diversity

MDPI

Article

Genetic Diversity and Population Structure in a *Vitis* spp. Core Collection Investigated by SNP Markers

Davide Bianchi, Lucio Brancadoro and Gabriella De Lorenzis *

Department of Agricultural and Environmental Sciences, University of Milan, via G. Celoria 2, 20133 Milano, Italy

* Correspondence: gabriella.delorenzis@unimi.it

Received: 21 February 2020; Accepted: 14 March 2020; Published: 16 March 2020

Abstract: Single nucleotide polymorphism (SNP) genotyping arrays are powerful tools to measure the level of genetic polymorphism within a population. The coming of next-generation sequencing technologies led to identifying thousands and millions of SNP loci useful in assessing the genetic diversity. The *Vitis* genotyping array, containing 18k SNP loci, has been developed and used to detect genetic diversity of *Vitis vinifera* germplasm. So far, this array was not validated on non-*vinifera* genotypes used as grapevine rootstocks. In this work, a core collection of 70 grapevine rootstocks, composed of individuals belonging to *Vitis* species not commonly used in the breeding programs, was genotyped using the 18k SNP genotyping array. SNP results were compared to the established SSR (Simple Sequence Repeat) markers in terms of heterozygosity and genetic structure of the core collection. Genotyping array has proved to be a valuable tool for genotyping of grapevine rootstocks, with more than 90% of SNPs successfully amplified. Structure analysis detected a high degree of admixed genotypes, supported by the complex genetic background of non-*vinifera* germplasm. Moreover, SNPs clearly differentiated non-*vinifera* and *vinifera* germplasm. These results represent a first step in studying the genetic diversity of non-conventional breeding material that will be used to select rootstocks with high tolerance to limiting environmental conditions.

Keywords: admixture; genetic diversity; non-*vinifera* germplasm; rootstock; viticulture

1. Introduction

Vitis vinifera cultivars, the most important economic fruit species in the modern world, is usually grown on rootstocks (a mixture of non-*vinifera* grapevine species and hybrids) due to its susceptibility to phylloxera attack, a homopteran insect (*Daktulosphaira vitifoliae* Fitch) that feeds on the *V. vinifera* roots [1]. Nevertheless, rootstocks play a key role in the adaptation of vines to the environmental conditions, affecting the production and the quality of grape and wines. Several studies report an effect of rootstocks on limestone tolerance [2,3], nutrients uptake [4], and water stress tolerance [5]. Although relevant efforts in grapevine rootstock selection were made on the turn of the 20th century, only a few genotypes found a large spread in vineyards and nowadays more than the 90% of *V. vinifera* varieties are grown grafted onto less than 10 rootstocks [6], with negative consequences on the tolerance to biotic and abiotic stresses [3] and on genetic diversity. It was already demonstrated that the genetic background of rootstock germplasm is narrow, traceable in a limited number of species. Based on Riaz et al. [7] results, three genotypes of three *Vitis* species contributed to the 39% of rootstock genetic diversity. In this context, rootstock collections represent a relevant starting-point for new breeding programs, aimed to select new promising genotypes able to face the environmental challenges of modern viticulture.

Whilst much work has been performed to study the genetic diversity of large *V. vinifera* germplasm collections [8–11], hardly any information is available on the genetic identification of non-

vinifera germplasm [7,12–18]. Recently, the grapevine rootstock collection of the University of Milan (Italy), composed of 379 accessions and including the largest part of the rootstock germplasm currently available worldwide, has been genotyped by SSR (Simple Sequence Repeats) to investigate genetic diversity, infer population structure, analyze pedigrees, and design a core collection. [19]. Molecular analysis identified 232 unique genotypes with a high level of admixture and a narrow genetic background. Among the 232 unique genotypes, 70 genotypes were selected to be included in a core collection designed to capture the entire allelic richness of the non-*vinifera* collection. Some of these genotypes are *berlandieri* × *rupestris* and *berlandieri* × *riparia* varieties (7%); some others have *labrusca* and *vinifera* parentage (30%), but most of them are individuals still not genetically identified or poorly characterized by the ampelographic and agronomic point of view, making this core collection even more interesting as new materials for the further breeding programs.

So far, SSR markers were one of the most reliable and robust tool used for the genetic characterization of *vinifera* and non-*vinifera* germplasm, widely adopted for their high degree of information provided by the large number of detected alleles per locus [8–20]. Recently, SNP (Single Nucleotide Polymorphism) markers have rapidly gained high popularity in the scene of *V. vinifera* molecular genetics [10,21–26]. The number of SNP loci used to study the genetic diversity increased as the technologies to detect them in the genome changed. Prior to the emergence of next-generation sequencing (NGS) technologies, SNP sets included tens [21] or hundreds [10] loci. With the coming of NGS technologies, the number of SNP loci rapidly increased up to thousands: 10k [27], 18k [24], and 37k [28] SNPs. Their popularity is mainly due to the abundance in the genome (they are the most abundant polymorphisms among the individuals of the same species), amenability to high-throughput detection and high reproducibility, since normalization with reference varieties is not required [23]. These molecular markers are widely used to study genetic diversity and to dissect complex traits via QTLs (Quantitative Traits Loci) or GWASs (Genome-Wide Association Studies) for a breeding program [24,29,30].

The most used SNP set is the Vitis18kSNP array, which was set up by the GrapeReSeq Consortium, re-sequencing the genome of 47 *V. vinifera* genotypes and 18 American genotypes, belonging to the species *Vitis aestivalis, Vitis berlandieri, Vitis cinerea, Vitis labrusca, Vitis lincecumii,* and *Muscadinia rotundifolia*. In this project, a total of 18,071 SNPs were selected, a third of which (4510 SNPs) identified in the Northern American species genome [24]. Several studies validated the 18k SNP set for the evaluation of genetic diversity in *V. vinifera* [19,23,24,26,31–34], but the Vitis18kSNP array could also represent a potentially effective tool for rootstock characterization, due to the consistent number of SNPs detected in the Northern American species genome.

The aim of this study was to validated the Vitis18kSNP genotyping array on non-*vinifera* germplasm core collection [19], representing the whole genetic diversity of grapevine rootstock collection housed at the University of Milan, and to evaluate the goodness of this new plant material. SNP and SSR profiles have been compared for their usefulness to detect genetic diversity and population structure.

2. Material and Methods

2.1. Plant Material

Seventy *Vitis* ssp. genotypes belonging to a core collection identified in Migliaro et al. [19] were genotyped using 18k SNP. The pedigree of 31 genotypes is unknown, and the others are 31 hybrids genotypes and 8 traced to pure *Vitis* species (Table 1). The core collection is located in Torrazza Coste, Pavia, Italy (44.984783 N, 9.089038 E, 133 m a.s.l.).

Table 1. List of the vine rootstock core collection. The genotype name and the breeding material is reported for each accession [19].

Accession ID	Genotype	Pedigree
1	101.14 Millardet et de Grasset	*V. riparia* × *V. rupestris*
4	110 Richter	*V. rupestris* × *V. berlandieri*
6	1202 C	*V. vinifera* × *V. rupestris*
10	161.49 Couderc	*V. berlandieri* × *V. riparia*
16	3309 Couderc	*V. riparia* × *V. rupestris*
17	333 Ecole de Montpellier or Tisserand	*V. vinifera* × *V. berlandieri*
19	41 B Millardet et de Grasset	*V. vinifera* × *V. berlandieri*
28	Cosmo 10	*V. berlandieri* × *V. riparia*
29	Dog Ridge	*V. rupestris* × *V. candicans*
38	Isabella	*V. labrusca* × *V. vinifera*
39	Jacquez	*V. aestivalis* × *V. vinifera*
40	Geilweilerhof V.348	*V. vinifera*
41	Kober 5BB	*V. berlandieri* × *V. riparia*
42	LN 33 or LLYOD'S NUMBER 33	*V. riparia* × *V. longii* × *V. vinifera*
43	Malegue 44.53	*V. riparia* × *V. cordifolia* × *V. rupestris*
46	Salt Creek	Unknown
55	*Vitis riparia* Fabre	*V. riparia*
56	*Vitis riparia* Gloire de Montpellier	*V. riparia*
68	Genotype 01	Unknown
69	Genotype 02	*V. berlandieri* × *V. riparia* × *V. cinerea*
70	Genotype 03	*V. berlandieri* × *V. riparia* × *V. cordifolia* × *V. rupestris*
71	Genotype 04	Unknown
81	Genotype 15	*V. berlandieri* × *V. riparia* × *V. rupestris*
83	Genotype 17	*V. berlandieri* × *V. riparia*
84	Genotype 18	Unknown
96	Genotype 29	*V. riparia* × *V. vinifera* × *?*
99	Genotype 33	*V. riparia* × *V. longii* × *?*
112	Genotype 46	Unknown
114	Genotype 48	Unknown
116	Vitis riparia Lombard	*V. riparia*
118	Genotype 52	Unknown
120	Genotype 54	Unknown
121	Genotype 55	Unknown
125	Vitis labrusca Muncy	*V. labrusca*
126	Genotype 60	Unknown
127	Genotype 61	Unknown
129	Genotype 63	*V. berlandieri* × *V. riparia* × *?*
132	Genotype 66	Unknown
134	Genotype 68	Unknown
136	Genotype 70	*V. vinifera* × *V. berlandieri* × *V. riparia* × *V. candicans*
140	Genotype 74	*V. berlandieri* × *V. riparia*
150	Kober 125 AA	*V. berlandieri* × *V. riparia*
152	Genotype 86	Unknown
153	Genotype 87	Unknown
154	Genotype 88	Unknown
155	Genotype 89	Unknown
161	Genotype 95	Unknown
162	Genotype 96	Unknown
163	Genotype 97	Unknown
164	Genotype 98	Unknown
166	Genotype 100	Unknown
169	Genotype 103	Unknown
171	Genotype 105	Unknown
172	Genotype 106	*V. riparia* × *V. rupestris* × *?*

173	Genotype 107	Unknown
176	Genotype 110	Unknown
177	Genotype 111	Unknown
184	Genotype 118	Unknown
187	143 B Millardet et De Grasset	*V. vinifera* × ?
192	202-4 Millardet et De Grasset	*V. riparia* × *V. longii* × ?
198	33 Ecole de Montpellier	*V. berlandieri* × *V. riparia*
199	420 B Millardet et De Grasset	*V. berlandieri* × *V. riparia*
206	Dufour 11 F	*V. riparia* × *V. rupestris*
214	Genotype 148	Unknown
215	*Vitis riparia* Sericea	*V. riparia*
216	*Vitis riparia* Sombre	*V. riparia*
217	*Vitis riparia* Tomenteux	*V. riparia*
223	Genotype 157	*V. labrusca* × *V. riparia* × *V. rupestris*
224	Genotype 158	Unknown
231	Genotype 165	Unknown

2.2. DNA Extraction and SNP Genotyping

One hundred milligrams of freeze-fresh young leaf tissue were ground with liquid nitrogen, and genomic DNA was extracted using NucleoSpin® Plant II (MACHEREY-NAGEL—Düren, Germany), according to manufacturer's protocol. Concentration of DNA and its quality were checked by electrophoresis on agarose gel by spectroscopy (260/230 and 260/280 ratios) using NanoDrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and the Quant-iT dsDNA HS assay kit for Qubit 3.0 Fluorometer (Thermo Fisher Scientific). SNP genotyping was performed on 200 ng of genomic DNA per sample using the Vitis18kSNP array (Illumina Inc., San Diego, CA, USA), containing 18,071 SNPs, by the laboratory of Fondazione Edmund Much (San Michele all'Adige, Trento, Italy).

2.3. Data Analysis

For SNP data, samples with call quality value (p50GC) lower than 0.54 and loci with a GenTrain (GT) score value lower than 0.6 [35] were filtered from the dataset, as well as those with more than 20% of missing data and monomorphic loci. Number of alleles and their frequency, observed heterozygosity (H_o), expected heterozygosity (H_e), and minor allele frequency (MAF) were assessed using PEAS V1.0 software [36]. In order to identify the minimum number of SNP loci able to explain the observed diversity in our data set, the accumulation curve approach implemented in the package *poppr* [37], and AMaCAID [38] for R software [39] were used. The results were viewed as a barplot.

The genetic structure of the core collection was analyzed using the LEA package [40] of R software by varying the number of ancestral genetic groups (K) from 1 to 10 in ten repetition runs for each K value. The Principal Components Analysis (PCA) was run by using the *adegenet* package of R software [41], and the first two components values were plotted on a 2-D scatterplot. The genetic distance among genotypes was set up on Nei's distance [42], performed in PEAS, and the clustering was performed using Unweighted Pair Group Method with Arithmetic Mean (UPGMA). A circular dendrogram was plotted using MEGA 7.0 software [43]. The validation of clustering results was performed considering the pairwise Nei's genetic distance. The values were calculated using the nei.dist function of R software.

In order to compare SNP and SSR results, the Migliaro et al. [19] core collection SSR profiles were used to detect the number of alleles, H_o and H_e, using the GenAlEx 6.5 software [44]. A structure analysis was performed using STRUCTURE 2.0 software [45]. Burn-in and MCMC (Markov Chain Monte Carlo) values were set on 100,000 replicate runs, the number of clusters (K) varied from 1 to 10, and 10 replicate runs were carried out to quantify the variation of the likelihood for each K. The most likely K value was chosen according to Evanno et al.'s [46] method. PCA was performed using the *adegenet* package, and the UPGMA circular dendrogram was drawn using MEGA 7.0 software on

the Nei's distance matrix assessed by GenAlEx 6.5. Clustering results were validated by pairwise Nei's genetic distance (GenAlEx 6.5).

To investigate the genetic relationship between non-*vinifera* and *vinifera* germplasm, our dataset was merged with those reported in De Lorenzis et al. [23], Laucou et al. [24], and De Lorenzis et al. [35]. The final dataset resulted in 1044 genotypes. PCA and parentage analysis were performed on the new dataset. Parentage analysis was performed to account for first-degree (parent-offspring) relationships among core collection genotypes and *V. vinifera* genotypes. The analysis was carried out by PLINK 1.07 software [47], calculating the identity-by-descent (IBD). The following parameters were set: MAF = 0.1 and r^2 of linkage disequilibrium = 0.05. The parent-offspring (PO) relationships among genotypes were assigned based on Z0 (probability of sharing 0 IBD allele identical-by-descent), Z1 (probability to share 1 IBD allele), Z2 (probability to share 2 IBD alleles), and PI-HAT (the relatedness measure measured as PI-HAT = P (IBD = 2) + 0.5 × P (IBD = 1)) parameters. To assign the PO relationships, the experimental values were compared to the theoretical ones: Z0 and Z2 values similar to 0, Z1 similar to 1, and PI-HAT to 0.5. Only relationships with core collection genotypes will be discussed.

3. Results

Seventy *Vitis* ssp. genotypes of a grapevine rootstock core collection [19] were genotyped using the Vitis18kSNP array. Filtering the genetic profiles for a call quality value (p50GC) higher than 0.54, 66 out of 70 genotypes were retained, probably due to a low quality of DNA. Accessions 152, 192, 216, and 231 were not considered for further analysis. The number of SNP loci with a GT score value higher than 0.6 was 16,495 (91.3% of the total), and the loci showing a percentage of missing data lower than 20% amounted to 15,688 (86.8%). Finally, 1508 monomorphic SNPs were removed, obtaining a final dataset of 14,180 SNPs (78.5%) suitable for genetic characterization of the analyzed genotypes (Table S1). The final dataset accounted for 11,717 *vinifera* SNPs (around 86% of SNPs identified in the *V. vinifera* genome) and 2463 non-*vinifera* SNPs (around 55% of SNPs identified in the genome of other species). Among the non-*vinifera* SNPs, the *M. rotundifolia* SNPs showed the lowest percentage of loci successfully amplified (19%), while *V. berlandieri* SNPs showed the highest (66%). Two R packages (*poppr* and AMaCAID) were used to identify the minimum number of loci able to distinguish the 66 genotypes (100% of genetic diversity). The genotype accumulation curves reported in Figure 1 indicated that by randomly sampling 64 or 49 SNPs, respectively, based on the simulation performed with *poppr* package (Figure 1a) and AMaCAID package (Figure 1b), 100% core collection genetic diversity is detected.

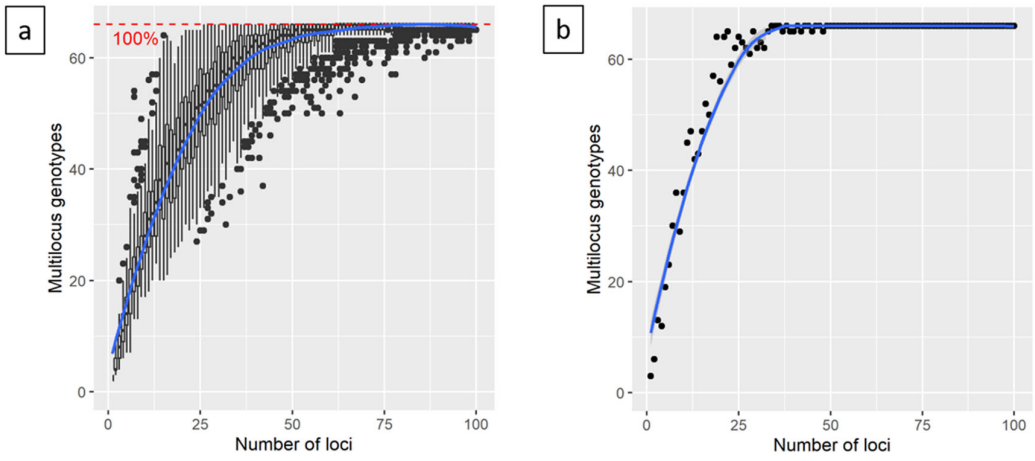

Figure 1. Genotype accumulation curve of 66 grapevine rootstock accessions, genotyped over 14,180 SNP loci, obtained with *poppr* package (**a**) and AMaCAID package (**b**). Value on Number of loci axis was limited to 100. The red dashed line represents 100% of the total observed genotypes.

Genetic diversity of core collection was evaluated using both single nucleotide polymorphism (SNP) and Simple Sequence Repeat (SSR) molecular markers. The average number of alleles for SNPs is 1.80, and the minor allele frequencies (MAF) is equal to 0.10. The percentage of SNPs reporting MAF higher than 0.05 was about 57%. For the largest part of SNPs (10,162), no difference ($p \leq 0.05$) was found between H_o and H_e values. In the other loci, H_o was lower than H_e in 3584 SNPs and higher in the remaining 434 SNPs. Both molecular markers showed similar H_o and H_e values (0.143 vs. 0.157 and 0.823 vs. 0.879, respectively, for SNP and SSR loci), as reported in Table 2.

Table 2. Observed (H_o) and expected (H_e) heterozygosity of the vine rootstock core collection and the ancestral groups identified by the structure analysis based on SNP and SSR profiles. N = number of genotypes.

Marker	Plant Material	N	H_o	H_e
SNP	Core collection	66	0.143	0.157
	Group 1	8	0.100	0.095
	Group 2	16	0.303	0.280
	Group 3	15	0.125	0.099
	Group 4	27	0.071	0.067
SSR	Core collection	70	0.823	0.879
	Group 1	10	0.882	0.730
	Group 2	24	0.779	0.771
	Group 3	36	0.836	0.891

In order to identify the ancestral population in the analyzed core collection, a structure analysis was performed with both SNP and SSR profiles. The analyses estimated the most likely number of ancestral populations at K = 4 for SNPs and K = 3 for SSRs. A bar plot representation of the two structures is shown in Figure 2. Based on SNP profiles, the percentage of admixed genotypes (reporting the predominant K values lower than 0.8) was about 53% (Figure 2a; Table S2). SNP-group 1 was the smallest group, where only 12% of genotypes were included, whereas SNP-group 4 was the biggest (41%). *V. berlandieri* × *V. riparia* genotypes and those having an unknown pedigree were grouped in all the four SNP-groups. The majority of unknown genotypes belonged to SNP-group 4 (11 out of 29). In the SNP-group 1, only three known genotypes were clustered, a *V. berlandieri* × *V. riparia* (ID 83), a *V. vinifera* × *V. rupestris* (ID 6), and a *V. labrusca* (ID 125). In the SNP-group 2, we had genotypes with other species in their genetic background (such as *V. cordifolia, V. labrusca, V. rupestris,* and *V. vinifera*), as well as in the SNP-group 3, where genotypes with *V. candicans, V. labrusca, V. longii,* and *V. rupestris* in their pedigree were clustered. Pure *V. riparia* genotypes were assigned to the SNP-group 4 together with *V. candicans, V. cinerea, V. longii, V. rupestris,* and *V. vinifera* genotypes. Based on SSR profiles, 31% of genotypes were admixed, and 69% of genotypes were grouped in three ancestral groups (10% of genotypes in SSR-group 1, 23% in SSR-group 2, and 36% in SSR-group 3) (Figure 2b; Table S2). The *V. berlandieri* × *V. riparia* genotypes were mainly grouped in ancestral SSR-group 1, whereas *V. rupestris* genotypes were assigned to SSR-group 2 and *V. vinifera* to SSR-group 3.

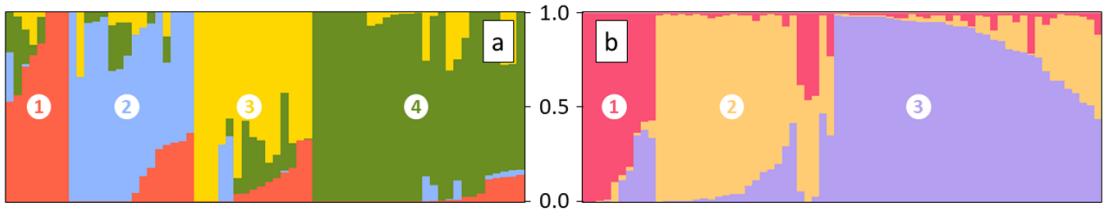

Figure 2. Genetic structure of the grapevine rootstock core collection (66 genotypes), defined using 14k SNP (**a**) and 17 SSR (**b**) molecular markers. Ancestral groups (1, 2, 3, and 4 for SNP and 1, 2, and 3 for SSR) are reported in different colors.

PCA was performed to identify correlations among structure groups (Figure 3). Regarding SNPs, the first two principal components (PCs) explained 37% of the total variability. PCA was able to discriminate among the ancestral groups identified by the structure analysis. PC1 separated SNP-group 2 from the others, whereas PC2 highlighted the differences among SNP-group 1, 3, and 4. SNP-group 2 was the group showing the highest diversity, with four out of five not-admixed genotypes (ID 39, 40, 71, and 120) clustered aside from the other genotypes (Figure 3a). Performing PCA on SSR profiles, the first two PCs described 49% of the total variability. As for SNPs, ancestral groups were discriminated by PCA. SSR-groups 2 and 3 were separated along the PC1 and SSR-group 1 along the PC2 (Figure 3b). Admixed genotypes were generally placed in between the genotypes of each ancestral group, independently from the used molecular markers.

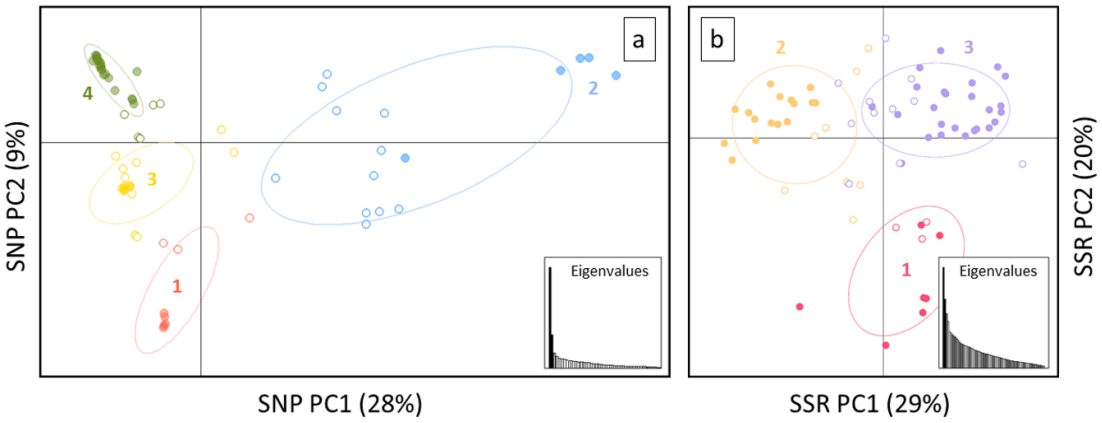

Figure 3. Principal component analysis (PCA) of the grapevine rootstock core collection (66 genotypes), defined using 14k SNP (**a**) and 17 SSR (**b**) molecular markers. Genotypes are classified according to the ancestral groups identified in the structure analysis. White filled dots are admixed genotypes. Colors are according to ancestral groups reported in Figure 2.

Genetic distance among genotypes of the grapevine rootstock core collection is reported in Figure 4. Based on the SNP analysis, the genotypes showed different levels of similarity, ranging from 85% to 98%. Using the threshold value of 95% for the similarity index, two main groups were identified, one grouping genotypes belonging to the structure SNP-groups 1, 3, and 4 and the other genotypes of SNP-group 2. In each cluster, genotypes were clustered according to the ancestral group they belong to. Similarly to the PCA analysis, genotypes of the SNP-group 2 were the most different, with samples ID 39, 40, 71, and 120 clustering as outgroups. SSR dendrogram showed similarity values ranging from 75% to 95%. Two main clusters were identified (threshold value = 83%), one grouping mainly genotypes of the ancestral SSR-groups 1 and 2, and the other, the genotypes of SSR-group 3. Each genotype was clustered according to their ancestral group. Genotypes of SSR-group 3 were the most different (Figure 4b).

H_o and H_e values per each ancestral group were similar to each other for both SNP and SSR groups, with H_o, generally, higher than H_e (Table 2). About SNPs, SNP-group 2 and SNP-group 4 were the groups showing the highest (0.303 vs. 0.280) and lowest (0.071 vs. 0.067) H_o and H_e values, respectively. The highest pairwise Nei's genetic distance was detected for the combination SNP-group 2–SNP-group 3 (0.197) and the lowest for the combination SNP-group 3–SNP-group 4 (0.037). About SSRs, H_o values ranged from 0.779 (SSR-group 2) to 0.882 (SSR-group 1) and H_e values from 0.730 (SSR-group 1) to 0.891 (SSR-group 3). Pairwise Nei's genetic distance showed the highest value for the combination SSR-group 2–SSR-group 3 (0.674), and the lowest for combination SSR-group 1–SSR-group 2 (0.484).

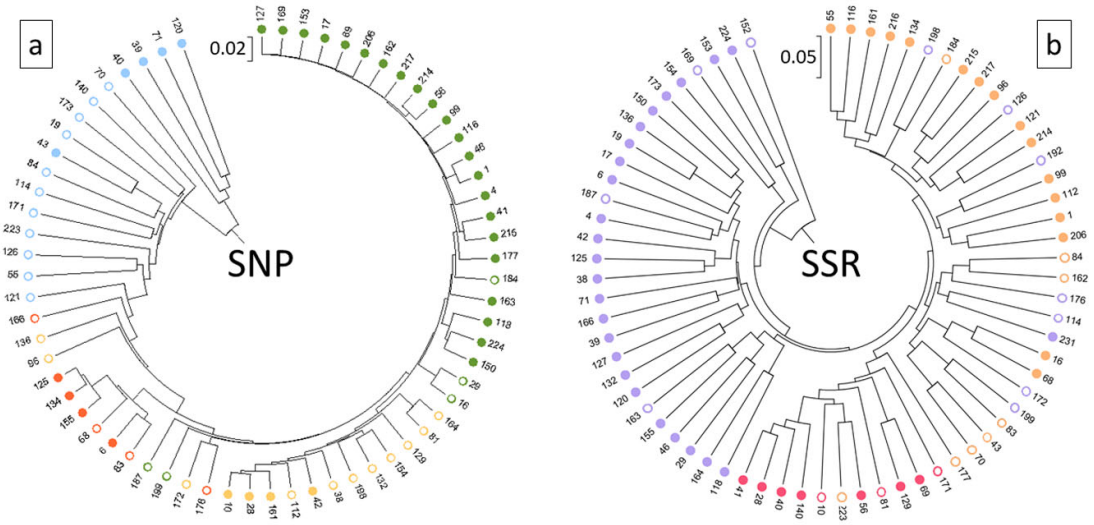

Figure 4. Clustering of the grapevine rootstock core collection (66 genotypes), according to Nei's distance, defined using 14k SNP (**a**) and 17 SSR (**b**) molecular markers. White filled dots are admixed genotypes. Colors are according to ancestral groups reported in Figure 2.

In order to highlight the genetic relationship between non-*vinifera* and *vinifera* germplasm [23,24,35], a second dataset was built, accounting for 1044 genotypes and 6375 SNPs. PCA results were plotted on a scatter plot (Figure 5). The first two principal components (PCs) explained 23% of total genetic variability (19% and 4% for PC1 and PC2, respectively). The genotypes were discriminated along the PC1 in two well distinct groups: (i) vinifera group; (ii) non-vinifera group. Moreover, some genotypes belonging to the non-*vinifera* dataset overlapped with the *vinifera* genotypes (ID 39, 40, 71, and 120, two genotypes with a *vinifera* background and two genotypes with an unknown pedigree), and some other were strongly differentiated from the rest of non-*vinifera* genotypes. The latter group of genotypes were mainly genotypes with a *riparia* background. Some *vinifera*-backgrounded genotypes and some unknown genotypes were included as well.

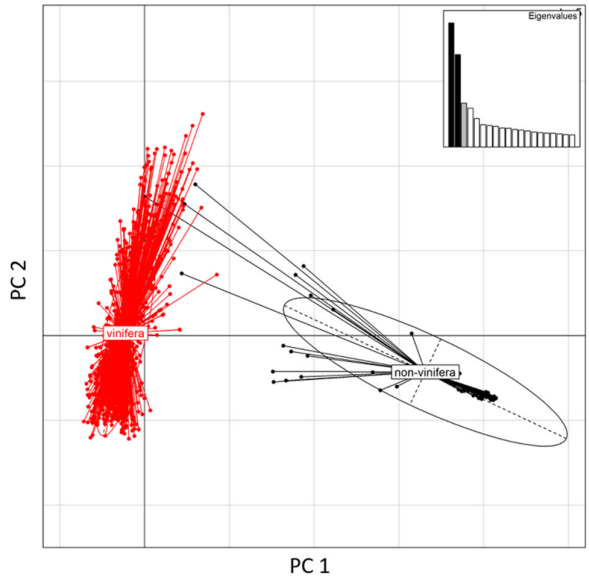

Figure 5. Principal component analysis (PCA) of non-*vinifera* (66 genotypes) and *vinifera* genotypes (978 genotypes [23,24,35]), defined using 6k SNP molecular markers.

On the same dataset, parentage analysis was performed to account for first-degree relationships among core collection genotypes and *V. vinifera* germplasm [23,24,35]. Only one PO relationship was observed in the new dataset: ID 40 (Geilweilerhof V.348) = Pinot noir × Riesling. The experimental values for relationship parameters were as follows: (i) ID 150—Pinot noir, Z0 = 0.012, Z1 = 0.923, Z2 = 0.092, PI_HAT = 0.553; ii) ID 150—Riesling, Z0 = 0.035, Z1 = 0.898, Z2 = 0.058, PI_HAT = 0.507.

4. Discussion

*4.1. The 18k SNP Genotyping Array is A Suitable Tool to Characterize Non-*vinifera *Germplasm*

Increasing efforts in new rootstock selection require effective tools able to investigate the diversity in the genus *Vitis*. Recently, an 18k SNP genotyping array has been developed, containing 13,561 SNPs isolated from *V. vinifera* and 4510 SNPs from other *Vitis* species [24]. So far, the Vitis18kSNP array was used in several studies on *V. vinifera* germplasm characterization, but its effectiveness on grapevine rootstocks has not been tested yet. In this work, the array was validated on 70 genotypes of a grapevine rootstock core collection, obtaining a final dataset of 14,180 SNP loci. This number of SNP loci was in line with the ones reported for *V. vinifera* germplasm, ranging from 10,041 to 16,501 SNPs [23,24,26,31,33–35], resulting in an informative tool for grapevine rootstock genetic characterization. Among the core collection genotypes with a known pedigree [19], the species mostly represented are *V. riparia* (29 genotypes), *V. berlandieri* (16), *V. rupestris* (11), and *V. vinifera* (10). Although, *V. riparia* and *V. ruprestris*, two species worldwide used in the breeding programs of grapevine rootstocks for their resistance trait to phylloxera [7], were not included in the panel of species used to identify and select the SNPs [24], genotypes having in their pedigree the genetic background of these two species were successfully analyzed. On the other hand, about 56% and 59% of SNPs identified in the genome of *V. aestivalis* and *V. cinerea*, respectively, were amplified, even though in the core collection, these two species appeared less represented (only two genotypes among the ones with a known pedigree). Nevertheless, it is not possible to exclude that among the unknown genotypes there are some individuals with an *V. aestivalis* and *V. cinerea* background. These results confirm that molecular markers identified in *V. vinifera* are appropriate for different genotype *Vitis* species and vice versa, as already verified for other molecular markers, such as SSR [48], REMAP [49], and iPBS [50]. Only the SNPs loci detected in the genome of *M. rotundifolia* were not useful for *Vitis* non-*vinifera* genotyping (only 19% of *M. rotundifolia* SNPs were successfully amplified). *Muscadinia* (2n = 40) and *Vitis* (2n = 38) are the two subgenera of the *Vitis* genus. The two subgenera are distinguishable based on morphological traits [51] and are nearly reproductively isolated, exhibiting significant divergence each other [52,53].

To genotype the *V. vinifera* germplasm, a set of nine SSRs has been established as reference tools to distinguish among the grapevine cultivars [54]. Seven out of these nine SSR loci were found to be suitable to distinguish among the non-*vinifera* genotypes [19]. Regarding the Visit18kSNP genotyping array, Mercati et al. [33] suggested a minimal set of 12 SNP loci to discriminate among Sicilian cultivars, and Laucou et al. [24] found 14 as the minimal number of SNP loci to distinguish among 783 grapevine cultivars. In this work, a minimum number of SNP loci has been proposed for the non-*vinifera* germplasm as well, using two different R packages. Both packages detected a number of loci (64 and 49 SNPs; Figure 1) higher than the one detected by Laucou et al. [24] and Mercati et al. [33]. Because the minimal set of loci can change depending on the genetic diversity of genotypes analyzed, the larger size of minimal SNP set detected for non-*vinifera* germplasm reflects the low genetic distance observed by SNPs in comparison to the SSR one (Figure 4).

4.2. SNP Profiles Reveal A High Level of Admixture

Genetic characterization of grapevine rootstocks can be performed by different marker types, in which the results do not always overlap [10]. In this study, SNP and SSR profiles were compared to assess the genetic diversity of the grapevine rootstock core collection. Differences between SNPs and SSRs were observed with respect to heterozygosity (Table 2). As expected, due to their multiallelic

nature and high level of polymorphism, SSR loci exhibited a significantly higher heterozygosity than bi-allelic SNP loci. The same trend was observed by Emanuelli et al. [10] comparing a set of 384 SNPs to 22 SSRs on 122 rootstock genotypes. In particular, they observed rootstock heterozygosity values of H_o = 0.099 and H_o = 0.734 for SNPs and SSRs, respectively, slightly lower than the heterozygosity observed in this work (Table 2), suggesting that 384 SNPs related to phenotypical traits have the same power than 14k SNPs unrelated to phenotypical traits to detect the heterozygosity. SSR H_e value of core collection (Table 2) was larger than the values detected in other studies about different rootstock material [10,15,17,18], confirming the uniqueness and preciousness of the analyzed germplasm collection. The H_o value of core collection was lower than the H_e for both molecular markers. This result can be addressed to a Wahlund effect due to population substructure [55]. Indeed, although the high percentage of admixed genotypes (Figure 2), structure groups were detected with both molecular markers. The average minor allele frequency among the 14,180 SNPs (MAF = 0.10) was slightly higher than rootstock germplasm studied by Emanuelli et al. [10] (MAF = 0.08), but lower than the *sativa* compartment (MAF = 0.26).

A different genetic structure was defined according to the marker type: using SSRs, three ancestral groups were identified (Figure 2b), whereas SNPs defined a more complex structure, consisting of four ancestral populations (Figure 2a). The same trend was reported by Laucou et al. [24] on *V. vinifera* cultivars genotyped with the same set of SNP and SSRs. A different result was described by Emanuelli et al. [10], where K = 6 and K = 5 were identified for SSRs and SNPs, respectively, probably due to the lower number of SNP loci used to genotype the individuals. As a result of the higher number of SNP ancestral groups, the percentage of admixed genotypes was lower for SSRs (most of them also resulted admixed for SNP analysis). According to Klein et al. [56], two main clades can be discerned among North American *Vitis* species: clade I comprised *V. riparia* and *V. rupestris* together with *Vitis acerifolia, Vitis arizonica,* and *Vitis monticola*; clade II consisted of *Vitis aestivalis, V. cinerea, V. labrusca,* and *Vitis mustangensis*. SNP and SSR profiles were not able to capture this division between *V. riparia* and *V. rupestris* and the other species, such as *V. labrusca*, probably due to the low number of genotypes having one species in their pedigree. The core collection was designed to maximize the genetic variation of our *non*-vinifera germplasm collection [19], and a high number of genotypes having a genetic background derived from three or four species have been included. The complex pedigree of selected genotypes supports the high level of admixture with any strong evidence of differentiation among species.

PCA (Figure 3) and cluster analysis (Figure 4) produced consistent results, which clearly discriminated the structure of ancestral groups for both markers types. Nei's genetic distances reflected structure, PCA, and cluster distribution, confirming the SNP-group 2 and SSR-group 3 as the most different, although SSR Nei's genetic distance values among groups were higher than SNP ones. Both groups clustered individuals having species different from *V. berlanderi*, *V. riparia,* and *V. rupestris* (the three species mostly used in the rootstock breeding programs [7]) in their genetic background, such as *V. aestivalis*, *V. candicans*, *V. cordifolia,* and *V. longii*. Some of these genotypes were clustered as much different in comparison to the individuals belonging to the same group (such as ID 39, 40, 71, and 120 for SNP analysis). Based on the comparison between non-*vinifera* and *vinifera* germplasm (Figure 5), these genotypes were assigned to the *vinifera* germplasm. Although *V. vinifera* was not used to breed rootstock material so far, due to their susceptibility of phylloxera [1], rootstock (non-*vinifera*) and scion (*vinifera*) do not always make up a successful graft. Indeed, the higher the inter-specificity between rootstock and scion, the higher the incompatibility. For this reason, it could be interesting to investigate these genotypes by the phenotypical point of view for further breeding programs.

In contrast to the trend of the whole core collection, H_o values within the SNP and SSR structure groups were slightly higher than the expected ones (except for the SSR-group 3) (Table 2). This result is due to the absence of clear discrimination among species based on the structure analysis (Figure 2) and low genetic variation (Figure 4) due to inbreeding among species [7]. The difficulty in finding a clear differentiation among genotypes with different genetic backgrounds can be traced back to the classification of the *Vitis* genus. Indeed, the *Vitis* species are interfertile with most of their distribution

areas overlapping, where natural hybridization can occur. This hybridization can mix the morphological traits and make the identification of a true species difficult [53].

*4.3. SNPs Performed Well in Discriminating Non-*vinifera *and* vinifera *Germplasm*

The Vitis18kSNP genotyping array was mainly developed to analyze *V. vinifera* germplasm [25], but in this work, it was also demonstrated as working well with non-*vinifera* germplasm, amplifying a high number of loci and discriminating well among non-*vinifera* and *vinifera* germplasm (Figure 5). It was already demonstrated that the two germplasms are clearly differentiated when analyzed with both SSR and SNP molecular markers [10,11]. The genotyping array strongly discriminated the two germplasms, even though some (four) core collection genotypes overlapped with the *vinifera*-genotypes. These four genotypes are two (ID 39 and 40) *vinifera*-backgrounded genotypes (ID 40 has a *V. vinifera* × *V. vinifera* pedigree) and two (ID 71 and 120) unknown genotypes, suggesting a likely *vinifera* background also for the latter genotypes. In the non-*vinifera* group, part of *riparia* genotypes were placed in between non-*vinifera* and *vinifera* genotypes, appearing as the less homogeneous genotypes. Because together with *riparia* genotypes, some *vinifera*-backgrounded and unknown genotypes were also placed, it can be suggested a "*riparia* × *vinifera*" background for those genotypes with an unknown pedigree. The strong differentiation among non-*vinifera* and *vinifera* germplasm was also confirmed by the lack of PO relationships between the two groups of genotypes.

4.4. From SSR to SNP Genotyping

SNPs are widely used to genotype crops [57–61] and are markers of choice for QTL and GWAS [24,29,30,62–64] due to their number, distribution, and density along the genome. In *V. vinifera*, the genotyping SNP array has been used successfully to investigate the genetic diversity of grapevine, to discriminate among the wild and cultivated compartments, to infer population structure, and to reconstruct the pedigree of cultivars [23,24,26,31,33–35]. Its attractiveness is due to a number of advantages, such as their high reproducibility among the laboratories, transferability, throughput, automatization, and inexpensiveness. Nevertheless, the success of this tool will be established once laboratories fully adopt SNPs as the genotyping method, instead of SSRs, and the number of individuals analyzed with SNP array raises, as well as the reference databases. If this shift appears difficult to be applied due to the great efforts made in genotyping the *vinifera* germplasm using a universal panel of nine SSR loci [54], for the non-*vinifera* germplasm, the genotyping is still at the beginning, making this shift a more feasible change.

5. Conclusions

The genetic base of available *Vitis* rootstocks derived from a restricted number of genotypes, selected among North American *Vitis* species at the end of the XIX century. Considering the relevant role of rootstocks on environmental stress tolerance, the low genetic diversity reduces the ability of grapevine cultivars to adapt to several environmental constraints. This issue can be faced by increasing the genetic and phenotypic diversity of the breeding material, including non-conventional material in the further breeding programs. Living germplasm collections are valuable resources for exploring the genetic and phenotypic diversity and providing new genetic resources to support plant breeding efforts. The non-*vinifera* collection housed at the University of Milan has been established with the purpose of collecting the diversity of non-*vinifera* germplasm as much as possible and to design a core collection where the putatively novel breeding material are included. Because the SNP genotyping is becoming even more popular for a number of advantages (rapid processing of large populations and data harmonization), 70 individuals of the non-*vinifera* germplasm core collection have been genotyped by *Vitis* SNP genotyping array. The SNP genotyping array has proved adequate to study the genetic diversity of non-*vinifera* germplasm. The genetic characterization provided the uniqueness and preciousness of the core collection as a source of plant breeding material not commonly used so far. The 18k SNP genotyping array will be a valid tool to assist the selection of the most promising individuals.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Table S1: SNP profiles of 66 non-*vinifera* rootstock accessions, genotyped at 18 K loci. Original dataset was filtered based on SNP call quality (p50GC < 0.54), GenTrain score (> 0.6), and monomorphic loci, resulting in 14,180 SNP loci. "A": homozygous for dominant allele; "B": homozygous for recessive allele; "H": heterozygous; "U": missing data. Table S2: Ancestry values at K = 4 and K = 3 for SNP and SSR profiles, respectively, of 66 non-*vinifera* rootstock accessions.

Author Contributions: Conceptualization, L.B. and G.D.L.; methodology, D.B. and G.D.L.; writing, D.B., G.D.L., and L.B. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by University of Milan.

Acknowledgments: This work is part of the European project titled "FREECLIMB—Fruit Crops Adaptation To Climate Change In The Mediterranean Basin" in the frame of the Programme Partnership For Research And Innovation In The Mediterranean Area (PRIMA; call 2018). The authors thank Samuele Bongiolatti for DNA extraction. The authors thank Asus for providing hardware.

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

References

1. Granett, J.; Walker, M.A.; Kocsis, L.; Omer, A.D. Biology and management of grape phylloxera. *Annu. Rev. Entomol.* **2001**, *46*, 387–412.
2. Brancadoro, L.; Rabotti, G.; Scienza, A.; Zocchi, G. Mechanisms of Fe-efficiency in roots of *Vitis* spp. in response to iron deficiency stress. *Plant Soil* **1995**, *171*, 229–234.
3. Ollat, N.; Peccoux, A.; Papura, D.; Esmenjaud, D.; Marguerit, E.; Tandonnet, J.P.; Bordenave, L.; Cookson, S.J.; Barrieu, F.; Rossdeutsch, L.; et al. Rootstocks as a component of adaptation to environment. In *Grapevine in a Changing Environment: A Molecular and Ecophysiological Perspective,* 1st ed.; Gerós, H., Chaves, M.A., Gil, H.M., Delrot, S., Eds.; John Wiley & Sons, Ltd.: Chichester, UK, 2016; pp. 68–108.
4. Köse, B.; Horuz, A.; Akınoğlu, G. Phenological changes of leaf nutrients in *Vitis labrusca* L. grape grafted on different rootstocks in heavy clay soil conditions. *Erwerbs-Obstbau* **2016,** *58*, 211–217.
5. Corso, M.; Bonghi, C. Grapevine rootstock effects on abiotic stress tolerance. *Plant Sci.* **2014**, *1*, 108–113.
6. Keller, M. *The Science of Grapevines: Anatomy and Physiology*, 2nd ed; Academic Press: Cambridge, MA, USA, 2015; pp. 1–552.
7. Riaz, S.; Pap, D.; Uretsky, J.; Laucou, V.; Boursiquot, J.M.; Kocsis, L.; Walker, A. Genetic diversity and parentage analysis of grape rootstocks. *Theor. Appl. Genet.* **2019**, *132*, 1847–1860.
8. Bacilieri, R.; Lacombe, T.; Cunff, L.; Le, Vecchi-staraz, M.; Di, Laucou, V.; Genna, B.; Péros, J.; This, P.; Boursiquot, J. Genetic structure in cultivated grapevines is linked to geography and human selection. *BMC Plant Biol.* **2013**, *13*, 25.
9. Cipriani, G.; Spadotto, A.; Jurman, I.; Di, 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 synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. *Theor. Appl. Genet.* **2010**, *121*, 1569–1585.
10. 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.
11. 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.
12. Crespan, M.; Meneghetti, S.; Cancellier, S. Identification and genetic relationship of the principal rootstocks cultivated in Italy. *Am. J. Enol. Vitic.* **2009**, *60*, 349–356.
13. Cseh, A.; Taller, J.; Podmaniczky, P.; Kocsis, L. Comparative analysis of the most widespread grapevine rootstock lines in the world, the teleki lines, with microsatellite markers. *Cereal Res. Commun.* **2006**, *34*, 773–776.
14. De Andrés, M.T.; Cabezas, J.A.; Cervera, M.T.; Borrego, J.; Martínez-Zapater, J.M.; Jouve, N. Molecular characterization of grapevine rootstocks maintained in germplasm collections. *Am. J. Enol. Vitic.* **2007**, *58*, 75–86.

15. Dzhambazova, T.; Hvarleva, T.; Tsvetkov, I.Y.; Atanassov, A. Characterization of grapevine rootstocks using microsatellite markers. *Biotechnol. Biotech. Eq.* **2007**, *21*, 58–62.
16. Jahnke, G.; Májer, J.; Szoke, B.; Varga, P.; Kocsisné Molnár, G.; Tarczal, E.; Kocsis, L. Analysis of grape rootstocks by microsatellite markers. *Acta Hortic.* **2014**, *1046*, 617–626.
17. Sefc, K.M.; Regner, F.; Glössl, J.; Steinkellner, H. Genotyping of grapevine and rootstock cultivars using microsatellite markers. *Vitis* **1998**, *37*, 15–20.
18. Upadhyay, A.; Saboji, M.D.; Reddy, S.; Deokar, K.; Karibasappa, G.S. AFLP and SSR marker analysis of grape rootstocks in Indian grape germplasm. *Sci. Hortic.* **2007**, *112*, 176–183.
19. Migliaro, D.; De Lorenzis, G.; Di Lorenzo, G.S.; De Nardi, B.; Gardiman, M.; Failla, O.; Brancadoro, L.; Crespan, M. Grapevine non-*vinifera* genetic diversity assessed by SSR markers as a starting-point for new rootstock breeding programs. *Am. J. Enol. Vitic.* **2019**, *70*, 4.
20. Lin, H.; Walker, M.A. Identifying grape rootstocks with simple sequence repeat (SSR) DNA markers. *Am. J. Enol. Vitic.* **1998**, *49*, 403–407.
21. 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.
22. 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.
23. 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.
24. Laucou, V.; Launay, A.; Bacilieri, R.; Lacombe, T.; Adam-Blondon, A.F.; Bérard, A.; Chauveau, A.; De Andrés, 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*, 1–27.
25. 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. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3530–3535.
26. 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*, 50.
27. Myles, S.; Chia, J.; Hurwitz, B.; Simon, C.; Zhong, G.Y.; Ware, D. Rapid genomic characterization of the genus *Vitis*. *PLoS ONE* **2010**, *5*, e8219.
28. Marrano, A.; Birolo, G.; Prazzoli, M.L.; Lorenzi, S.; Valle, G.; Grando, M.S. SNP-Discovery by RAD-sequencing in a germplasm collection of wild and cultivated grapevines (*V. vinifera* L.). *PLoS ONE* **2017**, *12*, e0170655.
29. Guo, D.-L.; Zhao, H.-L.; Li, Q.; Zhang, G.-H.; Jiang, J.-F.; Liu, C.-H.; Yu, Y.-H. Genome-wide association study of berry-related traits in grape *Vitis vinifera* L. based on genotyping-by-sequencing markers. *Hortic. Res*. **2019**, *6*, 11.
30. Zhang, H.; Fan, X.; Zhang, Y.; Jiang, J.; Liu, C. Identification of favorable SNP alleles and candidate genes for seedlessness in *Vitis vinifera* L. using genome-wide association mapping. *Euphytica* **2017**, *213*, 136.
31. De Lorenzis, G.; Squadrito, M.; Rossoni, M.; di Lorenzo, G.S.; Brancadoro, L.; Scienza, A. Study of intra-varietal diversity in biotypes of Aglianico and Muscat of Alexandria (*Vitis vinifera* L.) cultivars. *Aust. J. Grape Wine Res.* **2017,** *23*, 132–142.
32. Degu, A.; Morcia, C.; Tumino, G.; Hochberg, U.; Toubiana, D.; Mattivi, F.; Schneider, A.; Bosca, P.; Cattivelli, L.; Terzi, V.; et al. Metabolite profiling elucidates communalities and differences in the polyphenol biosynthetic pathways of red and white Muscat genotypes. *Plant Physiol. Biochem.* **2015**, *86*, 24–33.
33. Mercati, F.; De Lorenzis, G.; Brancadoro, L.; Lupini, A.; Abenavoli, M.R.; Barbagallo, M.G.; Di Lorenzo, R.; Scienza, A.; Sunseri, F. High-throughput 18K SNP array to assess genetic variability of the main grapevine cultivars from Sicily. *Tree Genet. Genomes.* **2016**, *12*, 59.
34. Sunseri, F.; Lupini, A.; Mauceri, A.; De Lorenzis, G.; Araniti, F.; Brancadoro, L.; Dattola, A.; Gullo, G.; Zappia, R.; Mercati, F. Single nucleotide polymorphism profiles reveal an admixture genetic structure of

grapevine germplasm from Calabria, Italy, uncovering its key role for the diversification of cultivars in the Mediterranean Basin. *Aust. J. Grape Wine Res.* **2018**, *24*, 345–359.
35. 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.
36. Xu, S.; Gupta, S.; Jin, L. PEAS V1.0: A package for elementary analysis of SNP data. *Mol. Ecol. Resour.* **2010**, *10*, 1085–1088.
37. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/ or sexual reproduction. *PeerJ* ***2014***, *2*, e281.
38. Caroli, S.; Santoni, S.; Ronfort, J. AMaCAID: A useful tool for accurate marker choice for accession identification and discrimination. *Mol. Ecol. Resour.* **2011**, *11*, 733–738.
39. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online: https://www.eea.europa.eu/data-and-maps/indicators/oxygen-consuming-substances-in-rivers/r-development-core-team-2006 (accessed on 15 March 2006)
40. Frichot, E.; Franc, O. LEA: An R package for landscape and ecological association studies. *Methods Ecol. Evol.* **2015**, *6*, 925–929.
41. Jombart, T. *adegenet*: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405.
42. Nei, M. Genetic distance between populations. *Am. Nat.* **1972**, *106*, 283–292.
43. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874.
44. 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.
45. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959.
46. 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.
47. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575.
48. Sefc, K.; Regner, F.; Turetschek, E.; Glössl, J.; Steinkeller, H. Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different Vitis species. *Genome* **1999**, *42*, 367–373.
49. D'Onofrio, C.; De Lorenzis, G.; Giordani, T.; Natali, L.; Cavallini, A. Retrotransposon-based molecular markers for grapevine species and cultivars identification. *Tree Genet. Genomes* **2010**, *6*, 451–466.
50. Guo, D.; Guo, M.; Hou, X.; Zhang, G. Molecular diversity analysis of grape varieties based on iPBS markers. *Biochem. Syst. Ecol.* **2014**, *52*, 27–32.
51. Mullins, M.G.; Bouquet, A.; Williams, L.E. *Biology of the Grapevine*, 1st ed; Cambridge University Press: Cambridge, England, 1992; pp. 80–146.
52. Aradhya, M.; Wang, Y.; Walker, M.A.; Prins, B.H.; Koehmstedt, A.M.; Velasco, D.; Gerrath, J.M.; Dangle, G.S.; Preece, J.E. Genetic diversity, structure, and patterns of differentiation in the genus *Vitis*. *Plant Syst. Evol.* **2013**, *299*, 317–330.
53. Wan, Y.; Schwaninger, H.R.; Baldo, A.M.; Labate, J.A.; Zhong, G.Y.; Simon, C.J. A phylogenetic analysis of the grape genus (*Vitis* L.) reveals broad reticulation and concurrent diversification during neogene and quaternary climate change. *BMC Evol. Biol.* **2013**, *13*, 141.
54. Maul, E.; Sudharma, K.N.; Kecke, S.; Marx, G.; Müller, C.; Audeguin, L.; Boselli, M.; Boursiquot, J.M.; Bucchetti, B.; Cabello, R.; et al. The European *Vitis* Database (www.eu-vitis.de): A technical innovation through an online uploading and interactive modification system. *VITIS* **2012**, *51*, 79–85.
55. Wahlund, S. Zusammensetzung von populationen und korrelationserscheinungen von standpunkt der vererbungslehre aus betrachtet. *Hereditas* **1928**, *11*, 65–106.
56. Klein, L.L.; Miller, A.J.; Ciotir, C.; Hyma, K.; Uribe-Convers, S.; Londo, J. High-throughput sequencing data clarify evolutionary relationships among North American *Vitis* species and improve identification in USDA *Vitis* germplasm collections. *Am. J. Bot.* **2018**, *105*, 215–226.

57. Chagné, D.; Crowhurst, R.N.; Troggio, M.; Davey, M.W.; Gilmore, B.; Lawley, C.; Vanderzande, S.; Hellens, R.P.; Kumar, S.; Cestaro, A.; et al. Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS ONE* **2012**, *7*, e31745.
58. Felcher, K.J.; Coombs, J.J.; Massa, A.N.; Hansey, C.N.; Hamilton, J.P.; Veilleux, R.E.; Buell, C.R.; Douches, D.S. Integration of two diploid potato linkage maps with the potato genome sequence. *PLoS ONE* **2012**, *7*, e36347.
59. Ganal, M.W.; Durstewitz, G.; Polley, A.; Berard, A.; Buckler, E.S.; Charcosset, A.; Clarke, J.D.; Graner, E.M.; Hansen, M.; Joets, J.; et al. A Large Maize (*Zea mays* L.) SNP genotyping array: Development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS ONE* **2011**, *6*, e28334.
60. Sim, S.-C.; Durstewitz, G.; Plieske, J.; Wieseke, R.; Ganal, M.W.; Van Deynze, A.; Hamilton, J.P.; Buell, C.R.; Causse, M.; Wijeratne, S.; et al. Development of a large SNP genotyping array and generation of high density genetic maps in tomato. *PLoS ONE* **2012**, *7*, e40563.
61. Verde, I.; Bassil, N.; Scalabrin, S.; Gilmore, B.; Lawley, C.T.; Gasic, K.; Micheletti, D.; Rosyara, U.R.; Cattonaro, F.; Vendramin, E.; et al.. Development and evaluation of a 9K SNP Array for peach by internationally coordinated SNP detection and validation in breeding germplasm. *PLoS ONE* **2012**, *7*, e35668.
62. Bérard, A.; Le Paslier, M.C.; Dardevet, M.; Exbrayat-Vinson, F.; Bonnin, I.; Cenci, A.; Haudry, A.; Brunel, D.; Ravel, C. High-throughput single nucleotide polymorphism genotyping in wheat (*Triticum* spp.). *Plant Biotechnol. J.* **2009**, *7*, 364–374.
63. Ha, B.K.; Hussey, R.S.; Boerma, H.R. Development of SNP assays for marker-assisted selection of two southern root-knot nematode resistance QTL in soybean. *Crop Sci.* **2007**, *47*, 73–82.
64. Tian, F.; Bradbury, P.J.; Brown, P.J.; Hung, H.; Sun, Q.; Flint-Garcia, S.; Rocheford, T.R.; McMullen, M.D.; Holland, J.B.; Buckler, E.S. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* **2011**, *43*, 159–162.