Combining genotyping by sequencing and genomic prediction within bi-parental crosses to speed up selection of grapevine cultivars

Timothée Flutre, Amandine Launay, Agnes Doligez, Eric Duchêne, Elisa Marguerit, P.F. Bert, Pierre Mournet, Loic Le Cunff

To cite this version:
Timothée Flutre, Amandine Launay, Agnes Doligez, Eric Duchêne, Elisa Marguerit, et al.. Combining genotyping by sequencing and genomic prediction within bi-parental crosses to speed up selection of grapevine cultivars. 12. International Conference on Grapevine Breeding and Genetics, Jul 2018, Bordeaux, France. hal-01848564

HAL Id: hal-01848564
https://hal.archives-ouvertes.fr/hal-01848564
Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Combining genotyping by sequencing and genomic prediction within bi-parental crosses to speed up selection of grapevine cultivars

Flutre T., Launay A., Dolbez A., Duchêne E., Marguerit E., Bert P.-F., Mournet P., Le Cunff L.

1 AGAP, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France
2 SVGV, Université de Strasbourg, INRA, Université de Strasbourg, F-68000 Colmar, France
3 UMR EGFV, Bordeaux Sciences Agro, INRA, Université de Bordeaux, F-33882 Villenave d’Ornon, France
4 IFV, Le Grau-du-Roi, France

timothée.flutre@inra.fr

1. Abstract

A major challenge facing viticulture consists in decreasing inputs, especially synthesized fungicides, and adapting to climate change, while maintaining berry quality and differentiated wine styles, as well as reasonable profits to the winemakers. In this endeavor, breeding new varieties remains an important lever. To speed up this process, genomic selection may be seen as advantageous to quickly test in the field candidates preciously selected for various complex traits.

We assess here the feasibility of this approach based on three bi-parental crosses (Syrah x Grenache, Riesling x Gewurztraminer and Cabernet-Sauvignon x Riparia Gloire de Montpellier), and several types of traits (yield components, phenology, terpenols, transpiration) whose phenotypes were already used for QTL mapping. The plant material was genotyped by sequencing (Keaygene patents), providing around 18000 filtered, uniformly distributed SNPs per genotype. Depending on genetic architecture and broad-sense heritability, the average accuracy of genomic prediction obtained by cross-validation can go from 0% (transpiration rate), 50% (budbreak) to 80% (berry weight).

In conclusion, such an integrated approach aims at exploiting the full potential of available phenotypic, genotypic and genealogical data in order to efficiently keep adapting the plant material. In parallel, the approach is also tested in a specific breeding program for the rosé wines, in order to speed up the selection of genotypes resistant to powdery and downy mildew, and adapted to the rosé vinification. Results will be compared to classical selection operated after offsprings are planted in the vineyard.

2. Crosses and experimental designs

<table>
<thead>
<tr>
<th>cross</th>
<th>offsprings</th>
<th>design</th>
<th>trait</th>
</tr>
</thead>
<tbody>
<tr>
<td>SxG</td>
<td>192</td>
<td>field</td>
<td>berry weight</td>
</tr>
<tr>
<td>RIxGW</td>
<td>256</td>
<td>field</td>
<td>budbreak</td>
</tr>
<tr>
<td>CSGRM</td>
<td>120</td>
<td>greenhouse</td>
<td>transpiration rate</td>
</tr>
</tbody>
</table>

Pedigree and field layout of the reciprocal cross, Syrah x Grenache, in Montpellier:

3. Statistical analysis of the phenotypes

\[ y = \beta S + \gamma G + \epsilon \text{ where } y \sim N(\mu, \sigma^2) \text{ and } \epsilon \sim N(0, \sigma^2_{\text{id}}) : \text{ fitted with lme4.} \]

1. fit by maximum likelihood (ML) several linear mixed models with different fixed effects (\(\beta\))
2. select the best based on the Akaike Information Criterion (AIC) and check its assumptions visually (e.g. residuals)
3. refit the best model by restricted maximum likelihood (REML) to obtain the broad-sense heritability (\(H^2\)) and the Best Linear Unbiased Predictors (BLUPs) of the total genotypic value (\(g\))

4. Genotyping

After classical DNA extraction, samples were digested with the ApaI restriction enzyme. Libraries multiplexing 36 samples were sequenced on Illumina HiSeq 3000 (paired-end - 2x75).

As this resulted in substantial non-uniformity, a second sequencing phase was performed (single-end - 1x150).

5. Genomic prediction

BLUP (\(g\)) = \(\mu + M_\alpha + \epsilon\) where \(\alpha \sim N(0, \sigma^2)\) and \(\epsilon \sim N(0, \sigma^2_{\text{id}}) : \text{ fitted with cgs3.} \)

The proportion of variance explained (PVE) shows that genotyped SNPs capture most of the genetic variance. Prediction accuracy is assessed by correlation (\(\rho\)) via cross-validation (50 random replicates).

6. Genomic selection

The IFV started in 2015 a pilot project to breed new varieties for rosé wine and experiment with a strategy to speed-up the whole process. The goal is to obtain several varieties, not only with two resistance genes for each of downy and powdery mildew, but also adapted to the peculiarities of rosé vinification.

Crosses between iconic varieties for rosé in two years to phenotype traits related to the ideotype defined by experts, viticulturists and winemakers, such as juice oxidability, berry acidity, etc. Prediction models can then be trained on this subset, and used to predict genotypic values of non-phenotyped offsprings produced in the subsequent years. At this stage, the idea is to select a subset of resistant genotypes with sufficiently high performance on the other traits so that they can bypass phase 2 and be directly assessed in several field sites (phase 3).

7. Conclusion and perspectives

Genomic prediction looks promising enough to be experimented in grapevine breeding programs, yet its usefulness depends on several key factors. Genotyping is now flexibly achieved via restriction-assisted DNA sequencing, and statistical models and software implementations are available. The crucial component remains phenotyping protocols and tools, which should be cheap enough to be performed on large progenies. The phenotyped traits should then also be closely related to the ideotype of interest.

8. Funding and acknowledgments

INRA (FruitSelGen project), CASDAR (EDGAR project), South Green and URGHi platforms.