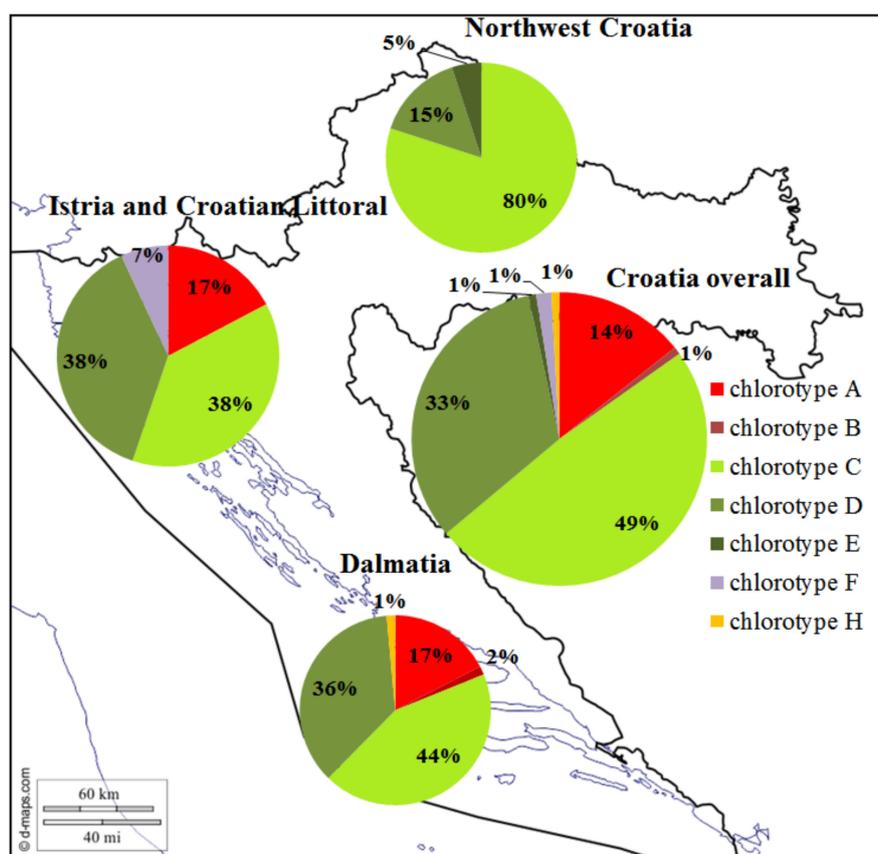
Marker Type	N loci	N	MD	A	N <sub>a</sub>	H <sub>o</sub>	H <sub>e</sub>	PIC	cum $P_{(ID)unrelated}$	cum $P_{(ID)sib}$
SSR	9 *	127	0	92	10.22	0.84	0.79	0.76	$4.07 \times 10^{-11}$	$1.44 \times 10^{-04}$
SNP	45	124	0.02%	90	2	0.416	0.39	0.31	$4.57 \times 10^{-16}$	$1.00 \times 10^{-08}$

\* set of nine recommended SSR loci [7]. N = number of observed accessions, MD = missing data, A = total number of amplified alleles, N<sub>a</sub> = average number of alleles per locus, H<sub>o</sub> = observed heterozygosity, H<sub>e</sub> = expected heterozygosity, PIC = polymorphic information content, P(ID) probability of identity = unrelated and sib.

### 3.4. Chlorotype Distribution in Croatian Germplasm

Fragment lengths for the four analyzed cpSSRs and associated chlorotypes are listed in Table S9. Chloroplast DNA analysis on a set of 127 putative Croatian varieties resulted in the reconstruction of the chloroplast haplotype for 93% genotypes (Table S6).

A total of seven different chlorotypes (A, B, C, D, E, F, and H) were detected. Graphical representation of chlorotype shares across regions within Croatia (Figure 2) shows differences in their distribution. Generally, in Croatia, the most prevalent chlorotype is C, which occurs in about half of the varieties, one-third of varieties have chlorotype D, and the third most common has chlorotype A, with differences among regions in their geographical distribution. In northwestern Croatia, chlorotype C dominates and is followed by chlorotype D. Interestingly, chlorotype A, in which occurrence in coastal regions (comprising regions Istria and Croatian Littoral, as well as Dalmatia) accounts for 17%, was not recorded in northwestern Croatia at all. For the region of Istria and the Croatian Littoral, four different chlorotypes (A, C, D, and F) have been observed, with chlorotypes C and D being dominant. A total of five different chlorotypes (A, B, C, D, and H) have been recorded in Dalmatia, making it the region of highest chlorotype diversity. The dominant chlorotype in this region was chlorotype C.



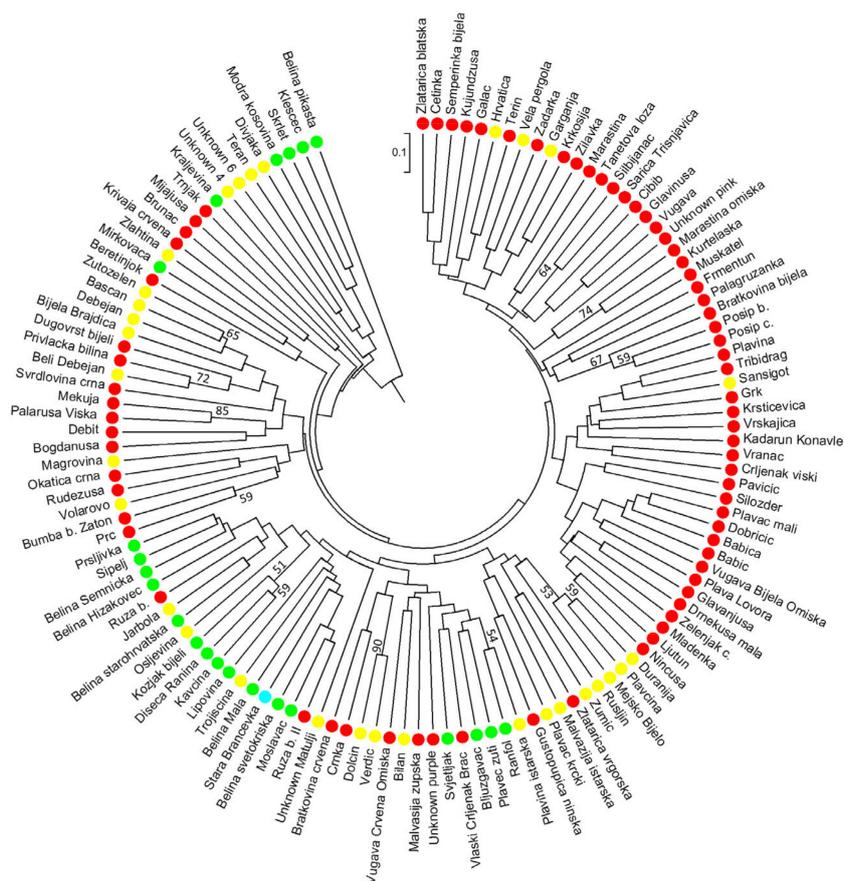
**Figure 2.** Graphical representation of chlorotype distribution across regions of Croatia.

### 3.5. Cluster Analysis and Genetic Structure

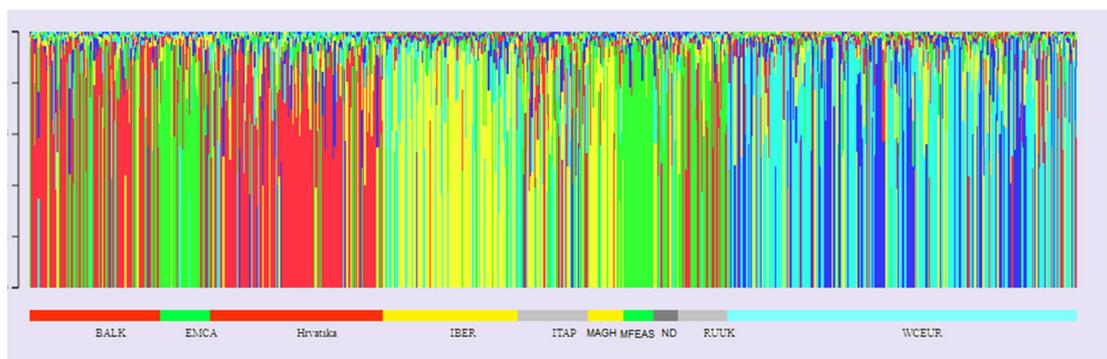
The grouping of 127 presumably Croatian varieties based on genotyping with 36 SSR loci is shown in Figure 3. Clustering was mostly in correlation with regions of their cultivation. Dalmatian varieties are grouped separately, and, within existing subgroups, it is possible to notice grouping due to kinship, e.g., subgroups with 'Plavac mali' or 'Tribidrag'. The same was true for the varieties of northwestern Croatia, for example, a subgroup around 'Belina starohrvatska'. Varieties from Istria and the Croatian Littoral also showed certain level of homogeneity within the sub-clusters, although their connections with both Dalmatian and continental varieties can be noticed.

Based on the Structure Harvester results, a decision was made on the best  $k$  ( $k = 5$ ) as the number of original populations (Figure S1). Membership coefficients ( $Q$  values) in each of the five proposed populations are given in Table S10 for all putative Croatian varieties.

The same is visualized in Figure 4, after the varieties have been sorted according to the sampling site or the assumed origin. With sampling location as the criterion, inferred original populations can be renamed as follows: population 1 (represented in red, marked Balkan) covers the largest number of varieties from Hungary, Croatia, Serbia, and Romania and could be broadly defined as the Balkans. Population 2 (represented in yellow and designated Iber) comprises the largest number of varieties from Spain, Portugal, Algeria, and Morocco and was defined as the Iberian Peninsula and Maghreb countries.



**Figure 3.** Circular UPGMA (Unweighted Pair Group Method with Arithmetic Mean) dendrogram for 127 presumably Croatian varieties based on 36 SSR loci. Dots represent varieties’ cultivation region: red = Dalmatia, yellow = Istria and the Croatian Littoral; green = northwestern Croatia; blue = Slavonia and Danube region.



**Figure 4.** Graphical representation of the proposed structure for  $k = 5$  according to Structure analysis. Each variety is represented by a vertical column colored according to the Q coefficient and the corresponding original population. The following abbreviations represent the sampling regions: BALK = Balkans; EMCA = Eastern Mediterranean and Caucasus; Hrvatska = Croatia; IBER = Iberian Peninsula; ITAP = Apennine Peninsula, MAGH = Maghreb region; MFEAS = Middle East and Far East; ND = a group of undetermined varieties; RUUK = Russia and Ukraine; WCEUR = Western and Central Europe. Balkan ancestral population is marked red, Iber ancestral population is marked yellow, East ancestral population is marked green, Biscay ancestral population is marked light blue, and Gallica ancestral population is marked dark blue.

Population 3 (represented in green and labeled East) encompasses the countries of eastern Mediterranean and the Black Sea, as well as Middle and Middle East countries, such as Greece, Turkey, Bulgaria, Syria, Cyprus, and Uzbekistan. Population 4 (visualized in light blue color coded as Biscay) includes primarily varieties of southwestern France, i.e., varieties from Bordeaux, and a smaller number of varieties of northern Portugal, Galicia, and northern Spain, as well as those cultivated along the Bay of Biscay (offshore Atlantic). Population 5 (visualized in darker blue and designated Gallica) was formed primarily around two varieties, 'Pinot' and 'Gouais blanc', and their numerous offspring scattered throughout northeastern France and Germany. Out of the 127 putative Croatian varieties, 57 had  $Q \geq 0.75$  within population 1, designated Balkan. An additional 34 varieties had more than 50% of their genome originating in the Balkan set but are still considered to be admixed. A total of eight varieties had at least 75% of their genome originating in another population: 'Bogdanuša', 'Maraština omiška', and 'Frmentum' originate from Iber population; 'Belina starohrvatska' (syn. 'Gouais blanc', 'Heunisch weiss') belongs to the Gallica group, while 'Cibib', 'Krivaja crvena', 'Mijajuša', and 'Trnjak' were assigned to the East group.

Varieties with  $Q$  values  $\geq 0.75$  (total of 455 varieties), distributed into five populations (Pop1 (Balkan) = 125, pop2 (Iber) = 87, pop3 (East) = 91, pop4 (Gallica) = 111, and pop5 (Biscay) = 41), were subjected to AMOVA. Results of AMOVA for the five subpopulations detected by Structure analysis (Table S11) confirm the proposed Structure clustering since  $\Phi_{PT}$  was highly significant indicating the existence of a general structure and confirming that most genetic diversity results from differences between populations.  $\Phi$  values for each population pair (Table S11) showed high significance for each pair of proposed populations and indicate that Balkan group is the most similar to East, while the most divergent is from the Gallica population.

### 3.6. Parentage Analysis

In Table 5, 24 proposed full parentages of Croatian varieties are presented, with 14 out of them having not been previously reported. The critical LOD values were determined by simulation of parentage (LOD = 14.5 for  $P = 0.05$ ).

When possible, genotyping was further extended so number of matches (based on visual inspection) of additional 14 three- and tetra-SSR loci, as well as 45 SNP loci, are indicated. Chlorotypes, if known, are highlighted so that maternal genotypes are indicated where possible. Although cv. 'Sokol' was not part of 127 Croatian set, his already well-known parentage is shown, as confirmation of his foreign origin. All proposed parentages were significant for  $P = 0.05$ . Offspring detected resulted from cross pollination of two different varieties, and self-pollination was not detected. For eight parentages ('Belina starohrvatska' included), both parents and offspring are part of Croatian germplasm, while, for the 17 proposed trios, at least one parent is of a foreign origin. The two most common varieties that emerge as a parent in six proposed trios are the Italian variety 'Bombino bianco' (syn. Slovenian Trevolina [30]) and 'Belina starohrvatska' (syn. French 'Gouais blanc', German 'Heunisch weiss', Table 1). 'Bombino bianco' can be characterized as a mother in at least half of the proposed trios. For two trios, 'Plavac mali' is a second proposed parent, with 'Bombino bianco' as a first, but since they share the same chlorotype, it is not possible to distinguish the maternal parent. The next most common parent in six crosses was 'Belina starohrvatska', which was in half of the cases a pollen recipient from donor 'Bljuzgavac' (syn. 'Blank blauer' [14]), while its crossing with the Romanian 'Alba imputotato' resulted in two more offspring. The aforementioned 'Bljuzgavac' is part of four parentages, and parents—'Bratkovina crvena' and Hungarian 'Gyöngy feher'—have been proposed for the same. The most represented Dalmatian variety was 'Plavac mali' (for which previously proposed parentage [64] was not confirmed [112]), and it was recorded as one of the parents in four proposed trios. For three varieties ('Kozjak bijeli', 'Muškattel', and 'Volovina'), both proposed parents were of foreign origin.

**Table 5.** Proposed full parentages of Croatian germplasm based on 20 SSR loci.

Offspring	First Candidate Parent	Second Candidate Parent	SSR Total/ Mismatch	LOD	add. SSR Total/ Mismatch	SNP	Previously Reported
Sokol (n.d.) *	Bermestia bianca	Madeleine Salomon = Agostenga blanc	20/0	46.01			[14]
Dolcin (D)	Verdić (D)	Maraština (D)	20/0	40.19	14/0	45/45	[19,90]
Ljutun (D)	Plavac mali (D)	Bombino bianco (D)	20/0	40.04	14/0		
Krstičevica (n.d.)	Plavina (C)	Bombino bianco (D)	20/0	39.62	13/0		
Plavina (C)	Tribidrag (D)	Verdeca = Lagorthi	20/0	37.72			[12,14]
Kurtelaška (D)	Kuč (A)	♀ Bombino bianco (D)	20/0	37.56	14/0		[14]
Belina svetokriška (C)	Alba imputotato	Belina starohrvatska (C)	19/0	36.38			
Debit (D)	Lasina (C)	♀ Bombino bianco (D)	20/0	36.34	8/0		[14]
Kozjak b. (C)	Bulanyi	Csomorika	18/0	36.19			
Ninčuša (D)	Plavac mali (D)	Bombino bianco (D)	20/0	35.70	14/0		[14]
Beli Debejan (D)	Privlačka bilina (n.d.)	♀ Bombino bianco (D)	19/0	35.35	15/0		
Muškatel (n.d.)	Dattier de Beyrouth = Afuz Ali	Perle de Csaba )	20/0	34.97			
Mejsko bijelo (A)	Duranija (A)	Žumić (A)	20/0	34.77	8/1	45/1	
Lipovina (n.d.)	Moslavac (C)	Tzimliansky belyi	20/1	34.67			
Pošip b. (C)	♀ Zlatarica blatska (C)	Bratkovina bijela (D)	19/0	34.47	13/1	45/1	[63]
Moslavac (C)	Alba imputotato	Belina starohrvatska (C)	20/0	34.43			
Glavinuša (A)	Plavac mali (D)	♀ Vugava (A)	20/0	34.05	14/2		[14]
Crnka (D)	Plavac mali (D)	Bratkovina crvena (D)	20/0	33.30	14/0	45/0	
Volovina (n.d.)	Sciaccarello = Mammolo	Muškat momjanski=Muscat a pet. grains	20/0	32.94			[14]
Belina šemnička (C)	Kovacs Kreger	Belina starohrvatska (C)	18/0	32.51			
Bljuzgavac (D)	Bratkovina crvena (D)	Gyöngy feher	20/0	32.31			[14]
Plavec žuti (C)	Bljuzgavac (D)	♀ Belina starohrvatska (C)	20/0	31.63	12/1		
Plavina istarska (C)	Pinella bianca	Bljuzgavac (D)	20/0	29.55			
Ranfjol (C)	Bljuzgavac (D)	♀ Belina starohrvatska (C)	20/0	28.84	13/2	45/0	[14]
Svjetljak (C)	Bljuzgavac (D)	♀ Belina starohrvatska (C)	20/0	27.80	12/2	45/1	

\* chlorotype; n.d. = not determined.

A total of 113 half kinships (varieties sharing at least one allele at each of the 20 SSR loci) was reconstructed, among which: (i) 103 cases of putative direct (first-degree) relationships (with no possibility to determine if a cultivar is a parent, an offspring, or a full sibling of the second cultivar), out of which 61 relationships were not previously reported; and (ii) five half-sib relationships, of which two were not reported previously, as well as (iii) five not possible relationships due to the already known parentages, out of which two were not reported previously. The full list of first-degree relationships are given in Table S12, with some examples highlighted (Table 6).

**Table 6.** Examples of possible direct (first-degree) relationships identified within Croatian grape germplasm using 20 SSR.

Cultivar 1	Cultivar 2	SSR Genotyped/ Mismatch	LOD Score	Comments	
Bljuzgavac	Žlahtina	20/0	13.22	Possible only if cultivar 1 is genitor of cultivar 2	
Bašćan	Bombino bianco	20/0	12.31		
Belina Hižakovec	Kozjak bijeli	19/0	15.42	Cultivar 2 grandparent of cultivar 1	
Belina Mala	Moslavac	20/0	7.63		
Belina Mala	Sacy [14]	20/0	9.16		
Belina Šemnička	Kozjak bijeli	18/0	12.13		
Bogdanuša	Palagružanka	20/0	11.85		
Bombino bianco	Glavanjuša	20/0	10.68		
Bombino bianco	Mladenka	20/0	15.63		
Bombino bianco	Žutozelen	19/1	13.61		
Bratkovina bijela	Pošip crni	19/0	13.61		
Dišeća ranina	Kozjak bijeli	19/0	7.24		Possible only if cultivar 2 is genitor of cultivar 1
Dišeća ranina	Urmi dinka [14]	19/0	10.8		
Draganela	Plavčina	20/0	15.22		
Drnekuša mala	Plavac mali	20/0	15.66		
Glavanjuša	Plavac mali	20/0	20.41		
Jarbola	Belina starohrvatska	20/0	9.37		
Klešćec	Argant [14]	20/0	19.87		
Malvazija župska	Belina starohrvatska	20/0	11.45		
Plava lovara	Plavac mali	19/0	16.37		
Plavac mali	Siložder	20/0	15.24		
Plavac mali	Šemperinka bijela	18/0	16.1		
Plavac mali	Vugava b. omiška	19/0	17.87		
Plavac mali	Zelenjak	20/0	18.64		
Pošip crni	Vugava bijela omiška	19/0	11.38		
Pršljivka	Alba imputotato [14]	20/0	11.28		
Sansigot	Tribidrag	20/0	13.87		
Silbijanac	Alba imputotato [14]	20/0	17.97		
Škrlet	Argant [14]	20/0	18.17		
Teran	Barbera [14]	17/0	6.05		
Teran	Greco nero [14]	20/0	19.77		

Varieties that are most interconnected within the presumable Croatian set through first-degree relationship are: 'Plavac mali' with a total of twelve assigned relationships, 'Bombino bianco' with a total of seven assigned relationships, 'Bljuzgavac' with a total of seven and 'Tribidrag' with a total of six assigned first-degree relationships. Within Croatian germplasm, for 71 varieties (56%) putative direct (first-degree) relationships was determined (most probably parent-offspring relationship) with one of the varieties from the total analyzed set. The aforementioned degree of relationship, but within Croatian germplasm only, was determined for 58 varieties (45%). For six varieties, 'Mijajuša', 'Mladenka', 'Muškattel', 'Silbijanac', 'Teran', and 'Žutozelen', possible first-degree relationships were determined exclusively with varieties of foreign origin.

## 4. Discussion

### 4.1. Determined Status of Croatian Grapevine Collections

Although Croatian germplasm has been genotyped via SSRs through various projects [36,57–61], this research represents the most comprehensive analysis of Croatian germplasm at the national level so far. Here, we covered all official grapevine repositories and defined latest status of all so far collected accessions in Croatia (Table S5). The process of germplasm collection requires years of

reasoned management with occasional genetic verification of true to typeness, e.g., intruders have been identified within the National Collection of Autochthonous Varieties (Faculty of Agriculture in Zagreb) and were consequently excluded. Fingerprinting has proven particularly useful for the Collection of Istrian varieties (Poreč). Although collected in local vineyards through various conservation projects, Istrian accessions have not been adequately genotyped and compared with international databases [113]. Open issues of the mentioned collection were answered and defined mostly through mislabeled accession / incorrect collection (Table S5). Each germplasm collection will be able, based on the fingerprinting data, to review their duplicates/redundant or unnecessary accessions internally and make cost rationalization decisions. From the same data, the occurrence of duplicates between collections is visible, but such duplications are preferred for security reasons and, moreover, are standard procedure in most germplasm collections [114]. Finally, it should be noted that determining the status of "duplicates" in a collection using microsatellites does not necessarily mean that these accessions are phenotypically identical since they can represent different clones. Identified errors in the internal SSR database, as well as the revision and complementation of data on profiles at nine common SSR loci, enabled the correction of data for 20% of Croatian varieties in which profiles are stored within the largest public database of grapevine varieties, i.e., the European *Vitis* database. The accuracy and completeness of genetic profiles ensures proper estimation of the overall grapevines genetic variability [7].

#### 4.2. Choice of Markers for Routine Fingerprinting

Using the M13-tailed primers, VViv37 locus could not be amplified. Successful amplification of the VViv37 locus was achieved only with primers that were both fluorescently labeled. Since this problem was addressed at a later research stage, weak amplification of the aforementioned locus (Table 2) was probably due to DNA degradation caused over time. Optimization of the PCR reaction with the M13 approach (proposed as cost effective) showed to be challenging due to the method requirements, primarily its temperature regime [76] and length difference of the reverse and forward primer.

On average, the amplification rates for tri- and tetra-nucleotide loci were poorer compared to other loci, as well as their detected allelic richness, PIC, and  $P_{(ID)}$  values (Table 2), although the last three parameters were similar to those previously recorded [73]. In addition, the most polymorphic loci showed a high occurrence of null alleles, which has been observed also in other studies [10,73]. This can be an issue, especially in pedigree reconstructions where such loci need to be neglected. One of the main advocated advantages of three- and tetra-nucleotide markers is easier allele determination due to the smaller stuttering effect and easier binning. However, for certain loci, alleles different from the product of basic repeat number were detected, such as Vchr2b, Vchr11b, and Vchr16a locus (Table S7). As confirmed in the present study, irregularities in allele lengths at the Vchr11b and Vchr16a loci were identified earlier [73].

Loci recommended [13] and used in genotyping of world's largest germplasm collection [14], in our study, proved to be equally effective, e.g., the two most informative loci in both studies were VVip31 and VVMD28. The cumulative  $P_{(ID)sib}$  value for the nine common SSRs today routinely used for genotyping was  $1.44 \times 10^{-4}$ , and the same efficacy can be achieved using top eight most informative microsatellite markers (VVMD28, VVip31, ssrVrZAG79, ssrVrZAG62, VViv67, VVMD5, VMC1b11, and VVMD32), except for the Vchr8b locus due to its high frequency of null alleles (Table 2). Given their high polymorphism, as well as other advantageous properties, like ease of amplification and well-known lab protocols and alleles expected, the recommended set of 9 SSRs proved once again to be a valuable and practical tool in grapevine fingerprinting.

#### 4.3. Will SNP Markers Replace SSRs for Routine Identification of Grapevine Varieties?

In the last five years, SNPs have been the markers of choice in grapevine research, as well as for other crops [43,115–117]; however, still, the most used type of markers for routine fingerprinting are SSRs. An alternative to the common set of 9 SSRs in the form of highly reproducible and well scattered

48 SNP markers was proposed [36]. In this research, 47 of the 48 proposed SNP loci were analyzed, and 45 of them have been used for further analyzes. Although SNP579\_187 locus was polymorphic in other studies [38,40] where minor allele frequency (MAF) value as 0.17 was reported [38], in this study, it was nonpolymorphic, i.e., uninformative for Croatian germplasm. This was less often the case with microsatellite loci since those selected for fingerprinting detected more than just two alleles per locus. PIC values were, on average, 2.5 times higher for 9 SSRs (PIC = 0.76) than for SNP markers (PIC = 0.31), which is in agreement with previous results [20,38,118]. Comparing the SNP and SSR loci based on  $P_{(ID) \text{ unrelated}}$  (Table 4), it is evident that one SSR locus provides on average information as 3–4 SNP loci, respectively, which is identical to earlier findings [38]. However, bearing in mind the complexity of grapevine kinship [41], it is more appropriate to use a  $P_{(ID) \text{ sib}}$  value indicating that 18 SSR markers provide the same information as 45 SNP loci, that is, one SSR locus provides information as 2.5 SNP loci. As genetic similarity among genotypes increases, so does the informative nature of the SNPs.

Large grapevine collections have genotyped their germplasm in the last decade [20,38,41], but resulting SNP profiles are generally not publicly available. It can be assumed that, without an international consortium initiative to form a public SNP profile database, international comparison will not be possible, as also stressed by others [43]. It is likely that smaller laboratories and research groups will continue to use microsatellite markers for routine genotyping (e.g., genotype checking in nursery production) due to lower equipment costs and the abundance of available microsatellite profiles of already analyzed grapevine varieties on a standard 9-loci set.

#### 4.4. Links to Other Countries through Synonyms

For a third of genotypes within Croatian collections, a synonym was detected in one of the neighboring or more distant countries (Table 1 and Table S5). For a total of 95 accessions, no synonym has been detected in any other country, so we can consider them to be unique for Croatian germplasm (Table 1 and Table S5). Although the analysis of the genetic structure suggested that origin of four unique varieties ('Bogdanuša', 'Trnjak', 'Krivaja crvena', and 'Cibib') was from foreign gene pools, since they were not found in areas of their presumed origin and have a long history in Croatia, they can still be considered as part of the Croatian grapevine heritage. Conservation of these genotypes should be accentuated given their uniqueness. In addition, additional 7 varieties ('Garganja', 'Draganela', 'Hrvatica', 'Dolčin', 'Bilan', 'Tribidrag', and 'Bratkovina bijela') might have Croatian origin because they have at least one Croatian variety as a first-degree relationship (most likely potential parent offspring pair), and other researchers have not reconstructed other first-degree relationships for the same.

For varieties of coastal area, links with Italian, Slovenian, and, to a lesser extent, Greek germplasm have been established through identified synonyms. The similar results were suggested in earlier comparisons of germplasms [11,57,60,62]. Historically and climatically, these Mediterranean countries are interconnected by the Adriatic Sea. The influence of the Greek germplasm probably extends from the time of the first ancient Greek colonization of Adriatic islands [65] till today in presence of varieties like 'Mijajuša' or 'Cipar'. Neighboring Bosnia and Herzegovina (B&H) is a country where many Croatian varieties are grown under the same name, e.g., as 'Plavac mali' or 'Plavina' from Croatia, or reciprocal example 'Žilavka' from B&H in Croatia [33]. Synonyms of varieties from the continental part of Croatia have been confirmed primarily in Austria and Germany. The fact that numerous Muscat accessions grown in Croatia are synonyms of some well-known Muscat varieties has been confirmed in this study, as well as in previous ones, such as 'Muškat ruža porečki' = 'Rosenmuskateller' [57], 'Muškat momjanski' = 'Muscat White' [119], or 'Muškat ruža omiški' = 'Muscat Hamburg' [59,60]. To what extent Italian or Greek germplasm had impact on the formation of Croatian germplasm is evident from the established genetic structure and kinship, which will be discussed later. Numerous established and confirmed synonyms are probably the result of longtime migration of varieties via trade or migration routes. Migration direction is often difficult to prove, so their exact origin remains an open question.

#### 4.5. Chlorotype Variation and Geographical Distribution in Croatia

The most extensive reconstruction of the chlorotypes of Croatian germplasm has detected seven different chlorotypes, revealed certain differences in their geographical distribution (Figure 2) and contributed to the affirmation of reconstructed pedigrees where chlorotypes of the two parents were distinct. Abundance in chlorotype diversity was found, including infrequent chlorotypes, such as E, F, or H, that were previously found specific for Near East *sylvestris* populations [53]. Chlorotypes of two accessions ('Teran' and 'Bogdanuša') did not match those previously established [53] (Table S6). Determined chlorotypes allowed us to discriminate two varieties, 'Belina starohrvatska' and 'Bombino bianco', as key maternal genitors in the formation of Croatian germplasm. Although the three most extensive grapevine pedigree reconstructions [14,19,41] did not include chloroplast DNA analysis in their methodology, in this research, it proved to be a useful diagnostic and supportive tool in pedigree reconstructions. Greater diversity of chlorotypes in coastal Croatia is correlated with greater genetic variability determined at nuclear DNA level (data not shown). Continental Croatia is dominated by chlorotype C, the one which is also dominant for a group of French and German varieties [53]. Chlorotype A was not recorded for varieties of continental Croatia, although its share in coastal germplasm was 17%. Chlorotype A is a dominant genotype of European *sylvestris* populations, and its distribution in *sativa* populations diminishes from west to east [53]. One possible explanation for the absence of this chlorotype in the continental part is the dominance of 'Belina starohrvatska' (chlorotype C) as the mother and the absence of *sylvestris* introgression. In contrast to our coastal part where chlorotype C is dominating, in Italy, chlorotype C is represented in a small percentage, with chlorotype D being the dominant one [53,120]. This is intriguing given the existence of numerous connections and links with Italian germplasm.

#### 4.6. Genetic Diversity and Structure of Croatian Germplasm

Even though most grapevine varieties are self-fertilizing, they are highly heterozygous as also confirmed in this paper. Average observed heterozygosity calculated for Croatian germplasm (36 SSRs) was 0.71 (Table 2), but when considering only 20 SSR common loci used for same parameter calculated for the Vassal Collection ( $H_O = 0.76$ ;  $n = 2096$ ) [15], the  $H_O$  value was the same. Average number of alleles per locus ( $N_a = 9.6$  for common 20 loci, data not shown;  $N_a = 9$  for 36 SSRs) was expectedly lower than that calculated for entire INRA collection ( $N_a = 16.1$ ). However, detected  $N_a = 8.9$  for group S-3.3 (Wine = Balkans & East Europe,  $n = 226$ ) [15] is somewhat lower than values of the Croatian set ( $N_a = 9.6$ ). In fact, only eastern populations of table grape exhibit significantly more alleles per locus, so we can state that genetic diversity of Croatian germplasm is rather high considering its size. This is also supported by SNP genotyping results, where  $H_O$  values in this study ( $H_O = 0.416$ ) were higher than those ( $H_O = 0.397$ ) previously determined [38] for same loci (but for a diverse panel of varieties from Imidra (Spain) repository) and those ( $H_O = 0.349$ ) determined by Reference [20] for a different SNP panel.

In order to gain insight into the genetic structure, it is necessary to include as many varieties as possible, as well as polymorphic, unlinked markers. Having in mind that the grape germplasm of a certain area does not represent necessarily a natural population, published profiles of only traditional varieties from the Vassal collection were included to reduce human impact through modern breeding. According to Evanno's  $\Delta K$  statistics (Figure S1)  $K = 5$  was indicated as the most pertinent level of population subdivision and the assumed genetic structure was confirmed by molecular variance analysis (Table S11). As the optimal subdivision(s) of the Vassal collection,  $k = 3$  and  $k = 5$  were found [15]. For  $k = 5$ , two groups were defined as in the present study (Balkan and Iber group), and the largest difference was obtained for grouping of Western and Central European varieties that, unlike in this study (populations Biscay and Gallica), grouped into one group. In this study, 41% of varieties were admixed according to  $Q$  values. If we would use stricter criteria,  $Q \geq 0.85$ , as in Reference [15], the proportion of varieties of admixed origin for our established  $k$  rises to 53% versus

62%, in accordance with Reference [15]. The reason for the large percentage of varieties of admixed origin can be found in complex pedigrees and migration of grapevine material.

Almost half of the Croatian germplasm was assigned to population 1 designated Balkan (Table S10), in accordance with the eco-geographic grouping to proles Pontica (subproles *Balkanica*) proposed by Reference [121], thus confirming previous reports for a limited number of Croatian varieties that classified them as part of Balkan and East Europe [15] or Balkan only [43]. It is important to note that both continental and coastal varieties were grouped into the same Balkan population. Given the many links with Italian germplasm, it was expected that some varieties would be grouped with Italian ones. However, analyses of genetic structure [15,20] define the Italian population as admixed one and emphasize the role of the Roman Empire in the expansion of viticulture and, consequently, of varieties itself. The Balkan group shows the greatest similarity to the ancestral population East (Table S11), which agrees with previous results [43] and is the result of the grapevine migration via Mediterranean trade routes throughout history.

Cluster analysis performed based on 36 SSRs (Figure 3) implies the existence of a structure within Croatian germplasm. A group of Dalmatian varieties is most homogeneous and formed around key varieties, such as 'Plavac mali' and 'Tribidrag', and the clusters follow established kinship relationships. In addition, a cluster containing both southern and northern Adriatic varieties grouped on the basis of kinship with the 'Bombino bianco' variety is clearly distinguished, as well as a large cluster consolidating descendants of 'Belina starohrvatska' and 'Bljuzgavac'. Varieties that were not part of Balkan ancestral group also exhibit lower genetic similarity with the rest of Croatian germplasm and are thus separated into different clusters, such as cluster of Eastern (presumably Greek) varieties: 'Mijajuša', 'Krivaja crvena', 'Trnjak', and 'Brunac'.

#### 4.7. Origin of Presumably Croatian Varieties

All proposed parentages of Croatian varieties published in reference [14] were confirmed in this study, partly at additional SSR and SNP loci (Table 5). Full parentage was reconstructed for 24 varieties that make up 19% of the putative set of 127 Croatian varieties. For the Vassal collection, this was possible for as much as 35% of varieties [14]. The fact that almost twice as much pedigree trios could be detected for the world's largest grapevine collection can be partly explained by the fact that half of them refer to modern varieties, created in the last 150 years, mostly through crossing of well-established and widespread varieties, which can, of course, still be found in germplasm collections. Discovery and confirmation of numerous parentages within French germplasm speaks in favor of well-preserved regional germplasm and suggests that inclusion of regional varieties within a candidate parent set increases the likelihood for successful parentage reconstruction. Published SSR profiles [14] originating from neighboring countries, such as Italy and Hungary, represent only a small fraction of their germplasm, whilst some countries, such as B&H or Serbia, are represented with just a few, and neighboring Slovenia with no accessions. In only five parentages, all three members were part of Croatian germplasm (not including 'Belina starohrvatska'), indicating that the number of reconstructed parentages could be higher if broader regional germplasm would be included. For a total of 45% of the varieties, at least one first-degree relationship (most probably parent-offspring relationship) with one of the supposed autochthonous varieties was detected, and for some of them first-degree relationships with foreign varieties were also detected. Therefore, assumption that most of the Croatian germplasm included in the research originated on the territory of present-day Croatia could not be confirmed, although most of the varieties today exist exclusively here. The low number of reconstructed parentages points to severe genetic erosion of Croatian germplasm since the late 19th century, as well as to the importance and justification for preserving neglected local varieties, although most of them are not economically interesting at present.

## 5. Conclusions

This study is so far the most representative and comprehensive analysis of Croatian grape germplasm. Although Croatia is a small wine growing country, the grape genetic diversity of its germplasm is high. SSR marker data enabled identification of many duplicates and synonyms both within and between collections, which will improve curating the germplasm repositories. The selected 45 SNP loci, previously recommended for fingerprinting, have proven to be a good identification tool, but their future relevance remains questionable due to the absence of public databases for comparison. So, we believe that the broadly accepted 9 SSRs set will remain the gold standard in varietal fingerprinting. The inferred genetic structure has classified 50% of Croatian germplasm as part of the Balkan group of cultivars in accordance with the eco-geographic grouping. The large number of admixed varieties is a consequence of complex pedigrees and migration of varieties over wider territory surrounding Croatia today. For a total of 95 accessions, no synonym has been detected in any other country, so we can consider them to be unique to Croatian germplasm. However, complete parentage reconstruction was possible for only 19% of varieties. This confirms the severity of genetic erosion in Croatian germplasm since the end of the 19th century. Contrary to the expectations, foreign varieties, i.e., genotypes that we failed to find within present day Croatian germplasm, participated in many parentages, and the number of cases where both parents are exclusively of Croatian origin is rather low. Thus, considering the rich political history of the entire region and changes of administrative borders, vicinity and similarity with particular geographical and agro-ecological areas of neighboring countries seems to be more relevant for genetic structure of germplasm than administrative borders of single countries. Studied varieties had been derived from spontaneous crosses of the same parental combinations or just with few key genitors, rather than out of many chance combinations. The empirical selection process that has been performed in the past by humans was probably a result of the unconscious preference for high yielding or more adaptive cultivars. Interestingly, almost all key parents are today critically endangered varieties or have disappeared from our vineyards. All this data points to the importance and justification of preserving neglected, local varieties in germplasm repositories, in spite of their present low economic significance.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4425/11/7/737/s1>, Figure S1: Determination of K values according to Reference [109], Evanno et al. (2005). The rate of change of the posterior probability of the data given the number of subgroups is plotted against K. The first peak (K = 5) corresponds to the optimum number of subgroups. Table S1: List of 212 *Vitis vinifera* L. subsp. *sativa* accessions including their codes, presumed varietal names, skin color, grape utilization, sampling collection, main cultivation region, conservation status and types of analyses conducted, Table S2: List of reference *Vitis* accessions, Table S3: Primer sequences, chromosome positions, fluorophore, range, associated annealing temperatures and PCR and capillary electrophoresis arrangement of SSR and cpSSR loci used in this study, Table S4: 47 SNP loci recommended for identification analyzed in this study and expected polymorphism, Table S5: Determined status of analyzed accessions, Table S6: Microsatellite profiles for 36 SSR loci and chlorotypes for 127 presumed Croatian varieties, Table S7: SSR allele frequencies for 127 presumed Croatian varieties, Table S8: 46 SNP profiles for 124 presumed Croatian varieties, Table S9: Determined chlorotype haplotypes based on 4 cpSSR markers. Table S10: Genome share of all varieties from the putative Croatian set in each of the five proposed gene pools. Q values > 0.75 are highlighted, Table S11: AMOVA results and pairwise population  $\Phi_{PT}$  values, Table S12: Complete list of Croatian cultivars and other *Vitis vinifera* cultivars in possible direct (first-degree) relationship.

**Author Contributions:** Conceptualization, M.Ž.M. and I.P.; formal analysis, M.Ž.M.; resources, E.M., G.Z., D.P., M.B. E.Z., I.P., M.Ž.M.; writing—original draft preparation, M.Ž.M.; writing—review and editing, I.P., E.M., G.Z., D.P., M.B., E.Z.; funding acquisition, I.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Ministry of Science and Education, Republic of Croatia, project “Genetic fingerprinting and genetic variability of grape and fruit cultivars” and partially by Cost Action FA1003 grant—“East-West Collaboration for Grapevine Diversity Exploration and Mobilization of Adaptive Traits for Breeding” and by the project KK.01.1.01.0005 Biodiversity and Molecular Plant Breeding Centre of Excellence for Biodiversity and Molecular Plant Breeding (CoE Crop-BioDiv), Zagreb, Croatia.

**Acknowledgments:** SNP fingerprinting was performed at the Julius Kuehn Institute, Federal Research Centre for Cultivated Plants, Institute for Grapevine Breeding - Geilweilerhof, Siebeldingen, Germany. The authors would like to acknowledge the given opportunity and technical assistance. Authors thank Carole P. Meredith for critical reading.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. FAOSTAT (2019) Statistics Division of the Food and Agriculture Organization (FAO) of the United Nations, Rome. Available online: [www.fao.org/faostat/en/#data/QC](http://www.fao.org/faostat/en/#data/QC) (accessed on 9 March 2020).
2. Adam-Blondon, A.F.; Martinez-Zapater, J.M.; Kole, C. *Genetics, Genomics, and Breeding of Grapes*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2011; pp. 1–30.
3. This, P.; Lacombe, T.; Thomas, M.R. Historical origins and genetic diversity of wine grapes. *Trends Genet.* **2006**, *22*, 511–519. [[CrossRef](#)]
4. Anderson, K.; Aryal, N.R. *Which Winegrape Varieties are Grown Where?* University of Adelaide Press: Adelaide, South Australia, 2015; pp. 4–5.
5. Hares, D.R. Selection and implementation of expanded CODIS core loci in the United States. *Forensic Sci. Int. Genet.* **2015**, *17*, 33–34. [[CrossRef](#)]
6. This, P.; Jung, A.; Boccacci, P.; Borrego, J.; Botta, R.; Costantini, L.; Crespan, M.; Dangl, G.S.; Eisenheld, C.; Ferreira-Monteiro, F.; et al. Development of a standard set of microsatellite reference alleles for identification of grape varieties. *Theor. Appl. Genet.* **2004**, *109*, 1448–1458. [[CrossRef](#)]
7. Maul, E.; Sudharma, K.N.; Kecke, S.; Marx, G.; Müller, C.; Audeguin, L.; Bosell, M.; Boursiquot, J.M.; Bucchetti, B.; Cabello, F.; et al. The European *Vitis* Database ([www.eu-vitis.de](http://www.eu-vitis.de))—a technical innovation through an uploading and interactive modification system. *Vitis* **2012**, *51*, 79–85.
8. European *Vitis* Database. Available online: [www.eu-vitis.de](http://www.eu-vitis.de) (accessed on 21 February 2012).
9. Maul, E.; Töpfer, R.; Carka, F.; Cornea, V.; Crespan, M.; Dallakyan, M.; de Andrés Domínguez, T.; de Lorenzis, G.; Dejeu, L.; Goryslavets, S.; et al. Identification and characterization of grapevine genetic resources maintained in Eastern European Collections. *Vitis* **2015**, *54*, 5–12.
10. Bergamini, C.; Perniola, R.; Cardone, M.F.; Gasparro, M.; Pepe, R.; Caputo, A.R.; Antonacci, D. The molecular characterization by SSRs reveals a new South Italian kinship and the origin of the cultivar Uva di Troia. *SpringerPlus.* **2016**, *5*, 1562. [[CrossRef](#)] [[PubMed](#)]
11. Schneider, A.; Raimondi, S.; Pirolo, C.S.; Torello Marinoni, D.; Ruffa, P.; Venerito, P.; La Notte, P. Genetic characterization of grape cultivars from Apulia (Southern Italy) and synonymies in other Mediterranean regions. *Am. J. Enol. Viticult.* **2014**, *65*, 244–249. [[CrossRef](#)]
12. Lacombe, T.; Boursiquot, J.M.; Laucou, V.; Dechesne, F.; Varès, D.; This, P. Relationships and genetic diversity within the accessions related to Malvasia held in the Domaine de Vassal grape germplasm repository. *Am. J. Enol. Viticult.* **2007**, *58*, 124–131.
13. Laucou, V.; Lacombe, T.; Dechesne, F.; Siret, R.; Bruno, J.P.; Dessup, M.; Dessup, T.; Ortigosa, P.; Parra, P.; Roux, C.; et al. High throughput analysis of grape genetic diversity as a tool for germplasm collection management. *Theor. Appl. Genet.* **2011**, *122*, 1233–1245. [[CrossRef](#)]
14. Lacombe, T.; Boursiquot, J.M.; Laucou, V.; Di Vecchi-Staraz, M.; Péros, J.-P.; Patrice, T. Large-scale parentage analysis in an extended set of grapevine cultivars (*Vitis vinifera* L.). *Theor. Appl. Genet.* **2013**, *126*, 401–414. [[CrossRef](#)] [[PubMed](#)]
15. Bacilieri, R.; Lacombe, T.; Le Cunff, L.; Di Vecchi-Staraz, M.; Laucou, V.; Genna, B.; Péros, J.P.; This, P.; Boursiquot, J.M. Genetic structure in cultivated grapevines is linked to geography and human selection. *BMC Plant Biol.* **2013**, *13*, 25. [[CrossRef](#)]
16. Nicolas, S.D.; Péros, J.P.; Lacombe, T.; Launay, A.; Le Paslier, M.C.; Bérard, A.; Mangin, B.; Valière, S.; Martins, F.; Le Cunff, L.; et al. Genetic diversity, linkage disequilibrium and power of a large grapevine (*Vitis vinifera* L.) diversity panel newly designed for association studies. *BMC Plant Biol.* **2016**, *16*, 74. [[CrossRef](#)] [[PubMed](#)]
17. Martín, J.P.; Luis Santiago, J.; Pinto-Carnide, O.; Leal, F.; Del Carmen Martínez, M.; Ortiz, J.M. Determination of relationships among autochthonous grapevine varieties (*Vitis vinifera* L.) in the Northwest of the Iberian Peninsula by using microsatellite markers. *Gen. Resour. Crop Evol.* **2006**, *53*, 1255–1261. [[CrossRef](#)]
18. Zarouri, B.; Vargas, A.M.; Gaforio, L.; Aller, M.; de Andrés, M.T.; Cabezas, J.A. Whole-genome genotyping of grape using a panel of microsatellite multiplex PCRs. *Tree Genet. Genomes* **2015**, *11*, 17. [[CrossRef](#)]

19. Cipriani, G.; Spadotto, A.; Jurman, I.; Di Gaspero, G.; Crespan, M.; Meneghetti, S.; Frare, E.; Vignani, R.; Cresti, M.; Morgante, M.; et al. The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymys and parentages, and reveals a large admixture amongst varieties of different geographic origin. *Theor. Appl. Genet.* **2010**, *121*, 1569–1585. [[CrossRef](#)]
20. Emanuelli, F.; Lorenzi, S.; Grzeskowiak, L.; Catalano, V.; Stefanini, M.; Troggio, M.; Myles, S.; Martinez-Zapater, J.M.; Zyprian, E.; Moreira, F.M.; et al. Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. *BMC Plant Biol.* **2013**, *13*, 39. [[CrossRef](#)]
21. Dangl, G.S.; Mendum, M.L.; Prins, B.H.; Walker, M.A.; Meredith, C.P.; Simon, C.J. Simple sequence repeat analysis of a clonally propagated species: A tool for managing a grape germplasm collection. *Genome* **2001**, *44*, 432–438. [[CrossRef](#)]
22. Aradhya, M.K.; Dangl, G.S.; Prins, B.H.; Boursiquot, J.M.; Walker, M.A.; Meredith, C.P.; Simon, C.J. Genetic structure and differentiation in cultivated grape *Vitis vinifera* L. *Genet. Res.* **2003**, *81*, 179–192. [[CrossRef](#)]
23. Sefc, K.M.; Regner, F.; Glössl, J.; Steinkellner, H. Genotyping of grapevine and rootstock cultivars using microsatellite markers. *Vitis* **1998**, *37*, 15–20.
24. Lopes, M.S.; Sefc, K.M.; Eires Dias, E.; Steinkellner, H.; Machado, M.L.C.; Machado, A.C. The use of microsatellites for germplasm management in a Portuguese grapevine collection. *Theor. Appl. Genet.* **1999**, *99*, 733–739. [[CrossRef](#)]
25. Lopes, M.S.; Dos Santos, M.R.; Dias, J.E.; Mendonça, D.; Da Câmara Machado, A. Discrimination of Portuguese grapevines based on microsatellite markers. *J. Biotechnol.* **2006**, *127*, 34–44. [[CrossRef](#)]
26. Lefort, F.; Roubelakis-Angelakis, K.K. Genetic comparison of Greek cultivars of *Vitis vinifera* L. by nuclear microsatellite profiling. *Am. J. Enol. Viticult.* **2001**, *52*, 101–108.
27. Halász, G.; Veres, A.; Kozma, P.; Kiss, E.; Balogh, A.; Galli, Z.; Szőke, A.; Hoffmann, S.; Heszky, L. Microsatellite fingerprinting of grapevine (*Vitis vinifera* L.) varieties of the Carpathian Basin. *Vitis* **2005**, *44*, 173.
28. Vouillamoz, J.F.; McGovern, P.E.; Ergul, A.; Söylemezoğlu, G.; Tevzadze, G.; Meredith, C.P.; Grando, M.S. Genetic characterization and relationships of traditional grape cultivars from Transcaucasia and Anatolia. *Plant. Genet. Resour. C* **2006**, *4*, 144–158. [[CrossRef](#)]
29. Štajner, N.; Korošek-Koruza, Z.; Rusjan, D.; Javornik, B. Microsatellite genotyping of old Slovenian grapevine varieties (*Vitis vinifera* L.) of the Primorje (coastal) winegrowing region. *Vitis* **2008**, *47*, 201–204.
30. Štajner, N.; Rusjan, D.; Korošek-Koruza, Z.; Javornik, B. Genetic characterization of old Slovenian grapevine varieties of *Vitis vinifera* L. by microsatellite genotyping. *Am. J. Enol. Vitic.* **2011**, *62*, 250–255. [[CrossRef](#)]
31. Dzhambova, T.; Tsvetkov, I.; Atanassov, I.; Rusanov, K.; Martínez-Zapater, J.M.; Atanassov, A.; Hvarleva, T. Genetic diversity in native Bulgarian grapevine germplasm (*Vitis vinifera* L.) based on nuclear and chloroplast microsatellite polymorphisms. *Vitis* **2009**, *48*, 115–121.
32. Bešlić, Z.; Todić, S.; Korać, N.; Lorenzi, S.; Emanuelli, F.; Grando, M.S. Genetic characterization and relationships of traditional grape cultivars from Serbia. *Vitis* **2012**, *51*, 183–189.
33. Tomić, L.; Štajner, N.; Jovanović-Cvetković, T.; Cvetković, M.; Javornik, B. Identity and genetic relatedness of Bosnia and Herzegovina grapevine germplasm. *Sci. Hortic.* **2012**, *143*, 122–126. [[CrossRef](#)]
34. Sefc, K.M.; Lopes, M.S.; Lefort, F.; Botta, R.; Roubelakis-Angelakis, K.A.; Ibáñez, J.; Pejić, I.; Wagner, H.W.; Glössl, J.; Steinkellner, H. Microsatellite variability in grapevine cultivars from different European regions and evaluation of assignment testing to assess the geographic origin of cultivars. *Theor. Appl. Genet.* **2000**, *100*, 498–505. [[CrossRef](#)]
35. Martín, J.P.; Arranz, C.; Castro, I.D.; Yuste, J.; Rubio, J.A.; Pinto-Carnide, O.; Ortiz, J.M. Prospection and identification of grapevine varieties cultivated in north Portugal and northwest Spain. *Vitis* **2011**, *50*, 29–33.
36. Žulj Mihaljević, M.; Šimon, S.; Pejić, I.; Carka, F.; Sevo, R.; Kojić, A.; Gaši, F.; Tomić, L.; Jovanović-Cvetković, T.; Maletić, E.; et al. Molecular characterization of old local grapevine varieties from South East European countries. *Vitis* **2013**, *52*, 69–76.
37. Štajner, N.; Tomić, L.; Ivanišević, D.; Korać, N.; Cvetković-Jovanović, T.; Beleski, K.; Angelova, E.; Maraš, V.; Javornik, B. Microsatellite inferred genetic diversity and structure of Western Balkan grapevines (*Vitis vinifera* L.). *Tree Genet. Genomes* **2014**, *10*, 127–140. [[CrossRef](#)]
38. Cabezas, J.A.; Ibáñez, J.; Lijavetzky, D.; Vélez, D.; Bravo, G.; Rodríguez, V.; Carreño, I.; Jermakow, A.M.; Carreño, J.; Ruiz-García, L.; et al. A 48 SNP set for grapevine cultivar identification. *BMC Plant Biol.* **2011**, *11*, 153. [[CrossRef](#)]

39. Zinelabidine, L.H.; Laiadi, Z.; Benmehaia, R.; Gago, P.; Boso, S.; Santiago, J.L.; Haddioui, A.; Ibáñez, J.; Martínez-Zapater, J.M.; Martínez, M.C. Comparative ampelographic and genetic analysis of grapevine cultivars from Algeria and Morocco. *Aust. J. Grape Wine Res.* **2014**, *20*, 324–333. [[CrossRef](#)]
40. Cunha, J.; Ibáñez, J.; Teixeira-Santos, M.; Brazão, J.; Fevereiro, P.; Martínez-Zapater, J.M.; Eiras-Dias, J.E. Characterisation of the Portuguese grapevine germplasm with 48 single-nucleotide polymorphisms. *Aust. J. Grape Wine Res.* **2016**, *22*, 504–516. [[CrossRef](#)]
41. Myles, S.; Boyko, A.R.; Owens, C.L.; Brown, P.J.; Grassi, F.; Aradhya, M.K.; Prins, B.; Reynolds, A.; Chia, J.-M.; Ware, D.; et al. Genetic structure and domestication history of the grape. *PNAS* **2011**, *108*, 3530–3535. [[CrossRef](#)]
42. De Lorenzis, G.; Chipashvili, R.; Failla, O.; Maghradze, D. Study of genetic variability in *Vitis vinifera* L. germplasm by high-throughput Vitis18kSNP array: The case of Georgian genetic resources. *BMC Plant Biol.* **2015**, *15*, 154. [[CrossRef](#)]
43. Laucou, V.; Launay, A.; Bacilieri, R.; Lacombe, T.; Adam-Blondon, A.F.; Berard, A.; Chauveau, A.; de AndreÂs, M.T.; Hausmann, L.; Ibáñez, J.; et al. Extended diversity analysis of cultivated grapevine *Vitis vinifera* with 10K genome-wide SNPs. *PLoS ONE* **2018**, *13*, e0192540. [[CrossRef](#)]
44. Sefc, K.M.; Steinkellner, H.; Lefort, F.; Botta, R.; Da Câmara Machado, A.; Borrego, J.; Maletić, E.; Glössl, J. Evaluation of the genetic contribution of local wild vines to EU grapevine cultivars. *Am. J. Enol. Viticult.* **2003**, *54*, 15–21.
45. Zdunić, G.; Preece, J.E.; Aradhya, M.; Velasco, D.; Koehmstedt, A.; Dangl, G.S. Genetic diversity and differentiation within and between cultivated (*Vitis vinifera* L. ssp. *sativa*) and wild (*Vitis vinifera* L. ssp. *sylvestris*) grapes. *Vitis* **2013**, *52*, 29–32.
46. Bowers, J.E.; Meredith, C.P. The parentage of a classic wine grape, Cabernet Sauvignon. *Nat. Genet.* **1997**, *16*, 84–87. [[CrossRef](#)] [[PubMed](#)]
47. Goryslavets, S.; Risovanna, V.; Bacilieri, R.; Hausman, J.F.; Heuertz, M. A parentage study of closely related Ukrainian wine grape varieties using microsatellite markers. *Cytol. Genet.* **2010**, *44*, 95–102. [[CrossRef](#)]
48. Ghaffari, S.; Hasnaoui, N.; Zinelabidine, L.H.; Ferchichi, A.; Martínez-Zapater, J.M.; Ibáñez, J. Genetic diversity and parentage of Tunisian wild and cultivated grapevines (*Vitis vinifera* L.) as revealed by single nucleotide polymorphism (SNP) markers. *Tree Genet. Genomes* **2014**, *10*, 1103–1112. [[CrossRef](#)]
49. Mena, A.; Martínez, J.; Fernández-González, M. Recovery, identification and relationships by microsatellite analysis of ancient grapevine cultivars from Castilla-La Mancha: The largest wine growing region in the world. *Gen. Resour. Crop Evol.* **2014**, *61*, 625–637. [[CrossRef](#)]
50. Carka, F.; Maul, E.; Sevo, R. Study and parentage analysis of old Albanian grapevine cultivars by ampelography and microsatellite markers. *Vitis* **2015**, *54*, 127–131.
51. Štajner, N.; Tomić, L.; Progar, V.; Pokorn, T.; Lacombe, T.; Laucou, V.; Boursiquot, J.M.; Javornik, B.; Bacilieri, R. Genetic clustering and parentage analysis of Western Balkan grapevines (*Vitis vinifera* L.). *Vitis* **2015**, *54*, 67–72.
52. Raimondi, S.; Ruffa, P.; Boccacci, P.; Abbà, S.; Schneider, A. A few main parents contributed to the traditional grapevine cultivar assortment in north-western Italy, as revealed by microsatellites and single-nucleotide polymorphisms. In Proceedings of the ISHS Acta Horticulturae 1248: XII International Conference on Grapevine Breeding and Genetics, Bordeaux, France, 15–20 July 2018; Delrot, S., Ollat, N., Gallusci, P., Eds.; Volume 1248, pp. 295–300. [[CrossRef](#)]
53. Arroyo-García, R.; Ruiz-García, L.; Bolling, L.; Ocete, R.; López, M.A.; Arnold, C.; Ergul, A.; Söylemezoglu, G.; Uzun, H.I.; Cabello, F.; et al. Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp. *sativa*) based on chloroplast DNA polymorphisms. *Mol. Ecol.* **2006**, *15*, 3707–3714. [[CrossRef](#)]
54. Zaninović, M. Iliri i vinova loza. *Godišnjak Centra Balkanološka Ispitivanja* **1976**, *13*, 261–272.
55. Bulić, S. *Dalmatinska ampelografija*, 1st ed.; Poljoprivredni nakladni Zavod: Zagreb, Croatia, 1949; pp. 5–343.
56. Pejić, I.; Maletić, E. Conservation, evaluation and revitalization of native grapevine varieties. *Mitt. Klosterneuburg.* **2010**, *60*, 363–368.
57. Maletić, E.; Sefc, K.M.; Steinkellner, H.; Kontić, J.K.; Pejić, I. Genetic characterization of Croatian grapevine cultivars and the detection of synonymous cultivars in neighboring regions. *Vitis* **1999**, *38*, 79–83.
58. Zdunić, G.; Pejić, I.; Karoglan Kontić, J.; Vukičević, D.; Vokurka, A.; Pezo, I.; Maletić, E. Comparison of genetic and morphological data for inferring similarity among native Dalmatian (Croatia) grapevine cultivars (*Vitis vinifera* L.). *J. Food. Agric. Environ.* **2008**, *6*, 333–336.

59. Zdunić, G.; Preece, J.E.; Dangl, G.S.; Koehmstedt, A.; Mucalo, A.; Maletić, E.; Pejić, I. Genetic characterization of grapevine cultivars collected throughout the Dalmatian region. *Am. J. Enol. Viticult.* **2013**, *64*, 285–290. [[CrossRef](#)]
60. Žulj Mihaljević, M.; Anhalt, U.C.M.; Rühl, E.; Tomić Mugoša, M.; Forneck, A.; Zdunić, G.; Preiner, D.; Pejić, I. Cultivar Identity, Intravarietal Variation, and Health Status of Native Grapevine Varieties in Croatia and Montenegro. *Am. J. Enol. Viticult.* **2015**, *66*, 531–541. [[CrossRef](#)]
61. Maletić, E.; Pejić, I.; Kontić, J.K.; Zdunić, D.; Preiner, D.; Šimon, S.; Stupić, D.; Andabaka, Ž.; Žulj Mihaljević, M.; Bubola, M.; et al. Ampelographic and genetic characterization of Croatian grapevine varieties. *Vitis* **2015**, *54*, 93–98.
62. Šimon, S.; Preiner, D.; Maletić, E.; Pejić, I. Genetic similarity among Croatian and Greek grapevine cultivars assessed by SSRs. In Proceedings of the 9th International Conference on Grape Genetics and Breeding, Udine, Italy, 2–06 July 2006; Peterlunger, E., Di Gaspero, G., Eds.;
63. Piljac, J.; Maletić, E.; Karoglan Kontić, J.; Dangl, G.S.; Pejić, I.; Mirošević, N.; Meredith, C.P. The parentage of a major white wine cultivar of Croatia. *Vitis* **2002**, *41*, 83–87.
64. Maletić, E.; Pejić, I.; Karoglan Kontić, J.; Piljac, J.; Dangl, G.S.; Vokurka, A.; Lacombe, T.; Mirošević, N.; Meredith, C.P. Zinfandel, Dobričić, and Plavac mali: The genetic relationship among three cultivars of the Dalmatian coast of Croatia. *Am. J. Enol. Viticult.* **2004**, *55*, 174–180.
65. Maletić, E.; Karoglan Kontić, J.; Pejić, I. *Vinova loza: Ampelografija, ekologija i oplemenjivanje*, 1st ed.; Školska knjiga: Zagreb, Croatia, 2008; pp. 62–109.
66. Thomas, M.R.; Scott, N.S. Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). *Theor. Appl. Genet.* **1993**, *86*, 985–990. [[CrossRef](#)]
67. Bowers, J.E.; Dangl, G.S.; Vignani, R.; Meredith, C.P. Isolation and characterization of new polymorphic simple sequence repeat in loci in grape (*Vitis vinifera* L.). *Genome* **1996**, *39*, 628–633. [[CrossRef](#)]
68. Bowers, J.E.; Dangl, G.S.; Meredith, C.P. Development and characterization of additional microsatellite DNA markers for grape. *Am. J. Enol. Viticult.* **1999**, *50*, 243–246.
69. Sefc, K.M.; Regner, F.; Turetschek, E.; Glössl, J.; Steinkellner, H. Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. *Genome* **1999**, *42*, 367–373. [[CrossRef](#)] [[PubMed](#)]
70. Zyprian, E.; Topfer, R. *Development of microsatellite-derived markers for grapevine genotyping and genetic mapping*; NCBI, GeneBank: Bethesda, MD, USA, 2005.
71. Di Gaspero, G.; Peterlunger, E.; Testolin, R.; Edwards, K.J.; Cipriani, G. Conservation of microsatellite loci within the genus *Vitis*. *Theor. Appl. Genet.* **2000**, *101*, 301–308. [[CrossRef](#)]
72. Merdinoglu, D.; Butterlin, G.; Bevilacqua, L.; Chiquet, V.; Adam-Blondon, A.F.; Decroocq, S. Development and characterization of a large set of microsatellite markers in grapevine (*Vitis vinifera* L.) suitable for multiplex PCR. *Mol. Breed.* **2005**, *15*, 349–366. [[CrossRef](#)]
73. Cipriani, G.; Marrazzo, M.T.; Di Gaspero, G.; Pfeiffer, A.; Morgante, M.; Testolin, R. A set of microsatellite markers with long core repeat optimized for grape (*Vitis* spp.) genotyping. *BMC Plant Biol.* **2008**, *8*, 127. [[CrossRef](#)] [[PubMed](#)]
74. Chung, S.M.; Staub, J.E. The development and evaluation of consensus chloroplast primer pairs that possess highly variable sequence regions in a diverse array of plant taxa. *Theor. Appl. Genet.* **2003**, *107*, 757–767. [[CrossRef](#)]
75. Weising, K.; Gardner, R.C. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* **1999**, *42*, 9–19. [[CrossRef](#)]
76. Schuelke, M. An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* **2000**, *18*, 233–234. [[CrossRef](#)]
77. Hunt, H.V.; Lawes, M.C.; Bower, M.A.; Haeger, J.W.; Howe, C.J. A banned variety was the mother of several major wine grapes. *Biol. Lett.* **2009**, *6*, 367–369. [[CrossRef](#)]
78. Macro Coding Genres. Available online: [http://www1.montpellier.inra.fr/grapegen06/technical/menu\\_activities\\_room\\_index/page\\_wpl/WP1-macro\\_coding\\_genres-25-07-08.xls](http://www1.montpellier.inra.fr/grapegen06/technical/menu_activities_room_index/page_wpl/WP1-macro_coding_genres-25-07-08.xls). (accessed on 21 February 2012).
79. Pastore, C.; Fontana, M.; Raimondi, S.; Ruffa, P.; Filippetti, I.; Schneider, A. Genetic Characterization of Grapevine Varieties from Emilia-Romagna (Northern Italy) Discloses Unexplored Genetic Resources. *Am. J. Enol. Viticult* **2020**, (in press). [[CrossRef](#)]

80. Bodor, P.; Szoke, A.; Toth-Lencses, K.; Veres, A.; Deak, T.; Kozma, P.; Denes Bisztraya, G.; Kiss, E. Differentiation of grapevine (*Vitis vinifera* L.) conculta members based on molecular tools. *Biotechnol. Biotechnol. Equip.* **2014**, *28*, 14–20. [[CrossRef](#)]
81. Costantini, L.; Monaco, A.; Vouillamoz, J.F.; Forlani, M.; Grandi, M.S. Genetic relationships among local *Vitis vinifera* cultivars from Campania (Italy). *Vitis* **2005**, *44*, 25–34.
82. Crespan, M.; Calò, A.; Giannetto, S.; Sparacio, A.; Storchi, P.; Costacurta, A. ‘Sangiovese’ and ‘Garganega’ are two key varieties of the Italian grapevine assortment evolution. *Vitis* **2008**, *47*, 97–104.
83. Crespan, M.; Fabbro, A.; Giannetto, S.; Meneghetti, S.; Petrusi, C.; Del Zan, F.; Sivilotti, P. Recognition and genotyping of minor germplasm of Friuli Venezia Giulia revealed high diversity. *Vitis* **2011**, *50*, 21–28.
84. González, M.F.; Gascueña, J.M.; Morales, A.M. Identification and relationships of grapevine cultivars authorized for cultivation in Castilla La Mancha (Spain). *Am. J. Enol. Viticult.* **2012**, *63*, 564–567. [[CrossRef](#)]
85. Pellerone, F.I.; Edwards, K.J.; Thomas, M.R. Grapevine microsatellite repeats: Isolation, characterisation and use for genotyping of grape germplasm from Southern Italy. *Vitis* **2001**, *40*, 179–186.
86. Ruffa, P.; Raimondi, S.; Boccacci, P.; Abbà, S.; Schneider, A. The key role of “Moscato bianco” and “Malvasia aromatica di Parma” in the parentage of traditional aromatic grape varieties. *Tree Genet. Genomes* **2016**, *12*, 1–14. [[CrossRef](#)]
87. Santana, J.C.; Hidalgo, E.; De Lucas, A.I.; Recio, P.; Ortiz, J.M.; Martín, J.P.; Yuste, J.; Arranz, C.; Rubio, J.A. Identification and relationships of accessions grown in the grapevine (*Vitis vinifera* L.) Germplasm Bank of Castilla y León (Spain) and the varieties authorized in the VQPRD areas of the region by SSR-marker analysis. *Gen. Resour. Crop Evol.* **2008**, *55*, 573–583. [[CrossRef](#)]
88. Maul, E.; Eibach, R.; Zyprian, E.; Töpfer, R. The prolific grape variety (*Vitis vinifera* L.) ‘Heunisch Weiss’ (=‘Gouais blanc’): Bud mutants, “colored” homonyms and further offspring. *Vitis* **2015**, *54*, 79–86.
89. De Lorenzis, G.; Imazio, S.; Rusjan, D.; Vouillamoz, J.F.; Nikolaou, N.; Failla, O.; Scienza, A. Genetic investigation of grapevine varieties ‘Ribolla Gialla’ (Italy), ‘Rebula’ (Slovenia) and ‘Robola’ (Ionian Islands). *Sci. Hortic.* **2013**, *150*, 425–431. [[CrossRef](#)]
90. Crespan, M.; Crespan, G.; Giannetto, S.; Meneghetti, S.; Costacurta, A. Vitouska is the progeny of Prosecco tondo & Malvasia b. lunga. *Vitis* **2007**, *46*, 192–194.
91. Šimon, S.; Maletić, E.; Kontić, J.K.; Crespan, M.; Schneider, A.; Pejić, I.C. Maraština a new member of Malvasia group. In II Int. Symp. "Mediterranean Malvasias", Salina (ME), 33. 2007. Available online: <https://www.bib.irb.hr/350268> (accessed on 22 April 2013).
92. Calò, A.; Costacurta, A.; Maraš, V.; Meneghetti, S.; Crespan, M. Molecular correlation of Zinfandel (Primitivo) with Austrian, Croatian, and Hungarian cultivars and Kratošija, an additional synonym. *Am. J. Enol. Viticult.* **2008**, *59*, 205–209.
93. Rusjan, D.; Bubola, M.; Janjanin, D.; Užila, Z.; Radeka, S.; Poljuha, D.; Pelengić, R.; Javornik, B.; Štajner, N. Ampelographic characterization of grapevine accessions denominated ‘Refošk’, ‘Refosco’, ‘Teran’ and ‘Terrano’ (*Vitis vinifera* L.) from Slovenia, Croatia and Italy. *Vitis* **2015**, *54*, 77–80.
94. Crespan, M.; Cabello, F.; Giannetto, S.; Ibáñez, J.; Kontić, J.K.; Maletic, E.; Pejić, I.; Rodríguez-Torres, I.; Antonacci, D. Malvasia delle Lipari, Malvasia di Sardegna, Greco di Gerace, Malvasia de Sitges and Malvasia dubrovacka-synonyms of an old and famous grape cultivar. *Vitis* **2006**, *45*, 69–73.
95. Park, S.D.E. Excel Microsatellite Toolkit. Computer Program and Documentation Distributed by the Author. 2008. Available online: <http://animalgenomics.ucd.ie/sdepark/ms-toolkit/> (accessed on 20 March 2012).
96. Botstein, D.; White, R.L.; Skolnick, M.; Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **1980**, *32*, 314. [[PubMed](#)]
97. Summers, K.; Amos, W. Behavioral, ecological and molecular genetic analyses of reproductive strategies in the Amazonian dart-poison frog, *Dendrobates ventrimaculatus*. *Behav. Ecol.* **1997**, *8*, 260–267. [[CrossRef](#)]
98. Waits, L.P.; Luikart, G.; Taberlet, P. Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines. *Mol. Ecol.* **2001**, *10*, 249–256. [[CrossRef](#)]
99. Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **2007**, *16*, 1099–1106. [[CrossRef](#)]
100. Peakall, R.; Smouse, P.E. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **2006**, *6*, 288–295. [[CrossRef](#)]
101. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **2012**, *28*, 2537–2539. [[CrossRef](#)]

102. Goudet, J. FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices Version 2.9.3.2. 2001. Available online: <http://www2.unil.ch/popgen/softwares/fstat.htm> (accessed on 20 May 2016).
103. El Mousadik, A.; Petit, R.J. High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. *Theor. Appl. Genet.* **1996**, *92*, 832–839. [[CrossRef](#)]
104. Bowcock, A.M.; Ruiz-Linares, A.; Tomfohrde, J.; Minch, E.; Kidd, J.R.; Cavalli-Sforza, L.L. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **1994**, *368*, 455–457. [[CrossRef](#)] [[PubMed](#)]
105. Minch, E.; Ruiz-Linares, A.; Goldstein, D.; Feldman, M.; Cavalli-Sforza, L.L. *Microsat (version 1.4 d): A Computer Program for Calculating Various Statistics on Microsatellite Allele Data*; University of Stanford: Stanford, CA, USA, 1995.
106. Kumar, S.; Banks, T.W.; Cloutier, S. SNP discovery through next-generation sequencing and its applications. *Int. J. Plant Genom.* **2012**, *2012*, 831460. [[CrossRef](#)] [[PubMed](#)]
107. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [[PubMed](#)]
108. Earl, D.A. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [[CrossRef](#)]
109. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)] [[PubMed](#)]
110. Jakobsson, M.; Rosenberg, N.A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23*, 1801–1806. [[CrossRef](#)] [[PubMed](#)]
111. Ramasamy, R.K.; Ramasamy, S.; Bindroo, B.B.; Naik, V.G. STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springer Plus* **2014**, *3*, 431. [[CrossRef](#)] [[PubMed](#)]
112. Žulj Mihaljević, M.; Maletić, E.; Zdunić, G.; Pejić, I. The key role of Plavac mali in the ancestry of Dalmatian grape cultivars and deconstruction of its parentage. (unpublished; manuscript in preparation).
113. Peršurić, Đ.; Ilak Peršurić, A.S.; Godena, S.; Sinčić, M.; Petrušić, D.; Užila, Z. Ampelographic Description and Sanitary Analysis of Four Istrian Grapevine Varieties (*Vitis vinifera* L.). *Agric. Consp. Sci.* **2012**, *77*, 113–117.
114. Van Hintum, T.J.; Visser, D.L. Duplication within and between germplasm collections. *Gen. Resour. Crop Evol.* **1995**, *42*, 135–145. [[CrossRef](#)]
115. Fresnedo-Ramírez, J.; Yang, S.; Sun, Q.; Cote, L.M.; Schweitzer, P.A.; Reisch, B.I.; Ledbetter, C.A.; Luby, J.J.; Clark, M.D.; Londo, J.P.; et al. An integrative AmpSeq platform for highly multiplexed marker-assisted pyramiding of grapevine powdery mildew resistance loci. *Mol. Breed.* **2017**, *37*, 145. [[CrossRef](#)]
116. Flutre, T.; Bacilieri, R.; Bécavin, I.; Berger, G.; Bertrand, Y.; Boursiquot, J.M.; Fodor, A.; Lacombe, T.; Laucou, V.; Launay, A.; et al. Genome-wide association study of a diverse grapevine panel: Example of berry weight. In Proceedings of the ISHS Acta Horticulturae 1248: XII International Conference on Grapevine Breeding and Genetics, Bordeaux, France, 15–20 July 2018; Delrot, S., Ollat, N., Gallusci, P., Eds.; pp. 227–234. [[CrossRef](#)]
117. De Lorenzis, G.; Mercati, F.; Bergamini, C.; Cardone, M.F.; Lupini, A.; Mauceri, A.; Caputo, A.R.; Abbate, L.; Barbagallo, M.G.; Antonacci, D.; et al. SNP genotyping elucidates the genetic diversity of Magna Graecia grapevine germplasm and its historical origin and dissemination. *BMC Plant Biol.* **2019**, *19*, 7. [[CrossRef](#)] [[PubMed](#)]
118. Lijavetzky, D.; Cabezas, J.A.; Ibáñez, A.; Rodríguez, V.; Martínez-Zapater, J.M. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. *BMC Genom.* **2007**, *8*, 424. [[CrossRef](#)] [[PubMed](#)]
119. Gluhić, D.; Peršurić, Đ.; Bubola, M.; Peršurić, A.S.I. Ampelographic characteristics of autochthonous population muškati momjanski bijeli. In Proceedings of the Drugi Hrvatski Oplemenjivački i Sjemenarski kongres, Poreč, Croatia, Sjemenarstvo, 23–27 May 2006; Kolak, I., Ed.; Hrvatsko Agronomsko Društvo: Zagreb, Croatia, 2006; Volume 23, Suppl. 1, pp. 62–63.
120. Salmaso, M.; Vannozzi, A.; Lucchin, M. Chloroplast microsatellite markers to assess genetic diversity and origin of an endangered Italian grapevine collection. *Am. J. Enol. Viticult.* **2010**, *61*, 551–556. [[CrossRef](#)]
121. Negrel, A.M. Evolucija kuljturnyx form vinograda. *Dokl. Akad. Nauk. SSSR* **1938**, *8*, 585.

