

Remiern



# The SSR Null Allele Problem, and Its Consequences in Pedigree Reconstruction and Population Genetic Studies in Viticulture

Gizella Jahnke <sup>1,\*</sup>, József Smidla <sup>2</sup>, Tamás Deák <sup>3,\*</sup>, Róbert Oláh <sup>4</sup>, Barna Árpád Szőke <sup>1</sup> and Diána Ágnes Nyitrainé Sárdy <sup>5</sup>

- <sup>1</sup> Badacsony Research Station, Institute for Viticulture and Oenology, Hungarian University for Agriculture and Life Sciences, Római út 181, H-8261 Badacsonytomaj, Hungary; szoke.barna.arpad@uni-mate.hu
- <sup>2</sup> Institute of Informatics and Mathematics, Faculty of Wood Engineering and Creative Industries, University of Sopron, Bajcsy-Zsilinszky Str. 9, H-9400 Sopron, Hungary; smidla.work@gmail.com
- <sup>3</sup> Department of Viticulture, Institute for Viticulture and Oenology, Buda Campus, Hungarian University for Agriculture and Life Sciences, Villányi Str. 29-43, H-1118 Budapest, Hungary
- <sup>4</sup> Kecskemét Research Station, Institute for Viticulture and Oenology, Hungarian University for Agriculture and Life Sciences, Katona Zsigmond Str. 5, H-6000 Kecskemét, Hungary; olah.robert@uni-mate.hu
- <sup>5</sup> Department of Oenology, Institute for Viticulture and Oenology, Buda Campus, Hungarian University of Agriculture and Life Sciences, Villányi Str. 29-43, H-1118 Budapest, Hungary; nyitraine.sardy.diana.agnes@uni-mate.hu
- \* Correspondence: gyorffyne.jahnke.gizella@uni-mate.hu (G.J.); deak.tamas@uni-mate.hu (T.D.)

**Abstract:** Null alleles are alleles that are recessive to codominant markers without any effect on the phenotype. In SSR assays, there are several reasons for the lack of amplification at a locus: the primer does not bind well, longer fragments do not amplify due to imperfections in the PCR reaction, or the amount of DNA in the sample is insufficient. In microsatellite studies, null alleles are mostly used in pedigree analysis and population genetics calculations such as diversity estimation. Null alleles in pedigree analysis can cause rejection of the true parent; if not recognized while in population genetics they distort the results in underestimating diversity. In this review, the effects caused by null-alleles in viticultural research and its possible solutions were summarized.

Keywords: microsatellite; pedigree; parentage; population structure; diversity

# 1. The Null Allele Problem

The grapevine, being one of humanity's oldest crops, has significant economic and cultural value. Climate change is the most major challenge confronting viticulture today, and its long-term impacts are likely to be substantially minimized by new vine varieties that are better adapted to the environment. This emphasizes the significance of breeding. However, breeding success is dependent on an accurate understanding of the genetic diversity in the starting material.

A null allele is an allele that results in the complete absence of a gene product or function. The best-known example of a null allele is the human AB0 blood group system, where allele "0" is considered a null allele because it does not produce a phenotype or the presence of allele "A" or "B" masks it (people with genotype "A0" and "AA" are the same, blood group "A") [1].

Molecular and genetic markers based on PCR reactions (e.g., SSR) often show codominant inheritance, meaning that there may be alleles at certain loci that cannot be detected. Such alleles can be considered null alleles. However, it should be noted that null alleles are not exclusively associated with codominant inheritance. Dominant PCR-based methods (RAPDs, etc.) may also contain null alleles. Array-based SNPs also have null alleles and many of the same considerations apply, as SSRs. GWS (genome Wide Sequencing), on the other hand, does not have this problem unless conducted at low depth and then the missing data are estimated through imputation.



Citation: Jahnke, G.; Smidla, J.; Deák, T.; Oláh, R.; Szőke, B.Á.; Nyitrainé Sárdy, D.Á. The SSR Null Allele Problem, and Its Consequences in Pedigree Reconstruction and Population Genetic Studies in Viticulture. *Horticulturae* 2022, *8*, 658. https://doi.org/10.3390/ horticulturae8070658

Academic Editor: Paolo Sabbatini

Received: 30 June 2022 Accepted: 14 July 2022 Published: 19 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). If  $A_n$  is a null allele, then the  $A_iA_n$  and  $A_iA_i$  genotypes are indistinguishable. If an individual is homozygous for such a null allele, no product is formed in the PCR reaction and genotyping fails [2].

The failure to detect an allele in a PCR-based genetic marker may be due to several reasons:

- The primers used in the PCR reaction fail to bind to the DNA because the DNA sequence is different from the conservative reference sequence on which the detection is based [3]. This problem can also be caused by inappropriate primer design.
- Amplification of alleles of different sizes may differ, with "longer" alleles sometimes not amplified [4].
- The amount of DNA in the sample can also cause the lack of detection, because using the same DNA extract yields a PCR product in some loci but not in others [5,6].

For several reasons, null alleles are a problem in the use of genetic markers. In the most common applications in pedigree analyses or paternity tests, they can cause the erroneous exclusion of one or both parents by implying the putative parent in a locus to be homozygous when in fact it is heterozygous for a null allele [1]. For example, crossing a mother with genotype  $A_iA_i$  and a father with genotype  $A_jA_n$  gives a 50% chance of the offspring being of genotype  $A_iA_n$ , which apparently excludes the true father [2,7].

In population genetics studies, the presence of null alleles can result in an apparent reduction in the proportion of heterozygotes, which can greatly confound the assessment of the genetic diversity of a population. A null allele can lead to an overestimation of the frequency of detectable (non-zero) alleles, which can lead to a misrepresentation of population structure [2].

# 2. Methods for the Estimation of Null Allele Frequencies

Several methods exist for estimating null allele frequencies. Each of these methods assumes that the individuals in the sample under study form an approximately ideal population (in nature, no ideal population exists, but most populations approximate it). Thus, the Hardy–Weinberg rule is used to estimate the frequency of null alleles [8].

#### 2.1. Chakraborty's Method

The method created by Chakraborty et al. to estimate the frequency of the null allele uses the expected (Hexp) and observed (Hobs) numbers of heterozygotes [9,10]:

$$\hat{p}_{n(Chakraborty)} = \frac{H_{exp} - H_{obs}}{H_{exp} + H_{obs}}$$

This formula gives a reasonably accurate estimate but cannot be considered a maximum likelihood estimate and does not take into account individuals in the sample where genotyping failed (no PCR product was detected), as this may be caused not only by the locus being homozygous for a null allele [2].

#### 2.2. Brookfield's Method

Brookfield's estimate is applicable when there are no missing data in the sample [4]:

$$\hat{p}_{n(Brookfield)} = \frac{H_{exp} - H_{obs}}{H_{exp} + 1}$$

Summers and Amos's (1997) method is also based on similar estimates. It gives a good estimate with simulated data, but all three estimates have the same flaw: they all estimate from the initial genotype counts and add up the data before giving an estimate of the null allele frequency [11].

In 2006, Kalinowski and Taper developed a new method for estimating null alleles [2]. The method is called maximum likelihood estimation and can handle incomplete data. Kalinowski and Taper developed their algorithm with a slight modification of the EM

(Estimation-Maximization) algorithm [12,13] (described later) for estimating the '0' allele of the AB0 blood group.

To deal with the fact that in cases where the PCR reaction does not give a result, it is not known whether this is due to a technical problem or to a genotype homozygous for the null allele, a new variable was introduced. This variable ( $\beta$ ) expresses the probability that the failure of the PCR reaction is not due to the null homozygous genotype but to some other problem. Their algorithm is implemented in C#, and the software can be downloaded from the Internet [14]. This algorithm was also implemented as a part of the software MolMarker [15,16].

# 2.3. The EM (Expectation-Maximization) Algorithm and Its Use for Estimating Null Allele Frequencies

The EM algorithm was first formulated by Dempster and colleagues. The EM algorithm is an iterative method that aims to provide maximum likelihood estimates of the parameters of statistical models where the model itself depends on missing or hidden data. The EM iteration consists of two steps:

Step 1 E (Expectation): in this step, the missing data are calculated by training a conditional expected value based on the estimated values of the parameters.

Step 2 M (Maximization): based on the data calculated in the previous step and the existing data, a new estimate of the model parameters is made by maximizing the likelihood function.

The iterations are continued until the difference between the previous and the current value of the likelihood function is less than a predefined, sufficiently small value.

The EM algorithm can be used to estimate the frequency of null alleles in PCR-based genetic markers. In this case, heterozygotes carrying the null allele are indistinguishable from homozygotes carrying the detectable allele, so in this case the null allele can be considered as hidden data. The other problem is that if no product (missing data) is obtained in the PCR reaction, there are two possible reasons for this: (1) the tested individual is homozygous for the null allele at the locus or (2) the genotyping failed due to some other error.

### 3. The Null Allele Problem in Pedigree Reconstruction in Viticulture

# 3.1. The Importance of Pedigree Reconstruction from the Grape Breeder's Point of View

Grape variety performance is a genetically based complex polyfactorial feature whose expression is greatly impacted by ecological and agronomic conditions and is eventually reflected in yield evolution and synthetic varietal value [17,18].

Crosses in grape breeding must be designed and executed with the chosen grape types in mind. The grape varietals predicted to pass on the desired characteristic should be known to the breeder [19,20].

Kozma [18] authored a book on grape breeding that goes into great length about inbreeding and the heterosis effect when crossing and selfing grape cultivars. Cross breeding of grape varieties was where he saw the highest opportunity for increased heterozygosity.

Negrul [21,22] analyzed the influence of crossing within and across ecological groups of varieties on progeny population variability. While crossing within convarietas did not result in considerable variety, he discovered that crossing across ecological groups could result in a progeny population with high diversity. The results revealed that cross between convar. pontica and convar. orientlis produced mostly intermediate offspring, with the pontica variation having a modest advantage. The hybrids grew faster than orientalis cultivars in general, but they produced tiny, luscious berries that are suited for winemaking. The traits of the occidentalis variety predominate in the progeny of crosses between orientalis and occidentalis varieties. The hybrids produced by crossing occidentalis and pontica convarietas have intermediate features. Some of them have proven to be quite prolific and produce excellent wine. Interspecific crosses can also promote genetic variability. The resultant interspecific hybrids are mostly employed in resistance breeding, as cultivated species frequently lack or have lower resistance than wild species. The so-called Franco-American hybrids have been the most widely utilized in grapevine breeding to boost resistance to mildews, and grey rot [23]. *Vitis amurensis* is also a significant source of genetic diversity. Early ripening, resistance to mildews, *Botrytis*, and *Agrobacterium*, and excellent frost tolerance are all significant qualities for grape breeding. This latter feature makes it a particularly useful source of genes in continental climate grape breeding [24].

The correct selection of parent pairs is a cornerstone of combination breeding; therefore, a wrong choice can set back the breeding programme by up to 30 years. In the past (and even today), so-called combining ability tests have been (and still are) carried out to address this issue. For example, in 1978 by crossing 'Bicane' with 20 different varieties from the pontica, orientalis, and occidentalis ecological groups, the cultivar's combining ability was calculated. A total of 4659 seedlings were investigated from these 20 crossings. Studies on the flower type's inheritance and other biological and economic characteristics have revealed that 'Bicane' has a high combining ability and it is heterozygous for the flower type, allowing homozygous varieties of "White Riesling" and "Muscat Hamburg" to be separated. Because of the large range of variation for numerous features, as well as the heterosis seen for traits including vigor, cold hardness of buds, cane maturity, crop level, cluster size, and berry size, high-valuable genotypes were chosen to be homologated [25].

In F1 offspring of 20 crossings between seeded and seedless table grape cultivars, the combining ability and genotype-environmental interaction were investigated in relation to average cluster weight. Significant distinctions between them were required for their employment in combination breeding to be successful. The importance of genotype-environmental interactions was clearly stated, and they should be taken into consideration in breeding practice [26].

It is clear from the foregoing that in grape breeding, the origin and pedigree of varieties is often more important than the phenotypic characteristics of a given variety. However, experiments to test combining ability are very time-consuming. We can save this work (and often time) by looking at the pedigree of varieties with valuable characteristics that are important to us. As recent research has shown, often valuable varieties with a wide range of ecological tolerance come from the same ancestor or ancestors such as 'Heunish' [27,28] or Pinots [29,30]. However, the progeny found in the study of Bowers et al. [31] are all historically related with northeastern France and not with any other locations, which suggests that the crossings took place in this region. It is obvious that 'Pinot' and 'Gouais blanc' make a good parental combination; on the other hand, any other varieties growing in the area are most likely to be relatives of 'Pinot' or 'Gouais blanc' and would be less fit as a result of inbreeding depression.

The breeder should then go back to the successful ancestor to save time and money.

# 3.2. Consequences of the Presence of Null Alleles in Pedigree Studies-Some Examples

The most important consequence of the presence of null alleles is that they can result in a true parent and its offspring appearing homozygous for different alleles. This can lead to a rejection of the real parent–offspring relationship [1,32].

It is essential to maintain precise pedigree records whenever a grape breeding program is being carried out. There is a possibility that the breeder's record will contain errors. It is feasible to identify and validate parent–offspring connections by the use of genetic markers, often known as "DNA Fingerprinting". Markers known as simple sequence repeats (SSR) were applied in order to validate or rectify the pedigrees of grape varieties developed through the Cornell breeding program. In this project, 'Ontario' was confirmed as the parent of the 'Glenora', 'Himrod', and 'Alden' scoring null alleles at the VVMD25 locus. 'Fredonia' × 'Black Kishmish' were also confirmed as parents for 'Suffolk Red' considering the possibility of being a null allele at VVMD6 [33]. To prove that 'Muscat of Hamburg', a fine black table grape variety with a muscat flavor, is the progeny from the crossing of 'Schiava Grossa' and 'Muscat of Alexandria,' researchers used 2 isozymes (GPI and PGM), 30 nuclear, and 5 chloroplastic microsatellite markers. The likelihood of null alleles was calculated and found to be extremely low or absent [34].

Fifty microsatellite markers were used to examine ancient and closely related grape cultivars from the Alps: "Cornalin," "Humagne Rouge," and "Goron" from Valais (Switzerland); and "Cornalin," "Petit Rouge," and "Mayolet" from the Aosta Valley (Italy).

The findings supported earlier research demonstrating the distinction between the Italian and Swiss "Cornalin" cultivars and the identity between "Humagne Rouge" and "Cornalin" from the Aosta Valley [35]. 'Goron', 'Petit Rouge', 'Mayolet', and 'Cornalin d'Aoste' all share at least one allele with 'Cornalin du Valais', suggesting parent/offspring relationships. Forty-nine out of fifty microsatellite loci support 'Cornalin du Valais' as the offspring of 'Petit Rouge' and 'Mayolet', but 'Humagne Rouge' has genotype 257–241 at locus VVMD 8 instead of 241–241 for 'Cornalin d'Aoste'. This clonal variant is likely caused by a null allele in 'Cornalin d'Aoste'. This was the first grapevine paternity research to deal with discrepancy at a microsatellite locus, demonstrating that the use of progressively large numbers of loci in generating parentage decisions leads to a proportional rise in the risk of meeting a locus with intra-cultivar variability throughout the analysis. It should be assumed that a single multiple repeat unit disagreement is not sufficient to invalidate a parentage hypothesis.

At first, it was assumed that a parent–offspring link existed between the red grape cultivar known as 'Sangiovese,' which is the most common red grape cultivar in Italy, and the ancient Tuscan variety known as 'Ciliegiolo' [36]. During the process of testing 'Sangiovese' as a parent of 'Ciliegiolo,' the putative other parent was looked for in a large, private, and standardized database; however, no candidate was found. After putting 'Ciliegiolo' through its paces as a potential parent for 'Sangiovese,' a total of four candidate cultivars were discovered. Only one of the fifty microsatellites was not consistent with this paternity test, leading researchers to conclude with a high level of confidence that the grape variety known as 'Sangiovese' is the offspring of 'Ciliegiolo' and 'Calabrese di Montenuovo' [37]. In the same year, Staraz et al., on the basis of their studies, suggested that 'Ciliegiolo' was not the parent but the offspring of 'Sangiovese' [38]. This hypothesis was later confirmed, with the addition that the 'Ciliegiolo' variety was probably the offspring of the 'Sangiovese' × 'Moscato violetto' varieties [39].

Bergamini et al. saw the discovery of two possible parents for the 'Sangiovese' grape that had not been mentioned earlier. The first variety that could be considered a putative parent is known as 'Ciliegiolo', and it has already been discussed as a relative of 'Sangiovese'. The second variety that could be considered a putative parent is 'Negrodolce', which is an old local variety that was considered lost over the course of the last century but was recovered by the authors. The newly postulated parentage held up well even after a comprehensive molecular examination, with the exception of a single inconsistency detected in one of the 57 different microsatellite markers that were examined. This discrepancy is certainly due to a null allele, and as a result, it should not impair the hypothesis. However, it does point out the limitations of the microsatellites profiling as a pedigree research method, considering that this was the third different kinship that had been proposed so far for the 'Sangiovese' grape variety [40,41].

In contrast to the modest number of markers required to establish the identity or nonidentity of two grapevine samples, a substantially higher number of markers are required to reconstruct parentage and pedigrees across cultivars to prevent incorrect relationship assignment. The majority of parentage and kinship reconstruction research included more than 25, and in some cases more than 50 markers [42].

Another example where the null allele caused a dilemma was in the study of Bowers et al. Microsatellite loci in 300 grape cultivars were applied to determine paternal relationships. Sixteen wine grapes grown in northeastern France, including 'Chardonnay',

'Gamay noir', 'Aligote ', and 'Melon', have microsatellite genotypes compatible with being the descendants of a single pair of parents, 'Pinot' and 'Gouais blanc', both of which were widespread in this region in the Middle Ages. 'Romorantin' does not share an allele with 'Pinot' at locus VVS2, expressing a 129-bp allele instead. 'Pinot fin teinturier,' a red-juiced variety of 'Pinot,' has this trait, but no other cultivars do. 'Dameron' does not share an allele with 'Gouais blanc' at locus VVMD36, supposing a mutation to a 254-bp or null allele [31].

SSR markers were utilized to identify muscadine cultivars and validate their pedigrees. Utilizing 20 SSRs from 13 linkage groups, 89 *Vitis* accessions were genotyped. Five SSR markers could differentiate all 81 subgenus *Muscadinia* accessions. Twelve cultivars' profiles did not match their previously reported parentage–offspring connections. MicroChecker v2.2.3 [43] was used to identify genotyping mistakes related to null alleles (nonamplified alleles), short allele dominance (large allele dropout), and stutter peaks [44].

With the goal of analyzing genetic diversity and examining parentages, a collection of 1005 grapevine accessions were genotyped at 34 microsatellite (SSR) loci [39]. After a preliminary simulation that permitted the estimate of crucial values of likelihood ratios, the parentage analysis was carried out using the CERVUS program. To accommodate for genotyping mistakes, the occurrence of null alleles, and mutations, mismatches at a maximum of two loci in each trio were permitted. Because of the high frequency of null alleles at locus VChr9b, the data analysis indicated that the majority of mismatches occurred there. As a result, this locus was eliminated and the analysis was rerun. In most cases, incompatible profiles were found at loci with a high frequency of null alleles, and could thus be explained by the presence of a null allele in either parent or offspring.

In Croatia, 36 nuclear SSR, 4 cpSSR, and 47 SNP investigations revealed a large number of admixed varieties and synonyms, which was attributed to complex pedigrees and migrations. The highest fixation index, divergence from the Hardy–Weinberg equilibrium, and highest prevalence of null alleles were determined for the Vchr8b and Vchr14b loci, and hence they were removed from future parentage analyses. The remaining set of markers revealed 24 full parentages and 113 half-kinships [45].

In the analysis of identification and parentage, the condition of HW equilibrium is a key underlying assumption to have [39]. Minor deviations from HW equilibrium or variations at a few loci do not have the potential to distort likelihood estimates; however, deviations at many loci may have this potential. In the event that this is the case, the certainty of identification and parentage designations should be interpreted with extreme caution. However, it is important to keep in mind that the discrimination power of the loci in HW equilibrium may be sufficiently strong to ensure the validity of the study as a whole.

# 3.3. Solutions for the Correct Pedigree Reconstruction

A novel approach was suggested by Mark R. Christi to detect parent–offspring pairs in large data sets; to allow for genotyping errors, null alleles and mutations, it is necessary to quantitatively estimate how many loci should be allowed to mismatch based upon the study-specific error rate. This approach was suggested for application to methods that determine the probability of identity among genotypes and suggested that one can additionally account for null alleles, missing data and mutation simply by adding estimates of those rates to the study-specific error rate [46].

It was suggested that well-established maximum likelihood approaches for estimating relationship and relatedness could be modified to take into account null alleles. This would be accomplished by differentiating between an observed genotype and the set of true genotypes that could have produced that observation. For example, the probability of observing the genotype pair ii/ii was calculated by adding the probabilities that the true genotypes are ii/ii, in/ii, ii/in, or in/in—the four true genotypes that would be observed as ii/ii. [7].

Genetic data can be used to estimate the genealogical link or relatedness of individuals of unknown ancestry. ML-RELATE is a computer software that calculates maximum likelihood estimates of relatedness and connection. This software can handle null alleles and is designed for microsatellite data. It employs simulation to identify which links are supported by genetic data and to compare suspected relationships to alternatives [47].

Pedigrees are used in many areas of genetic research because they enable a precise resolution of genealogical ties between individuals. The estimation of the short-term effective population size (Ne), which is important in domains such as conservation genetics, is one example of how pedigree information might be used. Despite their use, pedigrees are frequently unknown parameters that must be derived from genetic data. Using Markov Chain Monte Carlo, a Bayesian technique for jointly estimating pedigrees and Ne from genetic markers was proposed. With the use of a composite likelihood, this method allows for the examination of a large number of markers and individuals within a single generation, considerably increasing computational efficiency. Simulated data were used to demonstrate that the approach can accurately determine Ne and relationships up to first cousins [48].

#### 4. The Null Allele Problem in Population Genetic Studies in Viticulture

#### 4.1. Importance of Population Genetic Studies in Wild and Cultivated Grapevines

"Population genetics is the study of the genetic composition of populations, including distributions and changes in genotype and phenotype frequency in response to the processes of natural selection, genetic drift, mutation and gene flow" [49]. In the case of crops, population genetics research is mainly used to study and understand the process of domestication in wild relatives [50,51].

In the case of grapes, research in this area was initially limited to exploring the genetic diversity of the varieties grown in a country or region [52–56]. In the last decade, research has concentrated on assessing the genetic diversity of wild relatives, mainly *Vitis vinifera* ssp. *sylvestris* in the case of *Vitis vinifera* L. ssp. *sativa* cultivars [57–63], and of wild American species, mainly *Vitis rupestris*, *Vitis berlandieri* and *Vitis riparia* in the case of rootstocks [64–69]. At the same time, climate change has increased the need to search for wild relatives of different origins carrying resistance genes to mildews and, more recently, black rot [70].

The conservation of genetic diversity has become increasingly important in recent years due to climate change. There is now no doubt that genetic diversity is worth preserving for "worse times". In domestication, genetic diversity is always reduced by strong selection pressure, which can significantly reduce the survivability of cultivated species.

A low number of genetic markers was thought to be adequate to discover domestication and breeding loci due to extended LD at these loci induced by positive selection for hermaphroditism, lighter skin color, and muscat aroma [71]. Geographic isolation contributes to differences between Eastern table grapes and Western wine grapes. The reduction in genetic variation at this locus on chromosome 11 suggests that the size difference between table and wine grapes may not be attributable to geography alone, but to selection for larger table grapes in the East. Whether berry size became a breeding target was largely a result of the predominant religion in the geographic area where the grapes were bred: table grapes were bred to be large in the East where alcohol was prohibited, while wine grapes retained the ancestral smaller size that is more desirable for winemaking in predominantly Christian regions in the West. This means that religious constraints about alcohol intake may have altered nucleotide variability in a genomic area linked to berry size [71,72].

Wine and table grape cultivars have self-fertile hermaphrodite flowers, but wild European, American, and Asian varieties are dioecious. It was assumed by Battilana et al. [73] that 1 of 55 potential haplotypes for three SSR markers around the sex locus in 132 *V. sylvestris* accessions and 171 *V. vinifera* cultivars accounted for 66% of hermaphrodite individuals and it may be the product of domestication. Specific size variants of the VVIB23 microsatellite sequence inside the 3'-UTR of a putative YABBY1 gene are related with the sex alleles M, H, and f; these markers can help define the status of wild grapevine germplasm. As *Vitis vinifera* L. ssp. *sylvestris* is still regarded a gene pool for viticulture,

allelic diversity at VVIB23 can be used to define wild germplasm and integrate it into marker-assisted breeding programs.

# 4.2. Consequences of the Presence of Null Alleles in Population Genetic Studies

The overestimation of visible alleles consequently led to the underestimation of genetic diversity present in a population. Considering that the non-visible alleles can differ from each other, the scale of underestimation could be even higher. A relatively high frequency of null alleles could be present in grapevine (*Vitis vinifera L.* ssp. *sativa* and *Vitis vinifera* ssp. *sylvestris*), which could be estimated at more than 20% (Figure 1).

Instead of being heterozygous with a null allele, samples with only one allele per locus were considered homozygous by Martinez et al. [53]. Based on their data, we estimated the frequency of null alleles by omitting varieties containing three alleles (due to possible gene duplication), which could be as high as 15% (VVMD32 at locus 2 in Figure 2). This may even lead to an underestimation of genetic diversity, although as the study cited is pioneering and relatively few genotypes were examined, this does not detract from the value of the study.



Figure 1. Cont.



**(B)** 

**Figure 1.** Ratio of alleles in SSR loci with relatively high estimated null allele frequency. Data from Riaz et al. [74], null allele frequencies estimated by MolMarker [16]. (A): VVIn73 locus; (B): VVMD21 locus.



**Figure 2.** Ratio of alleles in VVMD32 SSR locus with probably null alleles. Data from the study of Martinez et al. [53]; cultivars with triple alleles were ignored, and allele frequencies were calculated by MolMarker [16].

Genetic diversity of wild populations and cultivated varieties of grapevine were studied by De Andres et al. [75]. The problem caused by null alleles was not addressed, even though our calculations indicate that the estimated proportion of null alleles among *Vitis vinifera* L. ssp. *sativa* cultivars was 26.46%, while the proportion of null alleles in the woodland grape (*Vitis vinifera* L. ssp. *sylvestris*) varieties was as high as 26.82% (Table 1).

**Table 1.** Ratio of null alleles (Null) and number of alleles (No) in *Vitis vinifera* L. ssp. *sylvestris* populations in 25, and among *Vitis vinifera* L. ssp. *sativa* cultivars in 20 SSR loci respectively, calculated by MolMarker [16] from the data published by De Andres et al. [75].

	V. vinifera ssp. sylvestris		V. vinifera ssp. sativa	
	Null	No	Null	No
VMC1B11	14.84%	12	4.97%	13
VMC4F3-1	10.21%	14	3.62%	13
VVIB01	11.98%	6	13.81%	6
VVIH54	4.17%	12	11.88%	12
VVIN16	8.59%	6	12.78%	6
VVIN73	13.28%	7	26.46%	7
VVIP31	8.65%	15	4.16%	13
VVIP60	12.79%	9	8.56%	13
VVIQ52	9.35%	5	15.47%	6
VVIV37	11.72%	10	4.97%	12
VVIV67	8.33%	10	0.29%	14
VVMD21	18.75%	7	12.50%	8
VVMD24	10.18%	8	8.03%	8
VVMD25	1.30%	10	8.84%	12
VVMD27	11.75%	10	3.88%	9
VVMD28	8.56%	15	7.45%	14
VVMD32	8.31%	11	5.01%	9
VVMD5	7.29%	10	4.43%	10
VVMD7	11.46%	14	10.19%	10
VVS2	12.50%	14	3.59%	13
ZAG112	16.45%	8	n.a.	n.a.
ZAG29	26.82%	4	n.a.	n.a.
ZAG62	6.77%	8	n.a.	n.a.
ZAG67	13.84%	14	n.a.	n.a.
ZAG83	23.38%	7	n.a.	n.a.

n.a.-not analyzed.

Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grapes were analyzed by Emanuelli et al. [76]. While not specifically addressing the estimation of null alleles, it was noted that SSR diversity may be underestimated because sequencing of some microsatellite loci suggested that the polymorphism did not only correspond to a variation in the number of repeats, but also to changes in their architecture and the flanking regions with substitutions and long indels. This suggests that the variation in the number of repeats is not the only factor contributing to the polymorphism. A broad germplasm collection of 1.5 hectares encompassing most Iranian grape varieties was investigated utilizing 23 SSR loci. The effective number of alleles ranged from 1.82 (VrZAG83) to 9.73 (VMCNG2G7) with a mean of 4.41. Due to significant polymorphism, five SSR markers with PIC values above 0.80 were chosen for fast fingerprinting of grape genotypes. Heterozygosity ranged from 0.49 (VVMD17) to 0.97 (VrZAG64) without significant variations from predicted values. Brookfield's method was used to estimate null alleles [54].

Null alleles were considered and problematic cases were treated by Cao et al. [44] when they studied the pedigree and genetic diversity of Muscadine (*Vitis rotundifolia* syn. *Muscadinia rotundifolia*) grapes. Duplicates of a cultivar with missing data at one or two loci but identical at all other loci were regarded the same, and the cultivar with the least missing data was chosen to represent it. When a representative accession was chosen, data from other accession(s) of that cultivar were used to fill in any gaps in the chosen representative's locus. The genotyping errors caused by null alleles (nonamplified alleles), short allele dominance (large allele dropout), and the scoring of stutter peaks were identified using MicroChecker v2.2.3 [43].

Modern viticulture coexists with traditional winemaking in Montenegro, which still follows historic traditions and uses local varietals. As a result, this region offers what could be a chance to investigate processes that increase genetic diversity. In total, 419 samples in situ around the country (cultivated plants from old orchards and wild vines) and 57 local types maintained in a grapevine collection were collected and analyzed to determine the diversity of Montenegrin grapevines and the processes involved in their development. More than 100 of the 144 genetic profiles corresponded to farmed grapevines, indicating a high level of diversity [77]. The software STRUCTURE was used, which maintains missing data and null alleles.

Null alleles were also estimated by Mihaljevic et al. [45] for the genetic diversity and population structure analyses of Croatian grapevines as well. The highest fixation index, divergence from the Hardy–Weinberg equilibrium, and the highest prevalence of null alleles were calculated for the Vchr8b and Vchr14b loci (excluded from further parentage analyses). The risk of a null allele occurrence was low for all other alleles utilized in the parentage study. The informativeness of the SSR loci was measured as polymorphic information content (PIC) and ranged from 0.29 for VVIn73 to 0.78 for VVMD28, with an average PIC value of 0.66.

#### 4.3. Solutions for Minimizing the Error of the Estimates

Nuclear SSRs are known for having a high frequency of null alleles, or alleles that do not amplify and hence are recessive and undetectable in heterozygotes. Maximum likelihood methods that compare observed and expected homozygote rates in the population under the assumption of Hardy–Weinberg equilibrium and direct null allele frequency estimations from progeny where parent genotypes are known were compared [78]. It was demonstrated that estimations from the two techniques are comparable, especially when the number of maternal half-sib progeny arrays sampled is substantial. Population genetic metrics such as genetic differentiation ( $F_{ST}$ ) may be generally unbiased, with null allele frequencies ranging between 5% and 8% on average across loci. However, in fine-scale population research and parentage analyses, utilizing markers with such a high average prevalence of null alleles (up to 15% for some loci) can be extremely misleading [78].

Microsatellite null alleles can be found in all taxa to various degrees. They are problematic since they can inflate genetic differentiation measures and can be erroneously scored as homozygotes. Although numerous methods for correcting null allele proportions and estimating FST exist, nothing is known about how null alleles effect assignment testing. The percentage of successfully assigned individuals in model-based clustering and Bayesian assignment approaches was marginally, but significantly, lowered in the presence of null alleles, according to data based on simulations (frequency range from 0.000 to 0.913). The power to appropriately assign individuals was slightly reduced due to the bias in assignment tests generated by null alleles (0.2 and 1.0 percent units for STRUCTURE- and 2.4 percent units for GENECLASS-based assignment tests). Furthermore, the presence of null alleles resulted in a little, but statistically significant, overestimation of  $F_{ST}$ . As a result, null allele-affected microsatellite loci would likely have no effect on the overall outcome of assignment testing and could thus be included in these types of studies. However, loci prone to null alleles should be utilized with caution because they reduce the power of assignment tests and affect FST accuracy, and loci that are less prone to null alleles should always be favored [79].

To achieve useful results, genetic clustering algorithms require a specific amount of data. How sample group information can be used to obtain better outcomes, when the amount of data is limited, was demonstrated in the common circumstance of individuals being sampled at multiple locations. For the structure program, new models are being built for both admixture and no admixture instances. The prior distribution for each individual's population assignment is modified in these models. The revised prior distributions allow for a variation in the proportion of individuals assigned to each cluster based on their location. The models were validated using simulated data, and microsatellite data from the CEPH Human Genome Diversity Panel was used to show them. The new models were shown to detect structure at lower degrees of divergence or with less data than the original structure models or principal components approaches, and they were not biased toward recognizing structure when it was not present. These models are included in a new version of STRUCTURE, which may be downloaded for free at http://pritch.bsd.uchicago.edu/structure.html (accessed on 28 June 2022) [80].

# 5. Conclusions

Ignoring SSR null alleles may lead to scientifically incorrect conclusions. This means that researchers working on SSR studies should always ensure that any problems that may arise are addressed in some way. Omitting loci carrying null alleles from the study only appears to solve the problem and should only be recommended if it does not lead to excessive data loss. This might be the case, for example, if the minimum number of loci that would otherwise be recommended for a given species in a pedigree analysis can be achieved without loci carrying undetectable alleles and the remaining loci are well distributed in the genome.

Failure to address the null allele problem adequately could postpone breeding against the abiotic effects that climate change will bring to the forefront.

**Author Contributions:** Conceptualization, G.J., J.S. and D.Á.N.S.; writing—original draft preparation, G.J., T.D. and R.O.; writing—review and editing, all authors; supervision, B.Á.S. and D.Á.N.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Ministry for Innovation and Technology within the framework of the Thematic Excellence Programme 2020-National Challenges Subprogram (TKP2020-NKA-16) for research on adaptation to the changing environment in the vitivinicultural sector. This research was funded by National Research, Development and Innovation Office (NKFIH) grant numbers: K131679 and K 125476.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

### References

- 1. Dakin, E.E.; Avise, J.C. Microsatellite Null Alleles in Parentage Analysis. Heredity 2004, 93, 504–509. [CrossRef]
- Kalinowski, S.T.; Taper, M.L. Maximum Likelihood Estimation of the Frequency of Null Alleles at Microsatellite Loci. Conserv. Genet. 2006, 7, 991–995. [CrossRef]

- Kwok, S.; Kellogg, D.E.; Mckinney, N.; Spasic, D.; Goda, L.; Levenson, C.; Sninsky, J.J. Effects of Primer-Template Mismatches on the Polymerase Chain Reaction: Human Immunodeficiency Virus Type 1 Model Studies. *Nucleic Acids Res.* 1990, 18, 999. [CrossRef] [PubMed]
- Wattier, R.; Engel, C.R.; Saumitou-Laprade, P.; Valero, M. Short Allele Dominance as a Source of Heterozygote Deficiency at Microsatellite Loci: Experimental Evidence at the Dinucleotide Locus Gv1CT in Gracilaria Gracilis (Rhodophyta). *Mol. Ecol.* 1998, 7, 1569–1573. [CrossRef]
- 5. Garcia de Leon, F.J.; Canonne, M.; Quillet, E.; Bonhomme, F.; Chatain, B. The Application of Microsatellite Markers to Breeding Programmes in the Sea Bass, Dicentrarchus Labrax. *Aquaculture* **1998**, *159*, 303–316. [CrossRef]
- Gagneux, P.; Boesch, C.; Woodruff, D.S. Microsatellite Scoring Errors Associated with Noninvasive Genotyping Based on Nuclear DNA Amplified from Shed Hair. *Mol. Ecol.* 1997, *6*, 861–868. [CrossRef]
- 7. Wagner, A.P.; Creel, S.; Kalinowski, S.T. Estimating Relatedness and Relationships Using Microsatellite Loci with Null Alleles. *Heredity* **2006**, *97*, 336–345. [CrossRef]
- Excoffier, L.; Slatkin, M. Maximum-Likelihood Estimation of Molecular Haplotype Frequencies in a Diploid Population. *Mol. Biol. Evol.* 1995, 12, 921–927. [CrossRef]
- 9. Chakraborty, R.; Jin, L. Heterozygote Deficiency, Population Substructure and Their Implications in DNA Fingerprinting. *Hum. Genet.* **1992**, *88*, 267–272. [CrossRef]
- 10. Chakraborty, R.; de Andrade, M.; Daiger, S.P.; Budowle, B. Apparent Heterozygote Deficiencies Observed in DNA Typing Data and Their Implications in Forensic Applications. *Ann. Hum. Genet.* **1992**, *56*, 45–57. [CrossRef]
- 11. 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]
- 12. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm. *J. R. Stat. Soc. B* **1977**, *39*, 1–38.
- 13. Ceppellini, R.; Siniscalco, M.; Smith, C.A. The Estimation of Gene Frequencies in a Randomly Mating Population. *Ann. Hum. Genet.* **1955**, *20*, 97–115. [CrossRef]
- 14. Kalinowski, S. ML-Null Freq-Steven Kalinowski | Montana State University. Available online: https://www.montana.edu/kalinowski/software/null-freq.html (accessed on 12 May 2022).
- 15. Jahnke, G.; Smidla, J. MolMarker-Molekuláris Markerekkel Kapott Kutatási Eredmények Értékelését Segítő Felhasználóbarát Szoftver (MolMarker-User-Friendly Software to Help Evaluate the Research Results Obtained by Molecular Markers); Georgikon Napok Keszthely, M., Kis, L.B., Lukács, G., Nagy, B., Tóth, G., Eds.; Pannon Egyetem Georgikon Mezőgazdaságtudományi Kar: Keszthely, Hungary, 2014; pp. 135–143. Available online: https://napok.georgikon.hu/hu/cikkadatbazis/cikkek-2012/doc\_download/203gyorffyne-jahnke-gizella-smidla-jozsef-molmarker-molekularis-markerekkel-kapott-kutatasi-eredmenyek-ertekeleset-segitofelhasznalobarat-szoftver (accessed on 28 June 2022).
- 16. Jahnke, G.; Smidla, J.; Poczai, P. MolMarker: A Simple Tool for DNA Fingerprinting Studies and Polymorphic Information Content Calculation. *Diversity* **2022**, *14*, 497. [CrossRef]
- 17. Kozma, P. A Szőlő Teljesítőképessége Növelésének Lehetőségei Keresztezéses Nemesítéssel—Possibilities of Increasing the Productivity of the Vine by Cross-Breeding-Academic Inaugural Lecture; Hungarian Academy of Sciences: Budapest, Hungary, 1973.
- 18. Kozma, P. A Szőlő Nemesítése. (Breeding of the Vine); Mezőgazdasági Kiadó: Budapest, Hungary, 1951.
- 19. Hajdu, E. A Kertészeti Növények Nemesítése a Szőlő Példáján. (The Horticultural Plant Breeding on the Example of Grape. *Gradus* **2016**, *3*, 378–383.
- Hajdu, E. 6—Grapevine breeding in Hungary. In Woodhead Publishing Series in Food Science, Technology and Nutrition, Grapevine Breeding Programs for the Wine Industry; Reynolds, A., Ed.; Woodhead Publishing: Cambridge, UK, 2015; pp. 103–134. ISBN 9781782420750. [CrossRef]
- 21. Negrul, A.M. Questions of Origin and Breeding of the Grape Vine on a Genetical Basis. *Genetika* 1968, 4, 84–97.
- 22. Negrul', A.M. Questions of the Origin and Breeding of Vines on a Genetical Basis. NI Vavilov Agric. Sci. 1969, 323–339.
- Kozma, P.I. A Szőlőrezisztencia-Nemesítés Szempontjai Es Módszerei Magyarországon. Int. J. Hortic. Sci. 2002, 8, 43–48. [CrossRef]
- 24. Korbuly, J. A Vitis Amurensis (Rupp.) Értékei a Szőlő Rezisztencia Nemesítés Számára. Int. J. Hortic. Sci. 2002, 8, 53–58. [CrossRef]
- 25. Gorodea, G.; Neagu, M.I. Combining Ability of the Cultivar "Bicane"; INRA: Paris, France, 1978.
- 26. Roytchev, V. Combining Ability and Genotype-Environmental Variability of Cluster Weight in F-1 Crosses between Seeded and Seedless Grape-Vine (*Vitis vinifera* L.) Cultivars. *Genet. Breed.* **2006**, *35*, 29–38.
- 27. 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 J. Grapevine Res.* **2015**, *54*, 79–86. [CrossRef]
- 28. Boursiquot, J.-M.; Lacombe, T.; Bowers, J.E.; Meredith, C.P. Le Gouais, Un Cépage Clé Du Patrimoine Viticole Européen. *Bull. L'O.I.V.* **2004**, *77*, 5–19.
- Vouillamoz, J.F.; Grando, M.S. Genealogy of Wine Grape Cultivars: 'Pinot' Is Related to 'Syrah'. *Heredity* 2006, 97, 102–110. [CrossRef]
- This, P.; Lacombe, T.; Thomas, M.R. Historical Origins and Genetic Diversity of Wine Grapes. *Trends Genet.* 2006, 22, 511–519. [CrossRef] [PubMed]

- 31. Bowers, J.; Boursiquot, J.M.; This, P.; Chu, K.; Johansson, H.; Meredith, C. Historical Genetics: The Parentage of Chardonnay, Gamay, and Other Wine Grapes of Northeastern France. *Science* **1999**, *285*, 1562–1565. [CrossRef] [PubMed]
- Jones, A.G.; Small, C.M.; Paczolt, K.A.; Ratterman, N.L. A Practical Guide to Methods of Parentage Analysis. *Mol. Ecol. Resour.* 2010, 10, 6–30. [CrossRef]
- Bautista, J.; Dangl, G.S.; Yang, J.; Reisch, B.; Stover, E. Use of Genetic Markers to Assess Pedigrees of Grape Cultivars and Breeding Program Selections. Am. J. Enol. Vitic. 2008, 59, 248–254.
- 34. Crespan, M. The Parentage of Muscat of Hamburg. Vitis 2003, 42, 193–197.
- 35. Vouillamoz, J.; Maigre, D.; Meredith, C.P. Microsatellite Analysis of Ancient Alpine Grape Cultivars: Pedigree Reconstruction of *Vitis Vinifera* L. "Cornalin Du Valais". *Theor. Appl. Genet.* **2003**, *107*, 448–454. [CrossRef]
- 36. Crespan, M.; Calo, A.; Costacurta, A.; Milani, N.; Giust, M.; Carraro, R.; di Stefano, R. Ciliegiolo and Aglianicone: Only One Variety in a Direct Relationship with Sangiovese [*Vitis vinifera* L.; Grapevine]. *Riv. Vitic. E Enol.* **2002**, *55*, 3–14.
- Vouillamoz, J.F.; Monaco, A.; Costantini, L.; Stefanini, M.; Scienza, A.; Grando, M.S. The Parentage of "Sangiovese", the Most Important Italian Wine Grape. Vitis 2007, 46, 19–22.
- Staraz, M.D.V.; Bandinelli, R.; Boselli, M.; This, P.; Boursiquot, J.M.; Laucou, V.; Lacombe, T.; Varès, D. Genetic Structuring and Parentage Analysis for Evolutionary Studies in Grapevine: Kin Group and Origin of the Cultivar Sangiovese Revealed. J. Am. Soc. Hortic. Sci. 2007, 132, 514–524. [CrossRef]
- 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 Synonymy and Parentages, and Reveals a Large Admixture amongst Varieties of Different Geographic Origin. *Theor. Appl. Genet.* 2010, 121, 1569–1585. [CrossRef]
- Bergamini, C.; Caputo, A.R.; Gasparro, M.; Perniola, R.; Cardone, M.F.; Antonacci, D. Evidences for an Alternative Genealogy of "Sangiovese". *Mol. Biotechnol.* 2013, 53, 278–288. [CrossRef]
- Bergamini, C.; Caputo, A.R.; Perniola, R.; Cardone, M.F.; Gasparro, M.; Antonacci, D. Genealogy of "sangiovese": A new hypothesis. In Proceedings of the 35th World Congress of Vine and Wine, Izmir, Turkey, 18–22 June 2012; OIV: Paris, France, 2012.
- Sefc, K.; Pejić, I.; Maletić, E. Microsatellite Markers for Grapevine: Tools for Cultivar Identification & Pedigree Reconstruction. In Grapevine Molecular Physiology & Biotechnology; Springer: Dordrecht, The Netherlands, 2009; pp. 565–596.
- 43. Van Oosterhout, C.; Hutchinson, W.F.; Wills, D.P.M.; Shipley, P. Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Mol. Ecol. Notes* **2004**, *4*, 535–538. [CrossRef]
- 44. Cao, S.; Stringer, S.; Gunawan, G.; McGregor, C.; Conner, P.J. Genetic Diversity and Pedigree Analysis of Muscadine Grape Using SSR Markers. *J. Am. Soc. Hortic. Sci.* 2020, 145, 143–151. [CrossRef]
- 45. Mihaljević, M.Ž.; Maletić, E.; Preiner, D.; Zdunić, G.; Bubola, M.; Zyprian, E.; Pejić, I. Genetic Diversity, Population Structure, and Parentage Analysis of Croatian Grapevine Germplasm. *Genes* **2020**, *11*, 737. [CrossRef]
- 46. Christie, M.R. Parentage in Natural Populations: Novel Methods to Detect Parent-Offspring Pairs in Large Data Sets. *Mol. Ecol. Resour.* **2010**, *10*, *115–128*. [CrossRef]
- Kalinowski, S.T.; Wagner, A.P.; Taper, M.L. ML-RELATE: A Computer Program for Maximum Likelihood Estimation of Relatedness and Relationship. *Mol. Ecol. Notes* 2006, *6*, 576–579. [CrossRef]
- 48. Ko, A.; Nielsen, R. Joint Estimation of Pedigrees and Effective Population Size Using Markov Chain Monte Carlo. *Genetics* **2019**, 212, 855–868. [CrossRef]
- Population Genetics-Latest Research and News | Nature. Available online: <a href="https://www.nature.com/subjects/population-genetics">https://www.nature.com/subjects/population-genetics</a> (accessed on 7 June 2022).
- Aravanopoulos, F.A.; Ganopoulos, I.; Tsaftaris, A. Population and Conservation Genomics in Forest and Fruit Trees. *Adv. Bot. Res.* 2015, 74, 125–155. [CrossRef]
- 51. Purugganan, M.D. Evolutionary Insights into the Nature of Plant Domestication. *Curr. Biol.* **2019**, *29*, R705–R714. [CrossRef] [PubMed]
- Marrano, A.; Grzeskowiak, L.; Sanz, P.M.; Lorenzi, S.; Prazzoli, M.L.; Arzumanov, A.; Amanova, M.; Failla, O.; Maghradze, D.; Grando, M.S. Genetic Diversity and Relationships in the Grapevine Germplasm Collection from Central Asia. VITIS J. Grapevine Res. 2015, 54, 233–237. [CrossRef]
- Martínez, L.E.; Cavagnaro, P.F.; Masuelli, R.W.; Zúñiga, M. SSR-Based Assessment of Genetic Diversity in South American Vitis Vinifera Varieties. Plant Sci. 2006, 170, 1036–1044. [CrossRef]
- Doulati-Baneh, H.; Mohammadi, S.A.; Labra, M. Genetic Structure and Diversity Analysis in *Vitis Vinifera* L. Cultivars from Iran Using SSR Markers. *Sci. Hortic.* 2013, 160, 29–36. [CrossRef]
- Jahnke, G.; Májer, J.; Lakatos, A.; Molnár, J.G.; Deák, E.; Stefanovits-Bányai, E.; Varga, P. Isoenzyme and Microsatellite Analysis of Vitis Vinifera L. Varieties from the Hungarian Grape Germplasm. Sci. Hortic. 2009, 120, 213–221. [CrossRef]
- 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.
- 57. Scali, M.; Zifferero, A.; Vignani, R. Distribution and Characterization of the *Vitis vinifera* L. Subsp Sylvestris in Southern Tuscany. *Recent Pat. Biotechnol.* **2018**, *12*, 208–220. [CrossRef]

- Zdunić, G.; Maul, E.; Dias, J.E.E.; Organero, G.M.; Carka, F.; Maletic, E.; Savvides, S.; Jahnke, G.G.; Nagy, Z.A.; Nikolić, D.; et al. Guiding Principles for Identification, Evaluation and Conservation of *Vitis vinifera* L. Subsp Sylvestris. *Vitis J. Grapevine Res.* 2017, 56, 127–131. [CrossRef]
- 59. Söylemezoğlu, G.; Ağaoğlu, Y.S.; Uzun, H.I. Ampelographic Characteristics and Isozymic Analysis of *Vitis vinifera* spp. Sylvestris Gmel. In Southwestern Turkey. *Biotechnol. Biotechnol. Equip.* **2001**, *15*, 106–113. [CrossRef]
- 60. Bodor, P.; Höhn, M.; Pedryc, A.; Deák, T.; Dücso, I.; Uzun, I.; Cseke, K.; Böhm, É.I.; Bisztray, G.D. Conservation Value of the Native Hungarian Wild Grape (*Vitis sylvestris* Gmel.) Evaluated by Microsatellite Markers. *Vitis J. Grapevine Res.* **2010**, *49*, 23–27.
- Jahnke, G.; Nagy, Z.A.; Koltai, G.; Hajdú, E.; Májer, J. Preservation and Characterization of Woodland Grape (*Vitis vinifera* ssp. Sylvestris GMEL) Genotypes of the Szigetköz, Hungary. In *Germplasm: Characteristics, Diversity and Preservation*; Walton, M., Ed.; Nova Science Publishers: New York, NY, USA, 2016; pp. 27–45.
- 62. Jahnke, G.; Nagy, Z.A.; Koltai, G.; Oláh, R.; Májer, J. Morphological, phenological and molecular diversity of woodland grape (*Vitis sylvestris* gmel.) genotypes from the szigetköz, hungary. *Mitt. Klosterneubg.* **2021**, *71*, 90–98.
- 63. Zoghlami, N.; Riahi, L.; Laucou, V.; Mliki, A.; Ghorbel, A.; This, P. Genetic Structure of Endangered Wild Grapevine *Vitis vinifera* ssp. Sylvestris Populations from Tunisia: Implications for Conservation and Management. *For. Ecol. Manag.* **2013**, *310*, 896–902. [CrossRef]
- Jahnke, G.; Nagy, Z.; Taller, J.; Májer, J.; Kocsis, L. Application of Isozymes and SSR Markers for the Analysis of the Genetic Background of Some Rootstocks Derived from Teleki's Seedlings (Teleki 5C, Kober 5BB, SO4). *Mitt. Klosterneubg.* 2015, 65, 221–236.
- 65. Arnold, C.; Schnitzler, A. Ecology and Genetics of Natural Populations of North American Vitis Species Used as Rootstocks in European Grapevine Breeding Programs. *Front. Plant Sci.* **2020**, *11*, 866. [CrossRef]
- 66. Poczai, P.; Hyvönen, J.; Taller, J.; Jahnke, G.; Kocsis, L. Phylogenetic Analyses of Teleki Grapevine Rootstocks Using Three Chloroplast DNA Markers. *Plant Mol. Biol. Report.* **2013**, *31*, 371–386. [CrossRef]
- 67. Lin, H.; Walker, M.A. Identifying Grape Rootstocks With Simple Sequence Repeat (SSR) DNA Markers. *Am. J. Enol. Vitic.* **1998**, 49, 403–407.
- 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 Simple Sequence Repeat Markers as a Starting Point for New Rootstock Breeding Programs. Am. J. Enol. Vitic. 2019, 70, 390–397. [CrossRef]
- 69. Jahnke, G.; Kocsisné Molnár, G.; Májer, J.; Szöke, B.; Tarczal, E.; Varga, P.; Kocsis, L. Analysis of Grape Rootstocks by SSR Markers. J. Int. Des Sci. Vigne Vin 2011, 45, 199–210. [CrossRef]
- 70. Băileștianu, N.A.; Mitrea, R. The black rot—A new challenge for vine crops. *Ann. Univ. Craiova Agric. Montanol. Cadastre Ser.* **2020**, *49*, 20–25.
- Migicovsky, Z.; Sawler, J.; Gardner, K.M.; Aradhya, M.K.; Prins, B.H.; Schwaninger, H.R.; Bustamante, C.D.; Buckler, E.S.; Zhong, G.Y.; Brown, P.J.; et al. Patterns of Genomic and Phenomic Diversity in Wine and Table Grapes. *Hortic. Res.* 2017, *4*, 17035. [CrossRef] [PubMed]
- 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] [PubMed]
- 73. Battilana, J.; Lorenzi, S.; Moreira, F.M.; Moreno-Sanz, P.; Failla, O.; Emanuelli, F.; Grando, M.S. Linkage Mapping and Molecular Diversity at the Flower Sex Locus in Wild and Cultivated Grapevine Reveal a Prominent Ssr Haplotype in Hermaphrodite Plants. *Mol. Biotechnol.* **2013**, *54*, 1031–1037. [CrossRef] [PubMed]
- 74. Riaz, S.; de Lorenzis, G.; Velasco, D.; Koehmstedt, A.; Maghradze, D.; Bobokashvili, Z.; Musayev, M.; Zdunic, G.; Laucou, V.; Andrew Walker, M.; et al. Genetic Diversity Analysis of Cultivated and Wild Grapevine (*Vitis vinifera* L.) Accessions around the Mediterranean Basin and Central Asia. *BMC Plant Biol.* **2018**, *18*, 137. [CrossRef] [PubMed]
- 75. De AndrÉs, M.T.; Benito, A.; PÉrez-Rivera, G.; Ocete, R.; Lopez, M.A.; Gaforio, L.; Muñoz, G.; Cabello, F.; MartÍnez Zapater, J.M.; Arroyo-GarcÍa, R. Genetic Diversity of Wild Grapevine Populations in Spain and Their Genetic Relationships with Cultivated Grapevines. *Mol. Ecol.* 2012, 21, 800–816. [CrossRef] [PubMed]
- 76. 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] [PubMed]
- 77. Maraš, V.; Tello, J.; Gazivoda, A.; Mugoša, M.; Perišić, M.; Raičević, J.; Štajner, N.; Ocete, R.; Božović, V.; Popović, T.; et al. Population Genetic Analysis in Old Montenegrin Vineyards Reveals Ancient Ways Currently Active to Generate Diversity in *Vitis vinifera*. Sci. Rep. 2020, 10, 15000. [CrossRef]
- 78. Oddou-Muratorio, S.; Vendramin, G.G.; Buiteveld, J.; Fady, B. Population Estimators or Progeny Tests: What Is the Best Method to Assess Null Allele Frequencies at SSR Loci? *Conserv. Genet.* **2009**, *10*, 1343–1347. [CrossRef]
- 79. Carlsson, J. Effects of Microsatellite Null Alleles on Assignment Testing. J. Hered. 2008, 99, 616–623. [CrossRef]
- 80. Hubisz, M.J.; Falush, D.; Stephens, M.; Pritchard, J.K. Inferring Weak Population Structure with the Assistance of Sample Group Information. *Mol. Ecol. Resour.* 2009, *9*, 1322–1332. [CrossRef]