Methods for conservation of intra-varietal genetic variability in ancient grapevine varieties

E. Gonçalves[1] and A. Martins[1,2]

[1] Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal

[2] Associação Portuguesa para a Diversidade da Videira – PORVID, Tapada da Ajuda, 1349-017 Lisboa, Portugal

Abstract. Conservation of intra-varietal genetic variability is a crucial strategy to preserve traditional viticulture and to face future challenges. Conservation can be performed through two processes: (1) conservation without other immediate objectives; (2) conservation which simultaneously ensures the evaluation of important quantitative traits (such as yield, must quality traits, tolerance to abiotic stresses). The second strategy permits the quantification of intra-varietal genetic variability for important quantitative traits and the selection of a superior group of genotypes (polyclonal selection). Based on the results obtained from the conservation of ancient Portuguese varieties, the abovementioned methods for conservation of intra-varietal diversity are compared, and the advantages and disadvantages of each one are discussed.

1. Introduction

A very high level of genetic diversity was naturally created and accumulated over centuries within ancient varieties, but it has been exposed to a process of accelerated erosion since the mid of the twentieth century. Therefore, the conservation of intra-varietal genetic variability is a crucial strategy to preserve traditional viticulture and to face future challenges (climate change, consumer demands, etc.).

To be efficient, the conservation must contemplate representative samples of the diversity within all the autochthonous varieties. The sample size depends on the total growing area as well as on the number of discontinuous growing regions of the variety. Simulations studies found a minimum sample size per region with around 70 genotypes [1]. As varieties are usually grown in several regions, samples with 400 or more genotypes are used. This strategy of conservation is currently conducted in Portugal by PORVID (Portuguese Association for Grapevine Diversity).

Conservation can be performed through two processes: (1) conservation without other immediate objectives; (2) conservation which simultaneously ensures the evaluation of important quantitative traits (such as yield, must quality traits, tolerance to abiotic stresses), including quantification of intra-varietal genetic variability for those traits, and selection of a superior group of genotypes. The first method of conservation is done in pots or/and in the field with a single plot per genotype. The second method includes the plantation of field trials (this means that involves experimental design). As regards the latter, the ultimate goal is to ensure that these collections of genotypes are not only repositories of genetic variability, but they can be simultaneously evaluated and utilized for selection.

When the conservation is done in pots or/and in field with a single plot per genotype (each genotype represented by 3–6 plants), the only evaluation that can be obtained is the mean phenotypic value of the plot for a target trait. Error and genotypic variances cannot be estimated as well as the predicted genotypic values cannot be obtained. For quantification of intra-varietal genetic variability and selection of quantitative traits this method is completely inefficient because the genotypic value is overwhelmed by environmental deviations. A simulation study was conducted using yield data from a field trial with 100 genotypes of Tinta Miúda variety with four blocks and four plants per plot [2]. Randomly extracting 10000 times a plant from each genotype (corresponds to simulating 10000 times a population of individual plants), estimates of broad-sense heritability between zero and 0.25 were obtained for phenotypic values from one plant. Therefore, as previously mentioned, the usefulness of this type of collection is only the conservation.

The field conservation with experimental design comprises partially replicated designs (augmented designs, alpha-alpha designs) or fully replicated designs (typically randomized complete block designs, or designs of the family of incomplete block designs, such as alpha and row-column designs).

Numerous variants of partially replicated designs (also known as unreplicated trials) are frequently used in plant breeding, mainly for preliminary assessment of the available germplasm [3]. These field trials use two types of treatments, check and test, which correspond to replicated and non-replicated treatments, respectively. The main objective is to select a sub-group of genotypes (normally around 1/3) for later more detailed studies. In grapevine, this type of experimental design was studied for conservation of rarely grown varieties [4]. The objective was to develop a method of preserving and quantifying intra-varietal genetic diversity which provides some guidelines for the future experimental procedures. Through simulation several experimental situations were generated,

varying the numbers of test and check treatments and the type of experimental design associated with the check treatments (augmented randomized complete block design and alpha-alpha design). A greater precision was achieved in the quantification of the intra-varietal genetic diversity with collections with an alpha-alpha design and of over 250 genotypes and a minimum of 33% of plots containing check genotypes. These authors also concluded that the partially replicated designs are useful in the context of ancient grapevine varieties to evaluate the intra-varietal genetic variability, but not to make a successful genetic selection. The main advantage of this type of experimental design is the smaller experimental area required.

In grapevine, the use of fully replicated designs means that the field trial would cover a large area (from 0.75 to 1.5 ha) which, by itself, causes large environmental variation. Therefore, the importance of the experimental design to control spatial variability is crucial to reach the objectives of quantification of genetic variability and selection. With the aim of identifying the most suitable experimental designs for these objectives, a study was carried out to assess via simulation the comparative efficiency of various experimental designs [5]: randomized complete block design, alpha design [6] and row-column design [7,8]. A higher efficiency of row-column designs compared with randomized complete blocks, or even with alpha designs, was found, enabling more precise estimates of genotypic variance, greater precision in the prediction of genetic gain and consequently greater efficiency in polyclonal selection.

Nowadays, all types of conservation above mentioned are used in Portugal. However, methodological studies are currently performed to reinforce the efficiency of the conservation which simultaneously ensures the evaluation of important quantitative traits and selection.

As regard to fully replicated field trials, traditionally they were planted according to a randomized complete block design. But, over the past decade, with the methodological advances in the area of experimental designs in grapevine, designs of the family of the incomplete blocks began to be implemented, particularly, resolvable row-column designs. The first data are now being available.

The efficiency of the fully replicated designs depends on the randomization process used to control environmental variation (particularly, spatial variation), and the number of replicates. The methodological study presented in this work aims to demonstrate the importance of the number of replicates to improve the accuracy and precision of quantification of intra-varietal variability and genetic selection.

2. Material and methods

The following example was focused on a field trial of Tinta Caiada variety where a large number of genotypes is evaluated concerning the most important traits. It is located in South Portugal (Reguengos de Monsaraz, Alentejo region), and it was established with 208 genotypes and a resolvable row-column design (RCD) with 5 replicates of 3 plants per plot. Each resolvable replicate comprises 13 rows and 16 columns. This design was generated in CycDesigN package [9].

The data analysis was based on the mean yield data per plot (kg/plant) from 2 years. The following model was fitted to yield data (the model for a RCD):

$$y_{ijlm} = \mu + u_{gi} + u_{rj} + u_{col(r)jl} + u_{row(r)jm} + e_{ijlm},$$

for $i = 1, \cdots, 208$, $j = 1, \cdots, 5$, $l = 1, \cdots, 16$, $m = 1, \cdots, 13$. The y_{ijlm} represent the observations, μ the population mean, u_{gi} the genotypic effects, u_{rj} the resolvable replicate effects, $u_{col(r)jl}$ the column effects within replicates, $u_{row(r)jm}$ the row effects within replicates, and e_{ijlm}, the random errors associated with individual plots. Model effects (with the exception of μ) were assumed independent and identically distributed normal variables with zero mean and corresponding variance. All random effects were assumed mutually independent. Residual maximum likelihood (REML) estimation method was used for variance parameter estimation. Linear mixed models were fitted using the ASReml-R package [10] in R software.

For comparing the effect of the number of replicates in the quantification of intra-varietal variability and in the selection, data from the 5 possible combinations of 4 repetitions and from the 10 possible combinations of 3 and 2 repetitions were used. Consequently, a resolvable row-column design (RCD) with 5 replicates (hereafter named as design RCD5) was compared to a RCD with 4, 3 and 2 repetitions (hereafter named as designs RCD4, RCD3, RCD2, respectively). The interpretation of the results was based on the indicators following described.

(1) Coefficient of genotypic variation (CV_G) (ratio between the estimate for the genotypic standard deviation and the overall mean) was computed for a better comparison of genetic variability estimate among experimental designs.

(2) A generalized measure of broad-sense heritability (H^2) [11] was calculated to evaluate how much phenotypic variability is due to genetic causes.

(3) The relative bias (RB) and relative mean square error (RMSE) associated to the estimates of the coefficient of genotypic variation and broadsense heritability for RCD4, RCD3, and RCD2 were obtained to evaluate the accuracy and the precision of the quantification of the intra-varietal variability and selection.

(4) Empirical Best Linear Unbiased Predictors (EBLUPs) of genotypic effects of the yield obtained from the fitting of the RCD5 model and from each of the RCD4, RCD3 and RCD2 models were ranked, and a set of 20 top-ranked genotypes (higher EBLUP values) were selected. The percentage of overlapping genotypes in each two sets of 20 selected clones (RCD5 and other design with lower number of replicates) was obtained. In this way, the final result of polyclonal selection performed according to RCD5 and each of the RCD4, RCD3 and RCD2 was compared.

3. Results

The estimates obtained for the genotypic coefficient of variation and the broad-sense heritability resulting from the fitting of a linear mixed model to the yield mean data for RCD5 design (5 repetitions) and for all possible combinations of 4, 3 and 2 repetitions are shown in Table 1.

Table 1. Quantification of intra-varietal genetic variability (CVG), broad-sense heritability (H2) estimates according to the different designs (RCD5, RCD4, RCD3, RCD2), the relative bias (RB) and relative mean square error (RMSE) associated to the estimates of the coefficient of genotypic variation and broad-sense heritability, and percentage of overlapping genotypes with the selection of 20 genotypes performed by the RCD5 and other design with lower number of replicates.

Number of replicates	CV_G(%)	H^2	Overlap of Genotypes (%)
RCD5	**27.711**	**0.719**	
RCD4	26.923	0.634	70.0
RCD4	27.891	0.672	80.0
RCD4	26.868	0.646	70.0
RCD4	28.658	0.706	70.0
RCD4	27.928	0.692	75.0
Mean	27.654	0.670	73.0
RB (%)	−0.208	−6.810	
RMSE (%)	0.075	0.642	
RCD3	27.176	0.559	60.0
RCD3	24.703	0.490	55.0
RCD3	26.533	0.564	60.0
RCD3	28.861	0.625	65.0
RCD3	28.672	0.642	70.0
RCD3	28.203	0.628	55.0
RCD3	26.923	0.581	55.0
RCD3	28.374	0.641	70.0
RCD3	27.106	0.604	55.0
RCD3	28.713	0.670	70.0
Mean	27.526	0.600	61.5
RB (%)	−0.667	−16.487	
RMSE (%)	0.226	3.260	
RCD2	22.869	0.324	40.0
RCD2	30.173	0.546	55.0
RCD2	28.141	0.487	55.0
RCD2	27.107	0.500	50.0
RCD2	27.390	0.484	55.0
RCD2	22.512	0.348	35.0
RCD2	27.593	0.523	40.0
RCD2	29.298	0.570	50.0
RCD2	28.701	0.589	60.0
RCD2	28.252	0.571	55.0
Mean	27.204	0.494	49.5
RB (%)	−1.832	−31.245	
RMSE (%)	0.879	11.360	

With regard to the quantification of the intra-varietal variability, for all possible combinations of replicates, the values of the CV_G were close to the value obtained with RCD5 design. The bias and the mean squared error associated to the estimate of CV_G increased with the decrease of the number of replicates, although, in a small level (RB and RMSE values were all close to zero). In sum, for the quantification of the intra-varietal variability, the use of a trial with a RC design with 2 repetitions may be acceptable to achieve that objective.

When the goal is the genetic selection, conclusions are completely different. As regards this latter goal, the number of replicates of the trial proved to be a
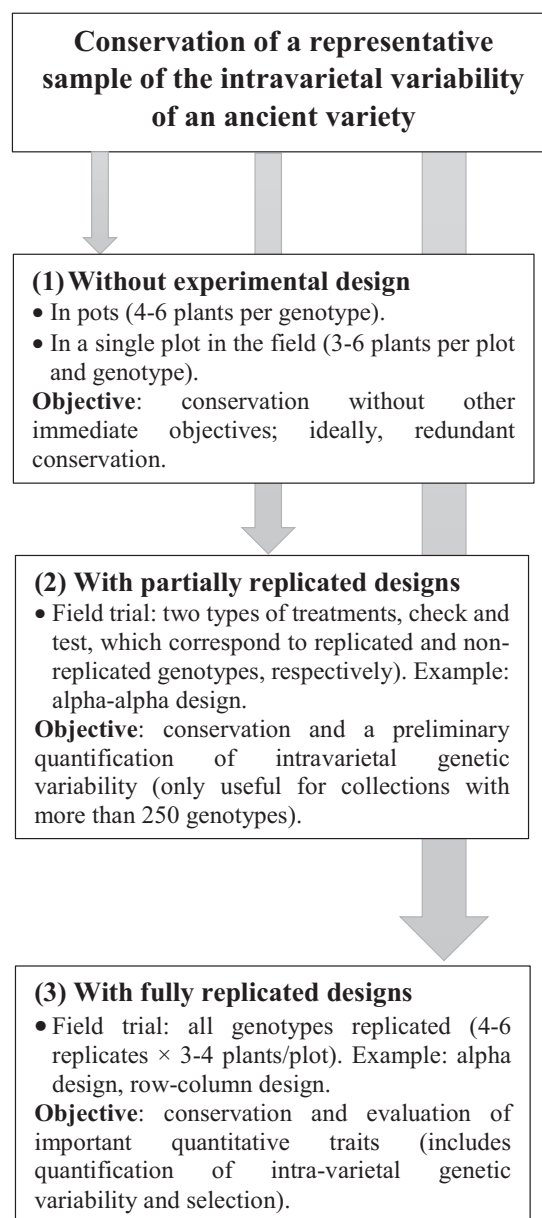
Conservation of a representative sample of the intravarietal variability of an ancient variety

(1) Without experimental design
- In pots (4-6 plants per genotype).
- In a single plot in the field (3-6 plants per plot and genotype).

Objective: conservation without other immediate objectives; ideally, redundant conservation.

(2) With partially replicated designs
- Field trial: two types of treatments, check and test, which correspond to replicated and non-replicated genotypes, respectively). Example: alpha-alpha design.

Objective: conservation and a preliminary quantification of intravarietal genetic variability (only useful for collections with more than 250 genotypes).

(3) With fully replicated designs
- Field trial: all genotypes replicated (4-6 replicates × 3-4 plants/plot). Example: alpha design, row-column design.

Objective: conservation and evaluation of important quantitative traits (includes quantification of intra-varietal genetic variability and selection).

Figure 1. Methodologies of intra-varietal variability conservation of ancient grapevine varieties.

very important issue. For all possible combinations of 4 replicates, the broad-sense heritability varied from 0.634 to 0.706, and was, on average, 0.67, which is below the reference value (RCD5, 0.719). This tendency was particularly pronounced with the decrease in the number of repetitions. In fact, when the results with 3 replicates are analysed, the values for this genetic parameter decreased, and ranged from 0.49 to 0.67, and was, on average, 0.60. Lower values were obtained with 2 replicates, varying from 0.32 to 0.59. For the all possible combinations of 2 replicates, in average, only about 50% of the phenotypic variability was explained by genetic causes, which compromises a successful genetic selection. Also, the bias (accuracy) and the mean squared error (precision) of the broad sense heritability estimates decreased as the number of replicates increased.

Selection consists in the selection of a top-ranked set of genotypes concerning target traits. The ranking of genotypes (the final result of polyclonal selection) was

also compared between RCD5 and the other designs. The results showed in Table 1 reveal the importance of the number of replicates in the group of genotypes selected. Performing a polyclonal selection (selecting a group of 20 top-ranked genotypes), the percentage of overlapping genotypes in the polyclonal selection according to RCD5 and RCD4 varied from 70% to 80%, and in average, 73% of the genotypes were coincident. This result indicates the loss of efficiency of the selection performed with the RCD4, but once again, this tendency was particularly pronounced with the decrease in the number of repetitions.

The polyclonal selection performed using the RCD3 shares with the RCD5 about 61.5% of the genotypes (ranging from 55% to 70%). When the RCD2 is used only about half of the genotypes are coincident with the selection performed with the RCD5.

In sum, it is obvious the decrease in the efficiency of selection due to the decrease of the number of replicates. Definitely, a fully replicated design with only 2 repetitions is no longer adequate to perform selection.

4. Conclusions

The methodologies of intra-varietal variability conservation of ancient grapevine varieties are summarized in Fig. 1.

There are several ways of ensuring proper conservation of the intra-varietal genetic variability of ancient grapevine varieties. Nevertheless, all possible options are not necessarily adequate to value the potential benefits of the varieties for grape growers in a short time through selection. This last goal implies necessarily the establishment of well-planned field trials, ideally, full replicated designs with not less than 4 replicates.

References

[1] E. Gonçalves, A. Martins, *Plant Breeding*, edited by A.Y. Ibrokhim (InTech, 2012), p. 333

[2] A. Martins, E. Gonçalves, *Grapevine Breeding Programs for the Wine Industry: Traditional and Molecular Techniques*, edited by A.G. Reynolds (Woodhead Publishing Elsevier, UK, 2015), p. 159

[3] R.A. Kempton, A.C. Gleeson, *Statistical methods for plant variety evaluation*, edited by R.A. Kempton, P.N. Fox (Chapman & Hall, 1997), p. 86

[4] E. Gonçalves, A.St. Aubyn, A. Martins, Tree Genet. Genomes **9**, 65 (2013)

[5] E. Gonçalves, A. St. Aubyn, A. Martins, Heredity **104**, 552 (2010)

[6] H.D. Patterson, E.R. Williams, Biometrika **63**, 83 (1976)

[7] E.R. Wiliams, J.A. John, Appl. Stat. **38**, 149 (1989)

[8] E.R. Williams, J.A. John, D. Whitaker, Biometrics **62**, 103 (2006)

[9] D. Whitaker, E.R. Williams, J.A. John, *CycDesigN: a package for computer generation of experimental designs* (2007)

[10] D.G. Butler, B.R. Cullis, A.R. Gilmour, B.J. Gogel, R. Thompson, *ASReml-R reference manual* version 4 (University of Wollongong, Wollongong, 2017)

[11] E. Gonçalves, I. Carrasquinho, A. St. Aubyn, A. Martins, Euphytica **189**, 379 (2013)