Supplemental Data for:
Migliaro D, De Lorenzis G, Di Lorenzo GS, De Nardi B, Gardiman M, Failla O, Brancadoro L and Crespan M. 2019. Grapevine non-*vinifera* genetic diversity assessed by simple sequence repeat markers as a starting point for new rootstock breeding programs. Am J Enol Vitic 70:390-397. doi: 10.5344/ajev.2019.18054.

Supplemental Text

Supplemental Text 1. Selection of the most suitable set of simple sequence repeat (SSR) loci for the genotyping of rootstocks.

Twenty-two SSR loci were used to fingerprint the entire University of Milan rootstock collection. Seventeen out of the 22 markers were retained for data analysis. An overview of the most common troubles we had during amplification and allele sizing is reported here.

VVMD32 and VMC6G1 peaks can be highly unbalanced. The two putative alleles can show strongly different heights giving rise to doubts in allele calling (Figure 1A). When the peaks are generally low, the already disadvantaged alleles may no longer be visible, and their presence/absence can be difficult to estimate.

VVMD32 also appeared to be critical in Laucou et al. (2011), showing a higher percentage of missing data in hybrids and rootstocks (6.89 and 15.56%, respectively) than in *sativa* and *sylvestris* accessions, compared with the whole set of 20 SSR used. The authors explained those missing data as possibly homozygous null.

Another kind of problem arose with VVMD25 marker. In this case, the new alleles determined a continuous variability in the allelic series, with differences of only one basepair (bp) among adjacent alleles (Supplemental Table 4 and Figure 1B), leading to a critical allele calling, even using the bin set. These presumed new alleles showed an intermediate length compared with those already recognized in *Vitis vinifera*. The new alleles should be verified by the crossing and genotyping segregating alleles in the progeny, but some combinations of very close alleles, differing by only one bp, were already present in our data set, such as Lindley (245-246), genotype 22, genotype 64, and genotype 57 (242-243). The allelic differences found in VVMD25 could not only be due to the number of tandem repeats of the microsatellite motif, but also to additional differences in the amplicon electric charge. Sequencing would be necessary to verify this hypothesis. We suggest discarding VVMD25 for non-*vinifera* genotyping because allele calling was shown to be objectively vulnerable. However, despite the great difficulties encountered, we decided to use this marker because we took great care in comparing allele lengths.

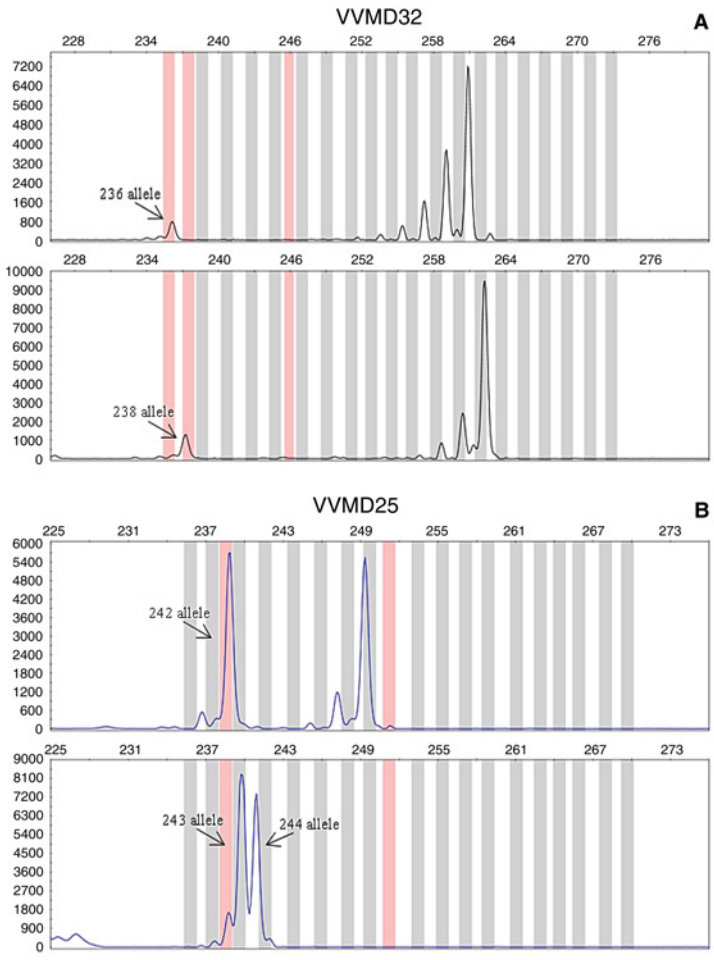

Figure 1 Critical signals for two "international" simple sequence repeats: strongly unbalanced height of peaks for VVMD32 (**A**) and two adjacent alleles, differing by only 1 basepair, for VVMD25 (**B**).

VMCNG4b9 is very good for *V. vinifera* genotyping but was shown to be problematic with rootstocks because of multiple signals obtained with silver staining (Crespan et al. 2009). It was retrieved because the use of fluorescence simplified the pattern and the reading compared to silver staining. This marker showed more than two alleles in many genotypes; those over 180 bp in length were not taken into consideration because, in the light of experience with *V. vinifera*, we hypothesized that they derive from the amplification of an additional locus. The presence of these extra alleles, however, is useful to classify unknown genotypes as non-*vinifera* grapevines. Like other SSRs, VMCNG4b9 showed a continuous variability in the allelic series (Supplemental Table 4), but for this marker, the assignment of allele length was very easy using the bin set.

Concerning VChr SSR markers, 26 were first tested on a subset of 26 rootstocks. Amplification data were used to select the most suitable seven loci. Four VChr SSRs were shown to be unsuitable for non-*vinifera* genotyping, as they showed no polymorphism (VChr1c), null alleles (i.e., no amplification in some samples [VChr15a and VChr17b]), or more than two alleles per genotype (VChr8a). Among the five VChr used to genotype the entire collection, the most polymorphic was VChr3a, showing a very well-separated allelic series and a nice distribution of frequencies, the most frequent allele (181) having $p = 0.17$. VChr3a was also the highest polymorphic in *V. vinifera* (Cipriani et al. 2010), but in rootstocks the list of alleles is longer, 37 versus 13. The other four VChr markers were less polymorphic, showing 12-13 alleles each. VChr10b and VChr19a showed the most regular allelic sequence, whereas, VChr4a and VChr7b showed some very close alleles, differing by only one bp.

Supplemental Text 2. Detailed description of genotype identification.

Obtaining the correct correspondence between molecular profile and variety name, when available, proved to be very difficult because of the poor information on many accessions. Even pedigree records, which could strongly support this exploration, evidenced many limits, mainly due to the following issues: (1) generic information on the parents used for crossing, (2) some rootstocks derived from open pollination, and (3) one or both parents of the bred variety were different from those declared.

A progressive genotype number was given to samples with missing information; the declared parents were checked using the available molecular profiles. We were aware that the number of SSRs used was too low for pedigree studies, but they can be useful for identification purposes if they support the information given by breeders. There are also some limits in using this kind of information because the pedigree given by breeders could be incorrect (Laucou et al. 2008). When at least one of the parents was compatible, a name was proposed for the genotype, even if still uncertain, awaiting the support of ampelography.

Our 3306 C profile (Supplemental Table 3, SSR profile ID 15) matched that of Garris et al. (2009) and of the VIVC website, so we concluded that the SSR profile of 3306 C, given as reference by de Andrés et al. (2007), was very probably wrong; moreover, it matched our genotype 1. Berlandieri Rességuier n. 2, given as reference by de Andrés et al. (2007), was also incorrect because many historical pedigrees (Supplemental Table 1) were compatible with our Berlandieri Rességuier n. 2 SSR profile, which in turn matched the one reported as reference in the VIVC Catalogue (Supplemental Table 3, SSR profile ID 232). Notably, two accessions named Rességuier n. 2, analyzed in our work and coming from Spain (genotype 166, Supplemental Table 3, SSR profile ID 53), matched the profile proposed by de Andrés et al. (2007).

In the case of Paulsen's crosses, some were identified using literature data, others were proposed here as reference by cross-referencing available information on accessions sharing the same name and coming from different collections, and using historical pedigree data. Paulsen 764, 775, 779, 782, and 1103 (Supplemental Table 3, SSR profiles ID 203, 23, 24, 204, 5) belong to the progeny Rességuier n. 2 × Rupestris du Lot and were consistent with the declared parents. Contrary to what was reported in the literature (Cosmo 1958, Galet 1988), 1045 and 1447 Paulsen (Supplemental Table 3, SSR profile ID 2, 9) were incompatible as offspring of Rességuier n. 2 × Ganzin 1 and Berlandieri × Rupestris Martin, respectively; our data showed that they may still belong to the series Rességuier n. 2 × Rupestris du Lot. 1043 Paulsen accession (genotype 113, Supplemental Table 3, SSR profile ID 179) was probably wrong because neither of the declared parents, Rességuier n. 2 and Ganzin 1 (Galet 1988), was consistent with our SSR profile.

As regards Richter's crosses, even if the history of these hybrids is complex and unclear, as explained by Galet (1988), it was not difficult identifying our samples (31, 57, 99, and 110 Richter, Table S3, SSR profiles ID 14, 21, 25, 4) using literature information. Interestingly, our data reject 110 Richter being a progeny of Rességuier n. 2 × Rupestris Martin; instead, it could be PO with Rupestris du Lot.

Some Castel crosses were identified through the literature: 196-17, 216-3, and 6736 (Supplemental Table 3, SSR profiles ID 12, 13, 22). Others were proposed here as probably correct because the pedigree information matched; these are 228-1, 215-1, and 227-1 (Supplemental Table 3, SSR profile ID 80, 229, 230).

1132 Grimaldi (Supplemental Table 3, SSR profile ID 182) could be correct, given that Uva di Troia is compatible as one of the genitors.

The best known Couderc hybrids were easily identified (161-49, 1613, 3306, 3309, 93-5, Supplemental Table 3, SSR profiles ID 10, 11, 15, 16, 63), while doubts remained for the others. Likewise, Millardet et de Grasset crosses 101-14, 106-8, 41 B, and 420 A were easily identified (Supplemental Table 3, SSR profiles ID 1, 3, 19, 20); 17-37 Millardet et de Grasset accession from CREA of Conegliano was verified thanks to the 17-37 clone I-V.G.V.A.33 kindly provided for comparison by Vivai Paulsen of Palermo (Italy) (Supplemental

Supplemental Data for:
Migliaro D, De Lorenzis G, Di Lorenzo GS, De Nardi B, Gardiman M, Failla O, Brancadoro L and Crespan M. 2019. Grapevine non-*vinifera* genetic diversity assessed by simple sequence repeat markers as a starting point for new rootstock breeding programs. Am J Enol Vitic 70:390-397. doi: 10.5344/ajev.2019.18054.

Table 3, SSR profile ID 62). Concerning Ruggeri hybrids, only 140 Ruggeri (Supplemental Table 3, SSR profile ID 8) was identified with certainty even if, given available SSR data, it cannot be a progeny of Rességuier n. 2 × Rupestris du Lot; only the male parent was consistent. The other Ruggeri varieties, 205, 240, and 300 (Supplemental Table 3, SSR profile ID 193, 194, 195) were proposed here as reference, the last two supported by the correspondence of the sex of flowers with literature information.

Genotype 151 (Supplemental Table 3, SSR profile ID 217) was identified as Riparia tomenteux (de Andrés et al. 2007), even if incompatible as the genitor of 3306 C and 3309 C. Genotype 104 and genotype 143 (Table S3, SSR profile ID 170 and 209) seemed to derive from self-fertilization of Golia, a hermaphrodite rootstock obtained by Angelo Pirovano crossing Castel 15-612 × Rupestris du Lot.

As expected, none of the 39 genotypes in the F series (FnPn) were identified because they were progenies obtained from the University of Milan breeding programs. Some believed different FnPn plants shared the same genotype, probably due to labeling errors during vegetative propagation. This result highlights the uniqueness of this plant material.

Supplemental Text 3. Detailed description of kinship results.

We found that the incorrect Rességuier n. 2 by de Andrés et al. (2007) (Supplemental Table 3, SSR profile ID 53), recorded here as genotype 166, could be a Rességuier n. 2 × Gloire de Montpellier cross. Concerning known rootstock varieties, 13-44 Evex and 13-5 Evex historical pedigree shows that both are open pollinated progenies of Rességuier n. 2. From our data, they may derive from a spontaneous Rességuier n. 2 × 34 EM cross. Martinez Zaporta 5A historical pedigree shows that it came from 41 B crossed with an unknown variety, from our data it may derive from 41 B × 3309 Couderc. The declared genitors of 99 Richter were Berlandieri Las Sorres × Rupestris du Lot, the first one missing in our data set; however, our data showed that this rootstock could be a progeny of Rességuier n. 2 × Rupestris du Lot. Concerning the genitors of the new rootstocks of the series “M” selected by the University of Milan (Supplemental Table 3, SSR profile ID 64, 65, 66, 67), some cannot be verified, because absent from the present research and the literature, others can be supported or excluded by the available molecular data.