Optimisation de la sélection de variétés de vigne à l'aide laprédiction génomique et phénomique

Charlotte Brault

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En Génétique et Amélioration des Plantes

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

Portée par

Unité mixte de recherche AGAP Institut
Amélioration génétique et adaptation des plantes méditerranéennes et tropicales

Grapevine breeding optimization with genomic and phenomic predictions

Présentée par Charlotte BRAULT

Le 6 décembre 2021

Sous la direction de Patrice This,
co-encadrée par Agnès Doligez, Timothée Flutre, Loïc Le Cunff et Vincent Segura

Devant le jury composé de

Jean-Luc JANNINK, Directeur de recherche, USDA-ARS	Rapporteur
Laurence MOREAU, Directrice de recherche, INRAE – GQE Le Moulon	Rapportrice
Laurent BOUFFIER, Chargé de recherche, INRAE - Biogeco	Examinateur
Jacques DAVID, Professeur, Institut Agro – AGAP Institut	Président du jury
Christopher SAUVAGE, Responsable projet génétique, Syngenta	Examinateur
Agnès DOLIGEZ, Ingénieur de recherche, INRAE – AGAP Institut	Co-encadrante
Patrice THIS, Directeur de recherche, INRAE – AGAP Institut	Directeur de thèse
Viviane BECART, Ingénieur viticole, Institut Rhodanien	Invitée

UNIVERSITÉ DE MONTPELLIER

l'institut Agro
agriculture • alimentation • environnement

Montpellier SupAgro

Remerciements

Comme quoi une thèse ça peut être une partie de plaisir, c'est peut-être rarement le cas mais ça l'a été pour moi grâce à toutes les personnes que j'ai côtoyées pendant ces trois belles années.

Je tiens tout d'abord à remercier vivement tous mes encadrants pour m'avoir fait confiance pour réaliser cette thèse. Un grand merci donc à Patrice pour avoir dirigé cette thèse, pour avoir été d'une bienveillance extraordinaire et avoir su arrêter les débats lorsque c'était nécessaire. Merci à Timothée d'avoir été là au début, je n'aurais pas pu mieux commencer qu'avec toi, tu as su poser les jalons d'une thèse que j'espère réussie ! Merci à Agnès d'avoir été si présente, si attentive à tous mes besoins, merci pour ta rigueur intellectuelle et scientifique dans nos discussions et dans tes relectures (il le fallait !). Merci à Loïc pour m'avoir fait découvrir les joies de l'IFV, pour avoir apporté ta bonne humeur, ton humour et ton grain de sel (breton ?) dans toutes les réunions (enfin celles où tu étais présent). Et enfin merci à Vincent d'avoir rejoint l'équipe d'encadrants et de m'avoir initiée à la NIRS, ta joie de vivre, et tes connaissances m'ont été d'une aide précieuse. C'était vraiment une chance pour moi de vous avoir tous en tant qu'encadrants, je garde un très bon souvenir de nos réunions de thèse de 3h où les discussions scientifiques n'en finissent jamais, où les graphiques se succèdent et où l'on se rend compte que je n'aurais jamais le temps de tout faire ! Encore un grand merci donc de m'avoir fait découvrir de la meilleure manière possible la recherche scientifique.

Je tiens également à remercier tout particulièrement Inter-Rhône pour avoir financé cette thèse tout en me laissant une totale liberté dans mon travail. Merci à Viviane Bécart pour les échanges constructifs sur les caractéristiques de la vigne idéale et merci pour les bouteilles !

Merci à toutes celles et ceux qui m'ont entourée durant ces trois années. Dans l'équipe DAAV, merci à Valérie d'animer chaque jour nos repas en râlant sur le menu, à Thierry pour ta gentillesse et ton calme légendaire, à Cathy qui m'a supportée cette dernière année, à Gautier et Amandine qui m'ont souvent débloquée dans mes analyses, à Philippe pour tes relectures avisées, à Cédric G et Angélique pour leur dépannage express, à Charles d'avoir pris le temps de m'expliquer la vie des baies, à Victoria qui deviendra bientôt une pro de R et de bash, mais aussi merci à Maryline, Virginie, Roberto, Béatrice, Jean-Pierre. Une spéciale dédicace à Cédric M. de m'avoir changé les idées, d'avoir toujours été là pour les cafés, les bières ou les burgers et plus encore :)

Mon travail de thèse a été rendu possible grâce aux données phénotypiques disponibles, merci donc à l'ensemble de l'équipe DAAV pour le phénotypage et à Yves, Gilles, Pierre et François pour l'entretien des vignes. Merci à Amandine pour le génotypage GBS. Un grand merci à toutes les personnes qui se sont motivées pour venir récolter des feuilles ou des sarments pour la NIRS, malgré la météo !

Mention spéciale à Matthieu qui est arrivé le même jour que moi et que je n'ai pas perdu de vue depuis le 1er jour ! Merci pour les très nombreuses pauses café (/cookies) qui se sont transformées en soirées jeux, puis en amitié sincère. J'en profite pour remercier Juliette A.

pour m'avoir accompagnée dans ce magnifique séjour de rédaction et à de nombreuses sorties/soirées. Merci à Romain d'avoir fait l'effort de rentrer dans notre monde de biologistes, merci de m'avoir initiée au sombre langage bash en retour. Viennent ensuite mes stagiaires préférés, Juliette et Miguel, qui m'ont appris à être encadrante, bon courage à vous pour la suite. Je souhaite remercier globalement les doctorants d'AGAP, Léo, Aurélie, Kelly, Cédric, Stella, Benjamin, Michel, Lison, Laïla et bien d'autres que j'oublie, merci de m'avoir fait me sentir moins seule dans les quelques galères du doctorat ! Merci à nos partenaires vigne du LEPSE, Aude et Adriaan pour l'échange de connaissances et de savoir-faire.

Merci à Jacques David pour m'avoir donné l'opportunité de donner des TD à SupAgro. Merci également aux organisateurs des groupes scientifiques R2D2, GQMS2 et Groupe Vigne, vous m'avez permis de présenter de nombreuses fois mes travaux et avez grandement contribué à ma culture et à ma curiosité scientifique. Je remercie également tous ceux qui ont aiguillé ma thèse, tout d'abord les membres de mes comités de thèse Florent Pantin, Thierry Simonneau, Hélène Muranty, Tristan Mary-Huard, Komlan Avia et Jean-Marc Lacape ; et les statisticiens Julien Chiquet et Marie Perrot-Dockès avec qui j'ai eu le plaisir de collaborer.

Quelques mots pour les non-montpellierains, que j'ai dû quitter pour cette thèse. Merci aux Fontenoises (Héloïse, Marie, Pauline, Amélie, Elodie, Pierrette, Fanélie, Camille) de répondre présentes pour des retrouvailles en Bourgogne ou ailleurs. Merci à Estelle pour nos trop rares sorties cheval (enfin ce n'est peut-être pas plus mal vu mon nombre de chutes au compteur !). Merci à la fine équipe grignonnaise (aka Lait Keep) pour nos fous rire qui me manquent, merci à Anaïs, Marie, Aline, Benjamin, Mariane, Basma, Grégoire, Juliette, Elsa et Alexandra la québécoise -et pour toujours dans nos cœurs, chacun de nous emporte le souvenir du vieux parc de Grignon-. Bon courage à celles et ceux qui sont encore en thèse, je pense fort à vous ! Bien sûr, toutes mes pensées vont à Paul qui m'a toujours soutenue et aimée pendant ces 3+3 années. Tu as été d'un soutien inconditionnel.

Merci enfin plus que tout à ma famille. A mes parents qui, sans comprendre le sujet de ma thèse me comprennent. A Delphine et Antoine pour ces moments de pur bonheur que vous m'avez apportés et pour qui ces trois années n'auront pas été de tout repos (et ce n'est pas fini) !

Contents

List of Figures

List of Tables

List of Abbreviations

AOP	**A**ppellation d'**O**rigine **P**rotégée
ArMV	**Ar**abic **M**osaic **V**irus
BLUP	**B**est **L**inear **U**nbiased **P**redictor
CV	**C**ross **V**alidation
DM	**D**owny **M**ildew
GxE	**G**enotype-by-**E**nvironment interaction
GxY	**G**enotype-by-**Y**ear interaction
GEBV	**G**enomic **E**stimated **B**reeding **V**alue
GFLV	**G**rapevine **F**an **L**eaf **V**irus
GBLUP	**G**enomic **B**est **L**inear **U**nbiased **P**rediction
GBS	**G**enotyping-**B**y-**S**equencing
GLOB	**G**enomic-**L**ike **O**mics-**B**ased
GLRV	**G**rapevine **L**eaf **R**oll associated **V**irus
GP	**G**enomic **P**rediction
GPGV	**G**rapevine **P**inot **G**ris **V**irus
GRGV	**G**rapevine **R**ed **G**lobe **V**irus
GRM	**G**enomic **R**elationship **M**atrix
GS	**G**enomic **S**election
LASSO	**L**east **A**bsolute **S**hrinkage and **S**election **O**perator
LD	**L**inkage **D**isequilibrium
MAS	**M**arker **A**ssisted **S**election
Mbp	**M**ega **b**ase **p**air
MET	**M**ulti-**E**nvironment **T**rial
MIM	**M**ultiple **I**nterval **M**apping
NIRS	**N**ear **I**nfra **R**ed **S**pectroscopy
OIV	**I**nternational **O**rganisation of **V**ine and wine
OLS	**O**rdinary **L**east **S**quare
PA	**P**redictive **A**bility
PCA	**P**rincipal **C**omponent **A**nalysis
PDO	**P**rotected **D**esignation of **O**rigin
PM	**P**owdery **M**ildew
PP	**P**henomic **P**rediction
PS	**P**henomic **S**election
QTL	**Q**uantitative **T**rait **L**ocus

RR	**R**idge **R**egression
SIM	**S**imple **I**nterval **M**apping
SNP	**S**ingle **N**ucleotide **P**olymorphism
SSR	**S**ingle **S**equence **R**epeats
TE	**T**able **E**ast
TS	**T**raining **S**et
UAV	**U**noccupied **A**erial **V**ehicle
VS	**V**alidation **S**et
WD	**W**ater **D**eficit
WE	**W**ine **E**ast
WW	**W**ine **W**est
WW	**W**ell-**W**atered

Chapter 1

Introduction

1.1 About grapevine

1.1.1 Economic importance of viticulture in the world and in France

Among horticultural crops, grapevine is one of the most economically valuable worldwide. With a total production of 14.3 million tons (Mt) out of the 77 Mt produced overall, spanned over 7.4 million hectares worldwide (https://www.oiv.int/en/statistiques/recherche, 2021). China was the first producing country in 2019, and France the fifth one with 5.5 Mt (http://www.fao.org/faostat/, 2019). However, a distinction needs to be made between table and wine grape productions. In 2018, Europe accounted for 37.8% of total world grape production, but for 62.9% of wine production (http://www.fao.org/faostat/, 2019). This reflects the fast growing of China's table grape production, while wine production remained stable across the last years in this country. France is the first country for grape production in terms of value, with $14.3 billion (Cantu & Walker, 2019; FranceAgriMer, 2019). In France in 2019, wine represented 31.2% of food industry exportation value (FranceAgriMer, 2019). In terms of surface area in 2019, 789,000 hectares were dedicated to grapevine in France, i.e., 3% of cultivated surface, and among this, 6,200 hectares were for table grape (FranceAgriMer, 2019).

1.1.2 Genetics and biology of cultivated grapevine

Phylogeny

The vast majority of cultivated grapevine belongs to *Vitis vinifera L.* species, from the *Vitaceae* botanical family and the *Vitis* genus, which encompasses around 70 species mainly in Asia and North America. In Eurasia, two *V. vinifera* sub-species still coexist: the cultivated one *V. vinifera* subsp. *vinifera* and the wild one *V. vinifera* subps. *sylvestris*. All the species from the *Vitis* genus, such as *V. rupestris*, *V. riparia* or *V. berlandieri* (in Figure 1.1), are inter-fertile with *V. vinifera* and are currently used in both scion and rootstock breeding programs for their genetic resistance to diseases.

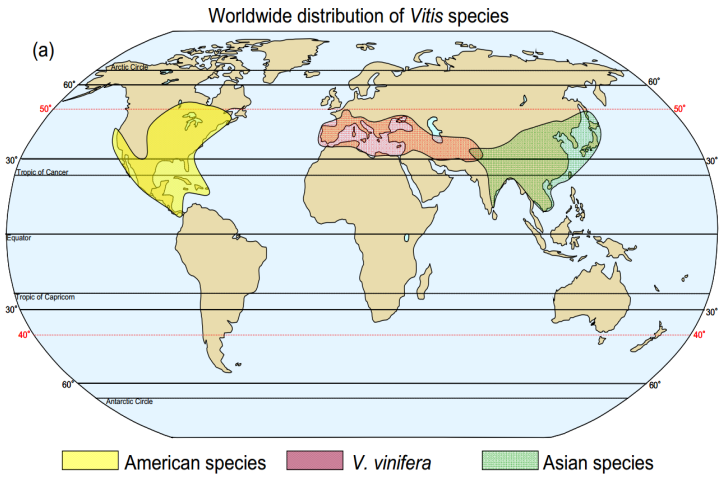


FIGURE 1.1 – Worldwide distribution of *Vitis* species (from Töpfer et al., 2010).

Domestication

Grapevine was first domesticated 8,000 to 10,000 years ago in the South Caucasus (Bacilieri et al., 2013; Péros et al., 2011) from *V. vinifera* subsp. *sylvestris*, with subsequent secondary domestication events around the Mediterranean Sea. Traits affected by domestication were sex, fertility, shoot architecture, leaf morphology, berry size and berry composition (This et al., 2006). Indeed, domestication led hermaphroditism to be the dominant form, compared to dioecious form in wild grapevine. Grapevine underwent differential secondary domestication events depending on local adaptation to environments and use of the grape. For direct fresh consumption (i.e., table grapes), a large bunch with big berries were preferentially selected, while for wine-making (i.e., wine grapes), higher sugar and anthocyanin content and higher skin / pulp ratio thus smaller berries was preferable (Liang et al., 2011; This et al., 2021).

Genetic diversity and structure

V. vinifera has a diploid genome with 2n=38 chromosomes, first sequenced in 2007, with a relatively small size of 487 Mbp, compared to other cultivated species (Jaillon et al., 2007; Laucou et al., 2011; Velasco et al., 2007). *V. vinifera* worldwide genetic diversity has been characterized notably in the French Vassal germplasm collection (Laucou et al., 2011; Laucou et al., 2018; Nicolas et al., 2016) and the American USDA germplasm collection (Myles, 2013; Myles et al., 2011). These studies underlined that *V. vinifera* is a highly heterozygous species with mean level of heterozygosity of 0.76, with strong inbreeding depression (Laucou et al., 2011; Reynolds, 2015).

Within *V. vinifera* subsp. *vinifera*, there is a low structuration according to both geographical origin (East / West) and grape use (wine / table), leading to three main genetic groups: WW (wine West), WE (wine East) and TE (table East) (Bacilieri et al., 2013; Nicolas et al., 2016

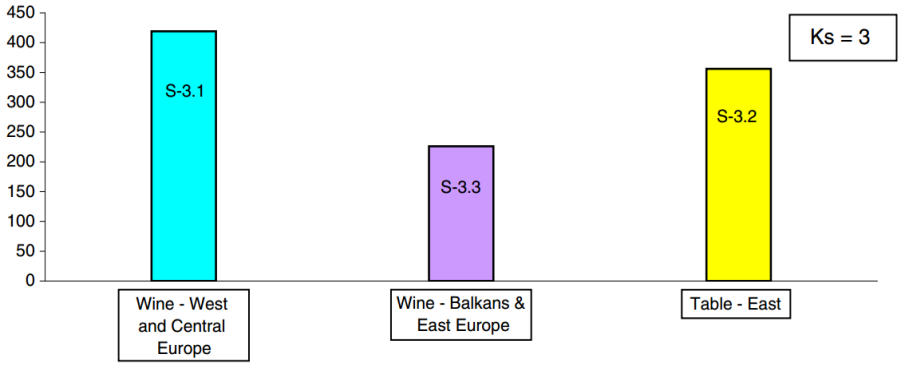


FIGURE 1.2 – Structuration of grapevine genetic diversity into three clusters, using STRUCTURE software (from Bacilieri et al., 2013).

and Figure 1.2). Further subdivision can be found within these groups (Bacilieri et al., 2013; Laucou et al., 2018).

V. vinifera is characterized by a large genetic diversity with 26.2 alleles on average, measured on 4,370 accessions with 20 SSRs (Laucou et al., 2011). *V. vinifera* shows rapid decay of linkage disequilibrium (LD, correlation between loci), with a distance between 9 to 458 Kb for a corrected r^2 between pairwise markers at 0.2 (approximatively half of the maximum r^2), however, large differences in this distance were observed between linkage groups, grapevine subpopulations and grapevine species (Barnaud et al., 2010; Barnaud et al., 2006; Flutre et al., 2020; Nicolas et al., 2016). Grapevine underwent a small genetic bottleneck due to domestication, nevertheless, genetic diversity is still probably weaker than before the phylloxera crisis, which destroyed most European vineyards and especially genotypes from the wild subspecies sylvestris. (Myles et al., 2011; This et al., 2006).

As cultivated grapevine could be multiplied through vegetative propagation, most of its genetic diversity has been conserved across centuries. However, due to market and consumer preferences, this genetic diversity remained underexploited, with only a few varieties being widely cultivated worldwide (Myles, 2013).

Grapevine biology and viticultural aspects

V. vinifera is a liana with woody shoots and fleshy fruits. However, domestication changed the shoot attitude of the vine to adapt to its culture. The scion is pruned every year, in order to counteract the continual shoot growth. Grapevine takes normally five to six years from fecundation to an offspring producing clusters, using grafting. This length can be reduced to three to four years by using a special pruning. However, wine-growers buy directly grafted plants that need around two to three years before the first harvest. Grapevine is a perennial species, but after 15 to 20 years the yield gradually decreases over the years, due to diseases or dieback. Thus, depending on the variety and the region, vine replacement normally occurs every 15 to 40 years. For example, in Occitanie (France), replacement occurs at 22 years on

average, from 17 years on average for Syrah to 42 years for Carignan (Tripiana & Mayoux, 2018).

1.1.3 Grapevine varieties and breeding

Grapevine historically cultivated varieties

History of grapevine breeding

— Ancestral grape cultivars
Most of the today's cultivated varieties arise from an original seedling from spontaneous crosses between grapevine cultivars (This et al., 2021). Then, vegetative propagation has been the main mean of diffusion of grapevine varieties since Roman ages (Mudge et al., 2009), essentially through cutting and layering. Indeed, using cuttings ensure the growers that grapes will have the same desirable traits with preserving the same genotype, whereas sowing seeds from sexual reproduction lead to highly diverse progenies. However, this did not exclude several crosses within cultivated grapes and/or rarely with wild grapes, followed by the occasional selection of some new improved varieties. In point of fact, many of the main cultivated varieties have existed for centuries, for example, Gamay and Chardonnay are from the Middle Age, Pinot Noir from Roman ages (Bowers et al., 1999; Lacombe et al., 2013), and Cabernet-Sauvignon is cultivated since the 17^{th} century in France (Bowers & Meredith, 1997).

— Phylloxera and mildew crises and inter-specific hybrids
At the end of the 19^{th} century, phylloxera, downy (DM) and powdery (PM) mildews were introduced in Europe from North America. The large majority of *V. vinifera* varieties were susceptible to these diseases. Accession of *Vitis* species were used as rootstock such as "*Rupestris du Lot*" or "*Riparia Gloire de Montpellier*" but crosses were also made between *Vitis vinifera* accessions and American *Vitis* species (such as *V. riparia*, *V. labrusca*, *V. rupestris*), which were resistant against phylloxera, and adapted to European soils. Latter, similar hybrids were developed in order to produce wines on their own roots (they were resistant against phylloxera, PM and DM) and were therefore called direct producer hybrids. They were grown until the middle of the 20^{th} century (This et al., 2006, Figure 1.3). However, for some of them, wine quality was poor (with a foxed flavor), with too high methanol concentration. Most direct producer hybrids were finally forbidden in France from 1934 on. This period saw the emergence of dedicated grapevine breeders and the creation of many hybrid grapevine varieties. Hybridization continued even after, with further backcrossing of F1 inter-specific hybrids with *V. vinifera* during several cycles, to achieve a proportion of *V. vinifera* genome close to 1 (the proportion of *V. vinifera* genome is theoretically doubled at each cycle). The other solution against phylloxera was to graft onto resistant rootstocks coming from American *Vitis* species, with traditional *V. vinifera* scions. This solution is still predominant today in the world.

Scion varieties

The *Vitis* international variety catalog registers 21,045 names of *V. vinifera* scion varieties, among which around 6,000 are distinct (OIV, 2017), yet 20 varieties represent 37% of cultivated area worldwide (Galet, 2000).

Variety	Colour	Destination	Area (in thousand ha)	Proportion
Kyoho	Black	Table	365	4.9
Cabernet-Sauvignon	Black	Wine	341	4.6
Sultanina	White	Table, drying and wine	273	3.7
Merlot	Black	Wine	266	3.6
Tempranillo	Black	Wine	231	3.1
Airen	White	Wine, Brandy	218	2.9
Chardonnay	White	Wine	210	2.8
Syrah	Black	Wine	190	2.5
Red Globe	Black	Table	159	2.1
Grenache Noir	Black	Wine	163	2.2
Sauvignon Blanc	White	Wine	123	1.7
Pinot Noir	Black	Wine	112	1.5
Ugni Blanc	White	Wine, Brandy	111	1.5

TABLE 1.1 – Distribution of the most cultivated grapevine varieties in the world. Source: OIV (2017).

Variety	Colour	Area (in thousand ha)	Proportion
Merlot	Black	112	13.9
Ugni Blanc	White	82	10.2
Grenache Noir	Black	81	10.0
Syrah	Black	64	7.9
Chardonnay	White	51	6.3
Cabernet Sauvignon	Black	48	6.0
Cabernet Franc	Black	33	4.1
Carignan Noir	Black	33	4.1
Pinot Noir	Black	32	4.0
Sauvignon Blanc	White	30	3.7
Other varieties		240	29.8
Total		806	100

TABLE 1.2 – Distribution of the most cultivated grapevine varieties in France. Source: OIV (2017).

— At the world level

At the world scale, the most widely cultivated grapevine variety (in terms of total grape growing area) is Kyoho, a tetraploid table grape, mostly grown in Asia; it represents 44% of the total growing area surface in China. Then come in order Cabernet-Sauvignon (wine grape), Sultanina (table grape and raisin) and Merlot (wine grape) (OIV, 2017 and Table 1.1).

— In France
In France, most surfaces (>99%) dedicated to grapevine are for wine production. The most widely cultivated varieties are Merlot, Ugni Blanc, Grenache Noir, Syrah and Chardonnay (OIV, 2017 and Table 1.2). In France, the Protected Designation of Origin (PDO) imposes a restricted set of varieties that can be grown in a given wine-growing area. More than half of the surface area in France is under a PDO (62%) (FranceAgriMer, 2019). The Occitanie region, with 263,000 hectares, is the biggest grape-growing area of France and the first vineyard in the world in terms of production under PDO (https://www.vignevin-occitanie.com).

Rootstock varieties

During the phylloxera crisis, the introduction of other *Vitis* species as rootstocks rescued the European vineyard. Main cultivated rootstocks are from other *Vitis* species or simple inter-specific hybrids involving species such as *V. rupestris*, *V. riparia*, *V. berlandieri* and/or *V. vinifera*. Rootstock varieties were mainly bred during the 19^{th} and early 20^{th} centuries yielding from 70 to 80 varieties available nowadays (Ollat et al., 2016). However, rootstock use does not only allow phylloxera resistance, but also confers resistance or tolerance to other diseases or abiotic stresses, adaptation to soils characteristics and have an impact on berry composition and yield (Migicovsky et al., 2021; Ollat et al., 2016). In France, the most widely grown rootstock varieties are Selection Oppenheim 4 (SO4), a *V. berlandieri* x *V. riparia* cross; 110 Richter, a *V. berlandieri* x *V. rupestris* cross; 3309 Couderc, a *V. riparia* x *V. rupestris* cross and Fercal, a cross between two inter-specific hybrids (Ollat et al., 2016, plantgrape.plantnet-project.org/en/).

Clonal selection

A special feature of grapevine is that plant selection is made also within the variety level, by selecting somatic mutations affecting phenotypic traits. Historically, clonal selection has been initiated in order to select clones without fanleaf virus, especially in the North-East of France and in Germany (Reynolds, 2015). Traits selected may be qualitative, such as berry color (e.g., Pinot Blanc and Pinot Gris are clones derived from Pinot Noir), or quantitative (e.g., phenology, titratable acidity, *Botrytis* tolerance and berry size (Reynolds, 2015)). Some well-known varieties such as Chardonnay display high phenotypic diversity among clones (Roach et al., 2018). For PDO growing areas, clonal selection is sometimes the only way to improve grapevine, because PDO, and more particularly the *Appellation d'Origine Protégée* (AOP) in France and Europe, imposes a set of varieties that can be grown in each region.

Recent French grapevine breeding programs

Several countries worldwide have developed breeding programs, but we will focus our review on French wine grape breeding programs for disease resistance over the past fifty years.

Known resistance genes against grape diseases

Powdery and downy mildews are the main fungal diseases of grape, receiving most the attention of grape breeders. To date, 14 resistance genes/QTLs against powdery mildew (PM) have been identified in other species than *V. vinifera* (Cantu and Walker, 2019, www.vivc.de), referred to as *Run* or *Ren* for Resistance to *Uncinula* or *Erysiphe necator*, respectively. *Run1* was the first locus characterized for this resistance, in a progeny between *Muscadinia rotundifolia* and *V. vinifera* (Bouquet, 1986; Pauquet et al., 2001). Comparatively, the first locus for resistance to downy mildew (DM), *Rpv1* for resistance to *Plasmopara viticola*, was also characterized in a progeny between *Muscadinia rotundifolia* and *V. vinifera* (Merdinoglu et al., 2003). Other major loci presently exploited in grape breeding are *Run1*, *Ren1*, *Ren3* and *Ren9* resistance genes. For DM, 32 resistance genes/QTLs are known (Bhattarai et al., 2020; Cantu & Walker, 2019; Sargolzaei et al., 2020). Breeders mostly work with *Rpv1*, *Rpv3*, *Rpv8*, *Rpv10* and *Rpv12* resistance genes. These PM and DM resistance genes confer major or minor / partial resistance. Other resistance genes are known against *Botrytis bunch rot*, *non-Botrytis bunch rot*, *grapevine fan leaf virus*, *black rot*, *phylloxera*, or *Xylella fastidiosa* (Krivanek et al., 2006; Rex et al., 2014; Xu et al., 2008; Zhang et al., 2009).

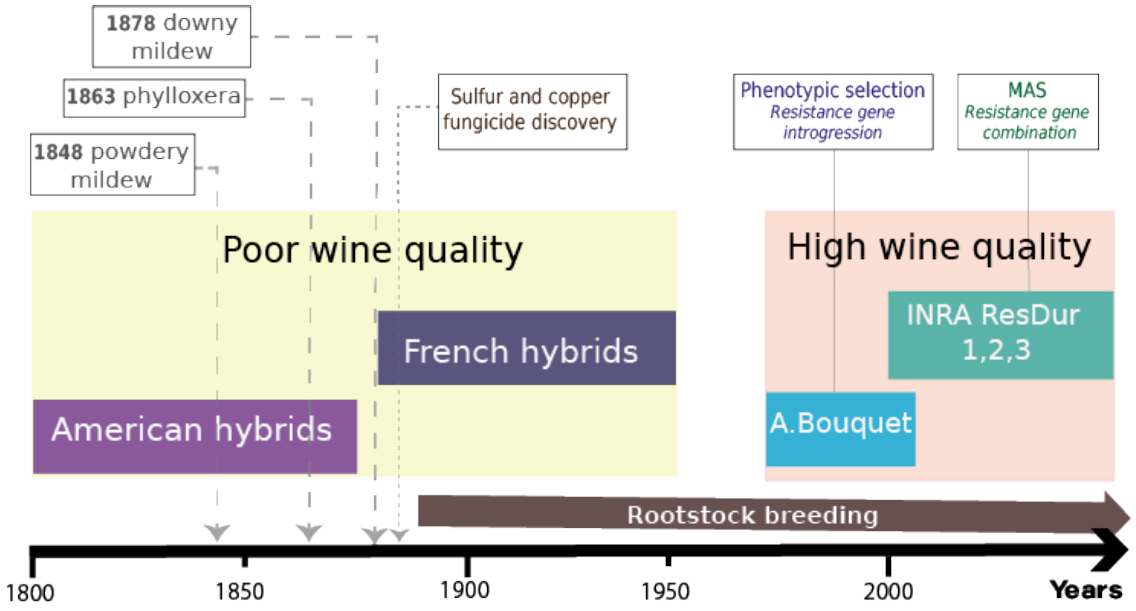


FIGURE 1.3 – Chronology of French grapevine breeding programs, adapted from Töpfer et al. (2010), MAS: marker-assisted selection.

Bouquet's program

Alain Bouquet, researcher and grape breeder at INRA, started a breeding program around 1970 to introgress genetic resistance to PM, by crossing *V. vinifera* with a F1 cultivar from a cross between *V. vinifera* and a resistant cultivar from *M.rotundifolia* (Bouquet et al., 2000). This cultivar was also bearing a DM resistance gene and these loci (*Run1* and *Rpv1*) were genetically linked. Then, progeny resistance was assessed through phenotypic evaluation, under severe disease pressure (Bouquet et al., 1981, Figure 1.3). Several successive crosses

were needed to recover a large proportion of *V. vinifera* genetic background, with phenotypes close to those of cultivated grapevines. As backcrossing to a single cultivar is not possible in grapevine due to strong inbreeding depression, recurrent cultivar had to be changed at each generation (Pauquet et al., 2001). Selection was not only focused on PM and DM resistance, but also targeted other quantitative traits, through the choice of *V. vinifera* parents. However, in each cycle, several years were needed to apply phenotypic selection. Finally, five pseudo-backcrosses were performed by A. Bouquet after the initial F1, yielding an expected proportion of 98.4% of *V. vinifera* genome; this program lasted 25 years. Seven wine varieties from this program have a temporary planting authorization in France and four of them will be used for crossing with Ugni Blanc for Cognac production (Pech-Rouge, 2017). They also have been used as parents for many breeding programs throughout Europe and US.

INRA-ResDur program

Total resistance conferred by a single gene for PM or DM, as obtained from Bouquet's program, could rapidly be circumvented (Peressotti et al., 2010). Along with the development of molecular markers, several resistance loci were progressively tagged (Pauquet et al., 2001), enabling Marker Assisted Selection (MAS) and thus making it possible to early determine if genotypes were bearing specific resistance genes, for total or partial resistance. Therefore, breeders could now select genotypes cumulating several resistance genes, for both PM and DM (Eibach et al., 2007; Merdinoglu et al., 2018).

The INRA-ResDur program started in 2000 at INRA, in collaboration with Julius Kuhn Institute (JKI, Germany), Agroscope (Switzerland), and Weinbau Institute Freiburg (WBI, Germany) (Schneider et al., 2019a; Schneider et al., 2019b and Figure 1.4). Starting from the varieties developed by Bouquet (4*th* backcross), the INRA-Resdur program aimed at selecting new grapevine varieties cumulating two (ResDur 1 and 2) or three (ResDur 3) resistance genes, for both PM and DM, in order to ensure long-lasting resistance (Figure 1.4). Presence of several resistance genes was assessed through MAS, and other traits were phenotypically selected (Schneider et al., 2019a). Official registration of ResDur 3 varieties is scheduled in 2025 and four ResDur1 varieties are already registered in the French catalogue: Artaban, Vidoc, Floreal and Voltis (https://observatoire-cepages-resistants.fr/, Schneider et al., 2019b).

ResDur2 and ResDur3 varieties are currently studied for their wine aptitude with different wine-making processes (https://observatoire-cepages-resistants.fr/, Guimier et al., 2019; Salmon et al., 2018), and tested for their tolerance to abiotic stresses at INRAE Pech-Rouge (Occitanie).

Current programs

As varieties with several resistance genes are currently available, the goal is now to cross ResDur2 varieties with emblematic varieties to each wine-growing region (Figure 1.4). For example, the regional Martell and EDGARR programs crossed their locally grown varieties (namely Monbadon, Montils and Vidal 36 for Martell; Vermentino and Cinsault for EDGARR) with resistant genotypes cumulating several resistance genes. The objective is to get resistant

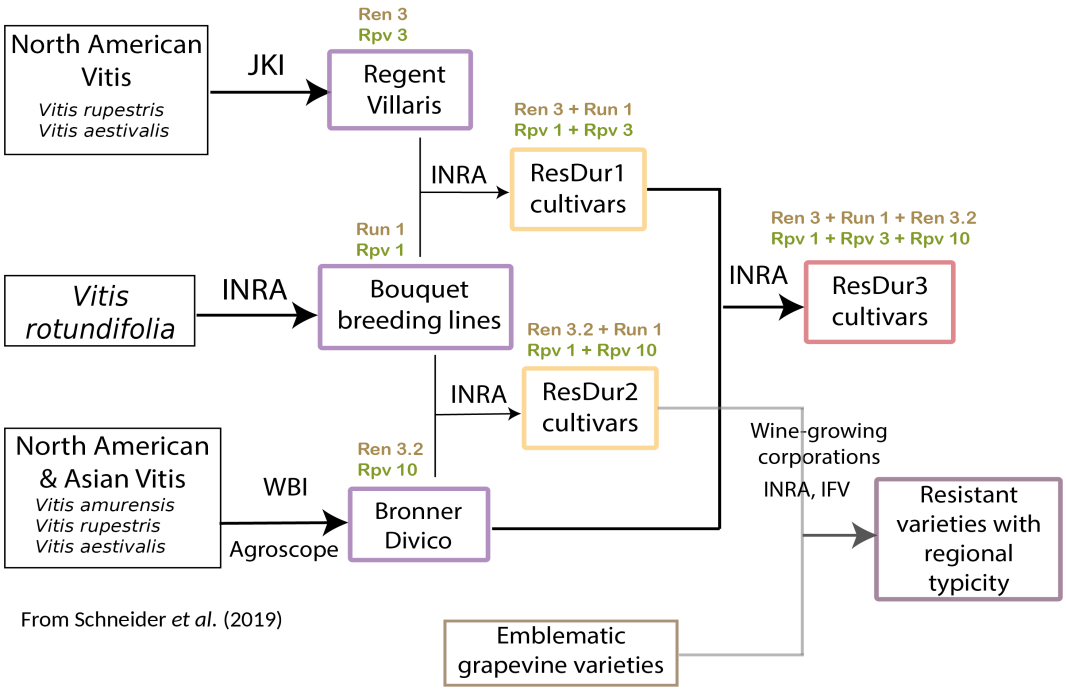


FIGURE 1.4 – INRA-ResDur and French grapevine breeding programs. JKI: Julius Kuhn Institute (Germany); INRA: Institut National de la Recherche Agronomique (France), WBI: Weinbau Institute Freiburg (Germany), Agroscope (Switzerland). Powdery and downy mildew resistance loci tracked by Marker Assisted Selection are given in green and brown, respectively.

varieties that are phenotypically close to emblematic cultivars. These programs are the first one in grapevine involving genomic selection (GS) for predicting quantitative traits of interest (see below). Offspring with less than two resistance genes for each of PM and DM, as sorted with MAS, are phenotyped to constitute the training set (TS) to calibrate the model (Figure 1.5). Then, the phenotypes of offspring bearing four resistant genes are predicted with this model.

However, in these programs, GS is mainly useful to decrease phenotyping effort, not to save time by avoiding phenotyping. Another regional program, from Inter-Rhône, will be able to make full use of GS by using an existing and widely studied bi-parental Syrah x Grenache cross (Coupel-Ledru et al., 2016; Coupel-Ledru et al., 2014; Doligez et al., 2006; Doligez et al., 2013; Fournier-Level et al., 2009; Fournier-Level et al., 2011; Huang et al., 2013; Huang et al., 2012; Huang et al., 2014) as TS, and crosses between Grenache and Syrah with resistant genotypes as target population.

Overall, the use of GS in grapevine breeding programs allows to skip the 2*nd* step of breeding, that is the phenotyping step, which occurs after having selected resistant genotypes with MAS. This step lasts around six years, which could be spared thanks to GS in grapevine breeding programs (Figure 1.6).

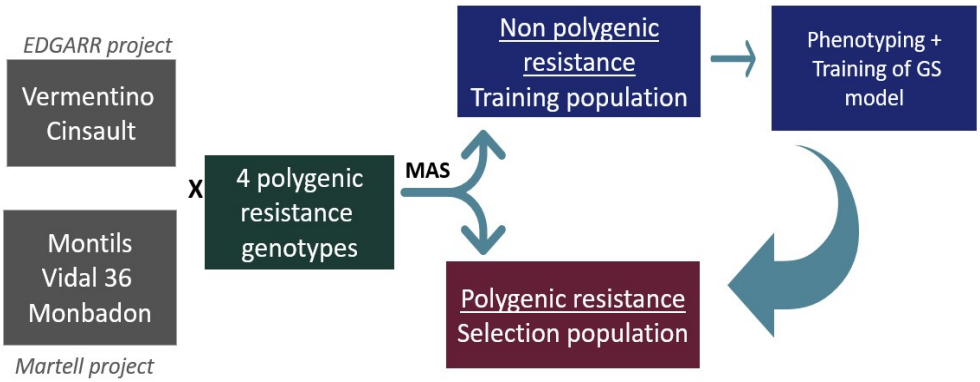


FIGURE 1.5 – Overlook of current EDGARR and Martell breeding programs (pers. comm. L. Le Cunff). MAS: marker assisted selection.

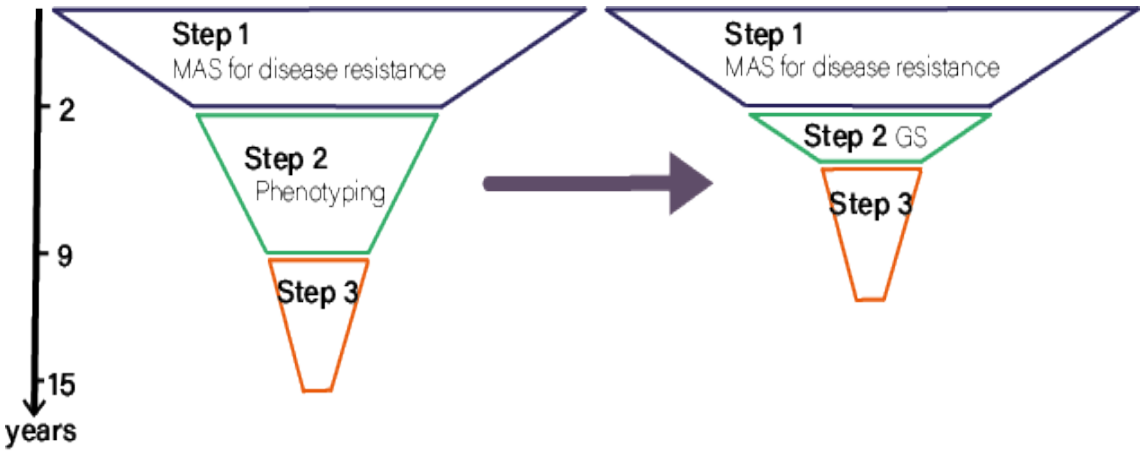


FIGURE 1.6 – Grapevine breeding steps without and with GS. MAS: marker assisted selection. Width of shape indicates number of genotypes under selection. Step 3 is the registration step.

1.1.4 Grapevine ideotype

The ideotype is the theoretical ideal variety that combines breeding objectives for all traits. For many crops, the ideotype is related to yield, composition of the final product (such as protein content in wheat grains) (Michel et al., 2019); and tolerance to biotic and abiotic stresses. Grapevine ideotype is not so easy to define. Few of the main reasons are the fact that the main final product is a transformed product (wine), the complexity of wine itself and the diversity of wines produced. However, grapevine, as others crops, is facing great challenges: reduced inputs in vineyards, in vinification process and foreseen climate changes. Thus, resistance to diseases such as PM, DM, Pierce's disease, phylloxera or grapevine fan leaf virus is crucial. Resistant varieties to PM and DM are already available to winegrowers. However, abiotic stress tolerance raises more and more concerns, especially with climate change. Winegrowers are looking for cultivars with a stable production across years, drought tolerance, with a synchronization between technological and physiological maturity dates (see below).

Moreover, besides biotic and abiotic stresses tolerance, wine quality is crucial in grapevine breeding, but wine quality itself is not clearly defined. The final product of grapevine is obtained by fermentation, thus genetic selection for wine quality needs to take into account interaction with fermentation practices. Moreover, wine quality being a subtle mix of different metabolites and flavors, breeding objective for wine quality will never be to maximize the accumulation of a given compound.

Grapevine diseases

Grapevine (*Vitis vinifera L.*) is susceptible to many diseases, caused by fungi, viruses or bacteria. The use of pesticides constitutes a risk for grape growers' and neighborhoods health and an environmental threat.

Fungal diseases

Grapevine is susceptible to several fungi or fungus-like eukaryotic microorganisms. As presented before, the most important diseases are powdery mildew (caused by *Erysiphe necator*) and downy mildew (caused by *Plasmopara viticola*). Cost for the management of these diseases is estimated on average at 979 €/ha in South of France (Ojeda et al., 2010), and pesticide use is less and less socially accepted because of the damage on the environment and health. Natural resistances are known and used into the breeding programs. PM, DM and Black rot (*Guignardia bidwellii*), concentrate most of the fungicide treatments in France. These three pathogens were imported from North America through infected plants.

PM was first reported in France in 1848 and DM in 1878. PM and DM affect leaves and berries and can cause substantial yield losses. Almost all *Vitis vinifera* traditional grapevine varieties are susceptible to both PM and DM, although partial genetic resistance to PM and DM was observed (Coleman et al., 2009; Riaz et al., 2013; Sargolzaei et al., 2020). Sulfur and copper are the most common organic pesticides applied against these fungi, but these treatments only prevent from the fungi multiplication, thus they are not curative treatments.

Other fungal diseases affect the woody part of the plant, they are called grapevine trunk diseases (GTD), and are caused by several fungi (Mondello et al., 2018). The three major GTDs are Esca, Botryosphaeria dieback, and Eutypa dieback (Claverie et al., 2020). Symptoms of GTDs are not yet fully understood, wood symptoms are not easily observable until apoplexy and foliar symptoms are not always correlated with disease aggressiveness. However, its incidence reached 11% in France in 2008 (Grosman & Doublet, 2012). Before 2001, sodium arsenate was used as curative treatment against GTDs but it has been forbidden because of its toxicity. Since then, there is no curative treatment available.

Insect diseases

In Europe, viticulture has also suffered from phylloxera (*Daktulosphaira vitifoliae*), a small insect (aphid) which feeds on grapevine roots and leaves. This insect was accidentally introduced in Britain from America around 1850 and spread all over Europe, destroying most

vineyards. In France, wine production declined from 54 MhL in 1865 to 29 MhL in 1885 (Stevenson, 1980). There is no chemical treatment, nor genetic resistance to phylloxera in *Vitis vinifera* genetic background. However, other *Vitis* species from America co-evolved with phylloxera and thus display resistance. Nowadays, most planted vines are grafted on resistant rootstocks, except those grown in sandy soils or in countries where phylloxera is absent, as in South America. Several phytophagous mite species are jeopardizing grapevine, among which the Tetranychidae and the Eriophyidae families (https://ephytia.inra.fr). Some insects are also attacking fruiting organs, such as cochylis (*Eupoecilia ambiguella*) and eudemis (*Lobesia botrana*). Damages from these species can be severe, and in addition other rots can develop, such as gray rot (*Botrytis cinerea*) (https://ephytia.inra.fr).

Other insects also indirectly cause diseases, by transmitting deleterious bacteria, such as the blue-green sharpshooter, which transmits the Pierce's disease (see below), or mealybugs, which are phytophagous and may transmit grapevine leaf roll virus (http://ephytia.inra.fr).

Bacterial diseases

Pierce's disease is caused by the bacterium *Xyllela fastidiosa* and transmitted by sharphopper, leafhopper or cicadas for example. This bacterium is not specific to grapevine and can infect for example citrus like in Brazil at the end of the eighties (Chang et al., 1993; Hartung et al., 1994) or olive trees like in Italia in Apulia region since 2013 (Saponari et al., 2013), but some strains also affect grapevine by blocking the xylem flow, leading to vine death. Pierce's disease caused major damages in the USA at the end of the 19th century. The strain that affects grapevine (subspecies *fastidiosa*) has been identified in Balearic Island in Spain but not elsewhere in Europe for now (Olmo et al., 2017).

Flavescence dorée is a phytoplasma disease transmitted by a leafhopper. The only cure against this disease is insecticide to eradicate the disease vector. In France, the uprooting of the contaminated vines is compulsory and if more than 20% of the vineyard surface is contaminated, the entire vineyard parcel must be uprooted.

Viral diseases

In grapevines (*Vitis* and *Muscadinia* sp.), about 70 viruses have been identified (Meng et al., 2017). The two major diseases caused to *V. vinifera* by these viruses are leafroll and fanleaf (Fuchs et al., 2017). Fanleaf is associated with 16 Nepovirus, especially GFLV (grapevine fan leaf virus) and ArMV (Arabic Mosaic Virus). Vectors of fanleaf viruses are nematodes and transmission in the field is done from one plant to the next. Grapevine leafroll is mostly associated with two viruses GLRV1 and GLRa-V3 (Grapevine Leaf Roll associated Virus 1-3). Other damaging viruses are *grapevine Pinot gris virus* (GPGV), a trichovirus which causes chlorotic mottling and leaf deformations, this virus is transmitted by a mite, *Colomerus vitis* and *grapevine red globe virus* (GRGV), a maculavirus which causes fleck on leaves (Fuchs et al., 2017). Symptoms include loss in yield, delay and heterogeneity in maturity. Only plants free of these viruses can be sold by nurseries.

Grapevine response to abiotic stresses

Worldwide, grapevine is mostly cultivated in temperate climate regions, between latitudes 4 and 51°C in Northern hemisphere and between 6 and 45°C in the Southern hemisphere (Santos et al., 2020). In Mediterranean areas, viticulture is sometimes the only agricultural production possible. Due to climate change, it is very *likely* that, in Mediterranean areas, temperature will be higher in summer, while precipitations will decrease (IPCC et al., 2021). Many factors related to climate change are likely to affect grapevine growth, such as temperature, water availability, light with photosynthetically active radiation (PAR), C02 concentration, etc (IPCC et al., 2021).

These abiotic factors impact grapevine phenology and yield but also berry composition and wine-making (Duchêne, 2016; Naulleau et al., 2021; Santos et al., 2020 and see below). In particular, there is a constant increase of sugar and a decrease of malic acid in harvested berries, yielding to higher alcohol degrees and less acidity in current wines (Mira de Orduña, 2010; Ruffner et al., 1976).

Several solutions concerning the crop system are currently under investigation to mitigate the susceptibility of grapevine to climate change effects (Naulleau et al., 2021; Santos et al., 2020; van Leeuwen & Destrac-Irvine, 2017). Moreover, there is a large genetic variability in grapevine response to abiotic stresses (Morales-Castilla et al., 2020), that is not yet fully used in breeding, mostly due to the PDO system which imposes a set of authorized varieties.

Grapevine breeding is a potential response to mitigate climate change consequences. Other solutions include changes in planting area, cropping system and/or winemaking processes. The change in planting area is not preferred by viticulturists in France, probably due to the rigid PDO/AOC system. Winemaking adaptation might not be sufficient to mitigate climate change effects. Thus, developing new adapted cultivars appears to be an attractive solution, in combination with previously mentioned ones.

Phenology

The first visible consequence of climate change in plants is phenology (Menzel et al., 2020). In grapevine the advancement of the onset of ripening (called véraison in French, this stage is linked to a metabolism change in berry which induces berry softening and anthocyanin synthesis for colored genotypes) and harvest dates over years is well documented and there is a clear shift. For example, during the 1950-2010 period, average harvest date in Chateauneuf-du-Pape (Rhône valley) advanced from 26^{th} to 10^{th} September (de Cortázar-Atauri et al., 2017; Mira de Orduña, 2010). In 2050, phenology stages are expected to occur 6 to 12 days earlier than today in France, no matter the climatic scenario, and up to 30 days by the end of the century, depending on the climatic scenario (de Cortázar-Atauri et al., 2017; Duchêne et al., 2010). This trend is not specific to France or Europe, in New-Zealand, maturity date is expected to occur 6 to 12 days earlier than today in 2050 on the four main cultivated varieties (Ausseil et al., 2021). Models have been developed in order to predict phenology stages according to weather conditions (Parker et al., 2011; Parker et al., 2020). This shift in phenology

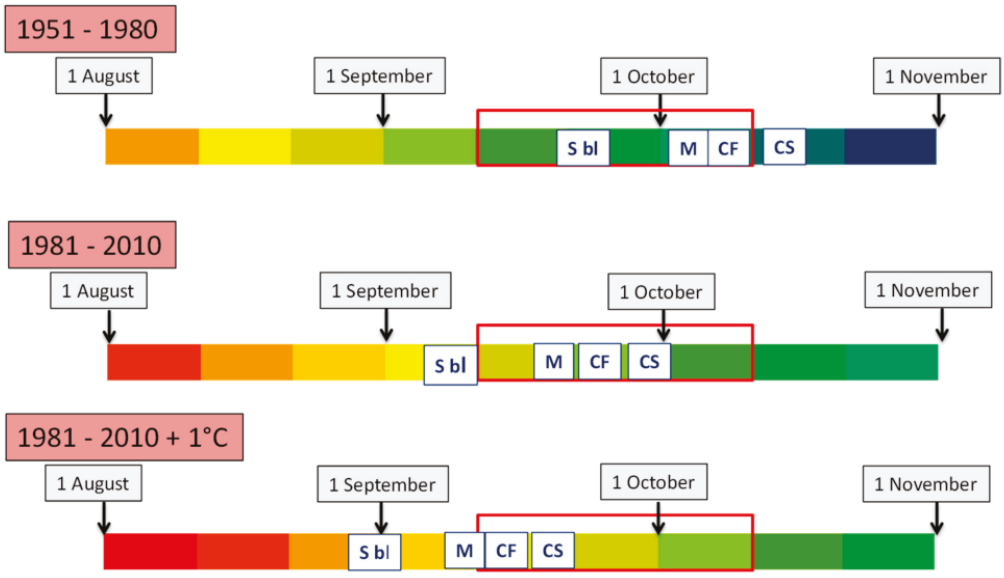


FIGURE 1.7 – "Modeled harvest dates for Sauvignon blanc (S bl), Merlot (M), Cabernet franc (CF), and Cabernet-Sauvignon (CS) in Bordeaux for the following periods: 1951–1980, 1981–2010, and 1981–2010+ 1 °C. Sugar ripeness is modeled with the grapevine Sugar Ripeness Model (GSR; Parker et al., 2019). Temperature data is from Bordeaux Mérignac weather station. Warm colors indicate higher temperatures and cold colors cooler temperatures". The scenario '1981-2010+1°C' corresponds to temperature projections for 2050 (from van Leeuwen et al., 2019).

dates implies that ripening happens under warmer weather conditions of summer with a direct impact on accumulation of aroma precursors and on wine color. Moreover, at any given date and due to climate change, temperatures have been increasing, including during berry development and ripening (van Leeuwen et al., 2019 and Figure 1.7). This shift impacts both water availability and berry composition (see below).

Genetic variability and determinism of phenological stages have been reported (Duchêne et al., 2012; Grzeskowiak et al., 2013; Parker et al., 2013; Vezzulli et al., 2019), but not yet their response to abiotic stresses.

In South areas, grapevine ideotype for drought tolerance would be to select varieties with an early maturity date, in order to avoid summer water stress for the plant. However, when considering berry development and composition, varieties with late maturity date would be preferred.

Water availability

Grapevine is already adapted to moderate drought stress levels typically encountered in Mediterranean regions. Moderate water deficit could be even beneficial for wine quality without decreasing yield. However, with climate change, water is expected to become scarcer

in some areas such as Southern France (IPCC et al., 2021). Thus, water deficit and the associated higher evaporative demand are expected to be a major threat for future viticulture.

Water deficit has several consequences at the plant level. Soil drying leads to a drop in plant water potential (potential energy of water, linked to the availability of water in a given area), thereby increasing the risk of cavitation (gaseous bubbles in the xylem which threaten plant survival by blocking water flow). Excessive drops may have catastrophic consequences for the plant, hence they have developed various adaptations to prevent them (Simonneau et al., 2017). Under water shortage, plants are closing stomata, in order to avoid water loss.

Both rootstock and scion are involved in the response to water stress through specific physiological adaptations, and it was shown that some genetic variability exists for water saving under soil drying conditions, through daytime and nighttime transpiration for example (Coupel-Ledru et al., 2016; Coupel-Ledru et al., 2014; Lucini et al., 2020; Marguerit et al., 2012). Most water losses are controlled by stomatal regulation, with a tradeoff between transpiration and photosynthesis (Simonneau et al., 2017). Some species or genotypes within a species have been observed to maintain higher leaf water potential compared to others when submitted to the same edaphic water deficit (Tardieu & Simonneau, 1998). Such genotypes are called isohydric, such as Grenache, while others called anisohydric cannot prevent their leaf water potential from dropping, such as Syrah or Chardonnay (Pou et al., 2012; Schultz, 2003). Although this classification is a bit Manichean, plant response to water stress is between these following extremes (Gambetta et al., 2020; Hochberg et al., 2018). Under mild water shortage, anisohydric behavior will maintain photosynthesis, thus leading to maintained yield and more vigorous plants, while isohydric behavior will have a reduced photosynthesis. However, under severe or long-lasting drought, anisohydric behavior is prone to severe damages due to dehydration, while isohydric behavior will be more resistant to such a climatic scenario (Pou et al., 2012). These two described behaviors are varying across season, drought characteristics and crop load (Lauri et al., 2016; Naor et al., 2013; Sade et al., 2012). The ideotype for water availability cannot be easily determined and depends on climatic scenario and risk management (Coupel-Ledru, 2015; Tardieu et al., 2018).

Berry composition

During ripening (Figure 1.9), phloem unloading of photo-assimilates (mostly sucrose) in the pericarp dramatically increases. The vacuolar accumulation of hexoses is intimately linked to the breakdown of malic acid at the beginning of ripening, then to sugar respiration (Savoi et al., 2021; Shahood et al., 2020). These authors have recently identified a set of sugar transporters and water channels transcripts definitively blocked at phloem arrest. Whatever, sugar concentration continues to increase following the arrest of phloem, due to the decrease in volume subsequent to evaporation. Consequently, the effects of temperature largely differ, depending on water availability, acting both at the photosynthetic level in the canopy, and at berry growth level.

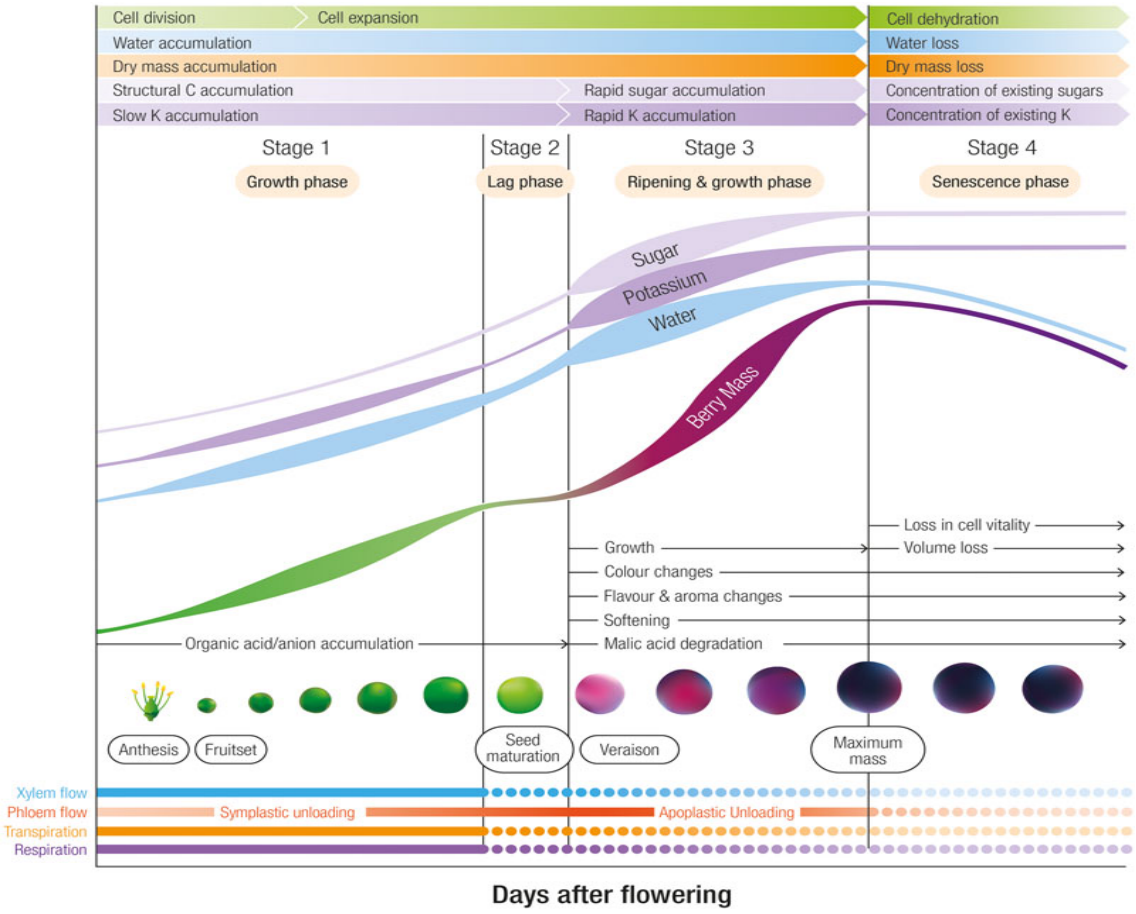


FIGURE 1.8 – Developmental stages of grape berries and their compounds (from Rogiers et al., 2017).

Comparatively to sugar, anthocyanin accumulation occurs at a slower rate during ripening at high temperatures. As a consequence, harvest is delayed to allow a higher concentration of anthocyanin in wine, leading to a higher sugar concentration (Sadras & Moran, 2012).

Decrease in malic acid concentration in turn increases wine pH, leading to shorter wine aging, while increased sugar content induces higher alcohol degree in wines. Aromas are also changed by higher temperature during ripening, with less herbaceous or vegetative notes in wines (Pons et al., 2017). On the other hand, moderate water deficit increases secondary metabolite accumulation in berries with more volatile thiol precursors, while strong water deficit decreases their synthesis (Pons et al., 2017). Ollat et al. (2018) showed that "Wines of climate change" were depreciated compared to current wines, based on tasting sessions.

Grapevine varieties can be selected to counteract these changes in wine composition. Objectives are to find varieties with more malic acid and anthocyanin and less sugar contents in berries (Bigard et al., 2018).

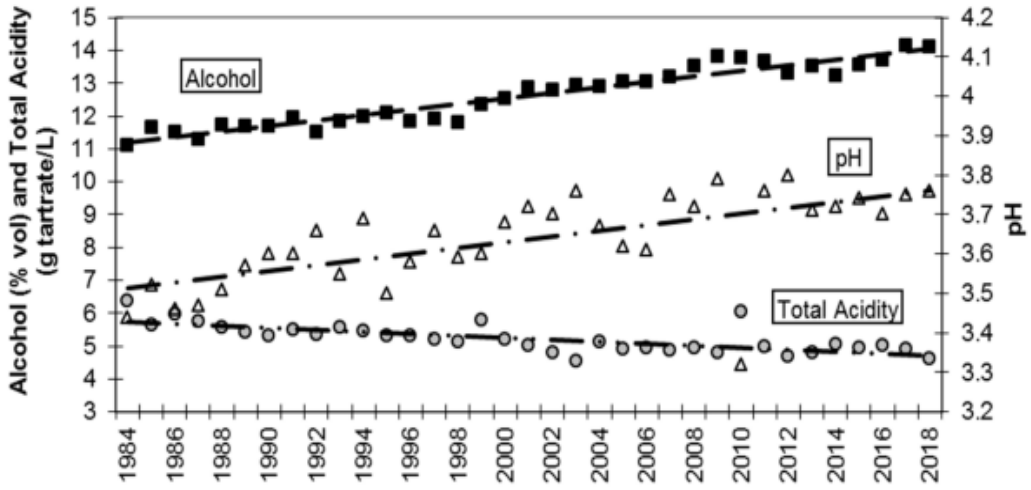


FIGURE 1.9 – 'Evolution of red wine composition in the Languedoc region (France) from 1984 to 2018. Each data point is the average of several thousands of analyses of red wines just after alcoholic fermentation (data: Dubernet laboratory, F-11100 Montredon des Corbières)' (from van Leeuwen et al., 2019).

Extreme climatic events

In addition to water stress, grapevine can be hit by extreme climatic events, such as heatwaves, hail or spring frost. These events will occur with higher frequency in the future (IPCC et al., 2021), but it is hard to conceive a grapevine variety resistant to all these events, e.g., severe freeze during burst, severe drought during summer, or heat burn. However, genetic variability and QTLs have been found for heat burn (A. Coupel-Ledru, pers. comm.)

Measure of drought-related traits

Drought tolerance is a highly complex trait affecting plant growth, berry composition and wine quality. Most of the time, traits related to drought tolerance are hard to phenotype, depending on environmental conditions and on cropping system (Tardieu, 2012). Traits linked to transpiration can be measured precisely in a greenhouse under semi-controlled conditions, but these results often cannot be transposed to the field. Indirect traits such as yield stability across years or production differential under water stress might be used to assess the drought tolerance level of genotypes. $\delta 13C$ is another indirect measure of water stress suffered by grapevine during the growing season. It refers to the relative amount of isotope 13 of carbon (over carbon 12) found in plant tissue (generally must or leaves in grapevine). $\delta 13C$ is related to intrinsic WUE, which itself is related to photosynthetic activity and stomatal conductance (Brendel et al., 2002; Farquhar & Richards, 1984).

1.2 Genotypic value prediction

1.2.1 Genomic prediction

Concept of genomic prediction

Before the advent of genomic prediction (GP), molecular markers were used to identify genetic polymorphisms associated with phenotypic variation through Quantitative Trait Loci (QTL) detection and mapping. These QTLs could then be used either to provide biological knowledge and insights into the underlying candidate genes or to select genotypes with favorable alleles without the need for phenotyping in a process called marker assisted selection (MAS). Through several generations, MAS essentially consists in following the allele to be introgressed or combined. MAS has been successfully applied in many plant species, such as for example for earliness and grain yield in maize (Bouchez et al., 2002), as well as for grapevine for disease resistance (see above) and sex of the flower. However, MAS becomes rapidly unfeasible when the number of QTLs to follow grows: with n QTLs, the probability to get a genotype harboring all favorable alleles is $1/2^n$. Moreover, many traits of interest are quantitative and thus typically controlled by many minor QTLs (Meuwissen et al., 2016). In addition, marker effects estimation from QTL detection studies is often upward biased and dependent on the study, the population (Beavis et al., 1994; Meuwissen et al., 2016; Xu, 2003), and only markers above a significance threshold are retained, many of which explain less than 10% of total genetic variation.

Genomic selection (GS), which consists in selection based on GP, has been proposed as a solution to this problem, since all marker effects are estimated without any significance threshold. It was first conceptualized by Meuwissen et al. (2001), who implemented several GP methods on simulated data. With the rapid increase of low-cost dense genotyping, the hypothesis that each causal gene or QTL is in linkage disequilibrium with a molecular marker tends to hold, making GP possible in many species of interest.

Implementation of GP in breeding programs requires that some genotypes will be both genotyped and phenotyped to constitute a reference population used to train the prediction model (training set, TS), while phenotypes of the target population (genomic estimated breeding values, GEBVs) will be predicted with the model, based on genotypic data only (Figure 1.10)

Statistically, the training step consists in fitting the following model: $y = X\beta + \epsilon$,

where y is the $n \times 1$ vector of genotypic values for the trait of interest over all n individuals of the TS, X is the $n \times p$ genotypic matrix for p markers, β is the $p \times 1$ vector of associated random marker effects, ϵ is the $n \times 1$ vector of residuals.

Marker effects are estimated in the TS with one of several possible methods (see below), then these effects are used along with the validation set (VS) genotypic matrix to predict genotypic values of the trait of interest in the VS: $\tilde{y} = X\hat{\beta}$.

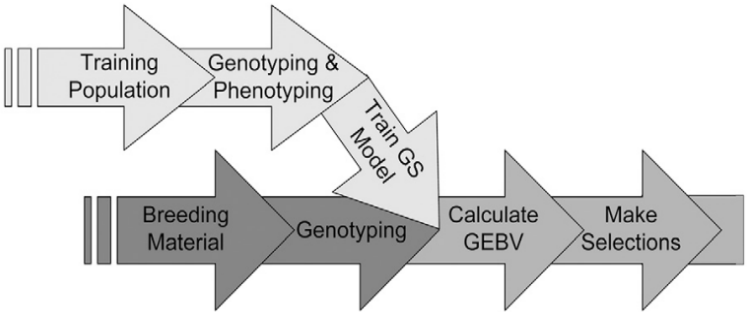


FIGURE 1.10 – Scheme of genomic selection implementation in breeding program. GEBV: genomic estimated breeding value (from Heffner et al., 2009).

Implementation of GP

Univariate vs multivariate models

The statistical model presented above is a univariate one, i.e., each trait is analyzed separately, but breeders and geneticists often need to handle several traits at a time. Most of the time, phenotypic correlations arise among studied traits. Phenotypic correlations are due to both environmental correlations (e.g., resource availability affect both yield and abiotic stress tolerance) and genetic correlations arising from pleiotropy (action of one locus on several traits) and LD.

Multi-trait, or multivariate, GP models allow to take into account such correlations by handling several traits jointly. The statistical model becomes: $Y = XB + E$, where $Y(n \times t)$, $B(p \times t)$ and $E(n \times t)$ are now matrices instead of vectors, adapted to t traits.

$B \sim MN(0, U_b, V_b)$, B follows a matrix-normal distribution, with U_b a $p \times p$ variance-covariance matrix between markers, and V_b a $t \times t$ variance-covariance matrix between traits (including genetic correlation).

$E \sim MN(0, U_e, V_e)$, E follows a matrix-normal distribution, with U_e a $n \times n$ variance-covariance matrix between individuals, and V_e a $t \times t$ variance-covariance matrix between traits (not due to genetic correlation).

Several authors compared multivariate GP models to univariate ones and found improvement in GP accuracy, but this improvement was not systematic and often small. Cases most favorable to multivariate GP were with correlated traits, many missing data, or with traits showing different heritability values (Calus & Veerkamp, 2011; Guo et al., 2014; Jia & Jannink, 2012; Marchal et al., 2016; Schulthess et al., 2018).

To enhance GP accuracy, some authors proposed to add a secondary trait easy to phenotype and measured on the test population to train the model for predicting a target trait. But, as test genotypes are used in the model, this leads to "data leakage" and thus to a bias on estimated accuracy (Runcie & Cheng, 2019).

Gene action and subsequent GP models

There are three kinds of gene actions: additive, dominant and epistatic (gene-by-gene interaction).

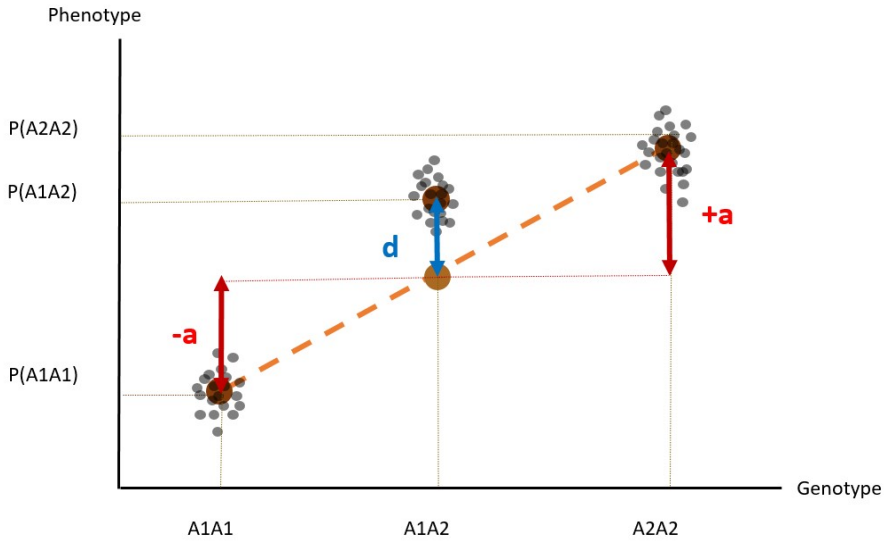


FIGURE 1.11 – Types of QTL effects. Regression of phenotype on genotype at a locus A with two alleles A1 and A2 in a given population. P(A1A1) represents the mean phenotypic value for the genotype A1A1. The brown point represents the mean phenotype in the population, the blue arrow ("d") the biological dominance effect and the red arrow "a" the additive biological effect.

The additive substitution effect can be measured as the change in phenotypic value when one allele is substituted to another one at a given locus (Lynch & Walsh, 1998). Another definition is the slope of the regression of phenotype on marker genotype (allelic doses) at the locus. Dominance deviation is defined as the difference between the heterozygote phenotype and the mid homozygote phenotype. The magnitude of the additive substitution and dominance deviation effects depends on the population mean, which relies on allelic frequencies in the considered population. These statistical effects are different from biological additive and dominance gene effects, defined as the "true" effects of the gene on the phenotype, i.e., independently of allelic frequencies in the population (Falconer and Mackay, 2009, Figure 1.11).

The breeder is interested in knowing and predicting the genetic value of genotypes. For inbred lines, genetic values correspond to breeding values (i.e., additive genetic values), because genotypes are homozygous, hence there is no dominance. For heterozygous plants such as grapevine, we are interested in predicting the total genetic values of genotypes, estimating both additive and dominance genetic effects, which may differ from breeding values. Moreover, grapevine as other outbreeding plants suffer from inbreeding depression which is directional dominance. Thus, taking into account dominance effects can prevent from getting

too much homozygous genotypes. However, when choosing parents to cross (mate allocation), only breeding values are useful, because dominance deviation effects are not transmitted to the progeny.

Epistatic effects are coming from interactions between loci, e.g., additive-by-additive, additive-by-dominance or dominance-by-dominance interactions. If more and more attention has been paid to epistasis in the past years, its integration into GP models remains scarce, mostly because huge datasets are needed to estimate interaction between loci. Estimation of variance partition for a given trait and species requires computation of specific relationship matrix (González-Diéguez et al., 2021; Varona et al., 2018; Vitezica et al., 2018). Epistasis effects may be induced by group-specific QTL effects when crossing heterotic groups together, as demonstrated in maize (Rio et al., 2019). Using haplotypes in GP may improve the capture of short-range epistatic effects, while capturing more clearly the LD pattern between individuals (Hess et al., 2017; Won et al., 2020).

Hence, disentangling genetic variances and effects remains statistically challenging. Moreover, Huang and Mackay (2016) showed that if genetic variance partitions often estimate negligible non-additive variance components, this is mainly due to statistical coding and this does not imply that most gene effects are additive. Thus, accounting for non-additive effects in GP requires specific statistical computing, either by adding specific terms in the equation, or by using a dedicated non-linear method (see below). The first option requires a specific matrix for coding heterozygote genotypes, for example coding $\{aa : 0, aA : 1, AA : 0\}$. But, since this coding corresponds to biological dominance effects, then one needs to use other formula to recover breeding values and dominance deviation effects (Vitezica et al., 2013; Vitezica et al., 2018). Nevertheless, this investment of taking into account for non-additive QTL effects could prove useful especially for vegetatively propagated crops.

Genomic prediction methods

Estimating marker effects is a high-dimension problem, also called "small n, big p problem", because there are typically much more predictor variable effects (marker effects) to be estimated than observations available (phenotypic records). Thus, ordinary least squares (OLS) estimation of marker effects cannot be directly performed and other methods are needed, such as regression with shrinkage. We will consider here three kinds of GP models: frequentist, Bayesian and other non-linear models (classified in Figure 1.12, adapted from Azodi et al. (2019)). In frequentist GP methods, genetic variance associated with marker effects is assumed to be fixed, whereas in Bayesian inference, there is a prior on marker effects distribution. Non-linear GP models encompass multi-kernel and deep-learning GP methods.

Frequentist methods

Commonly used frequentist methods for GP are penalized regression methods, such as RR-BLUP (Ridge Regression – Best Linear Unbiased Predictor or RR) (Henderson, 1985), LASSO (least absolute shrinkage selection operator) (Tibshirani, 1996) or EN (elastic net) (Zou &

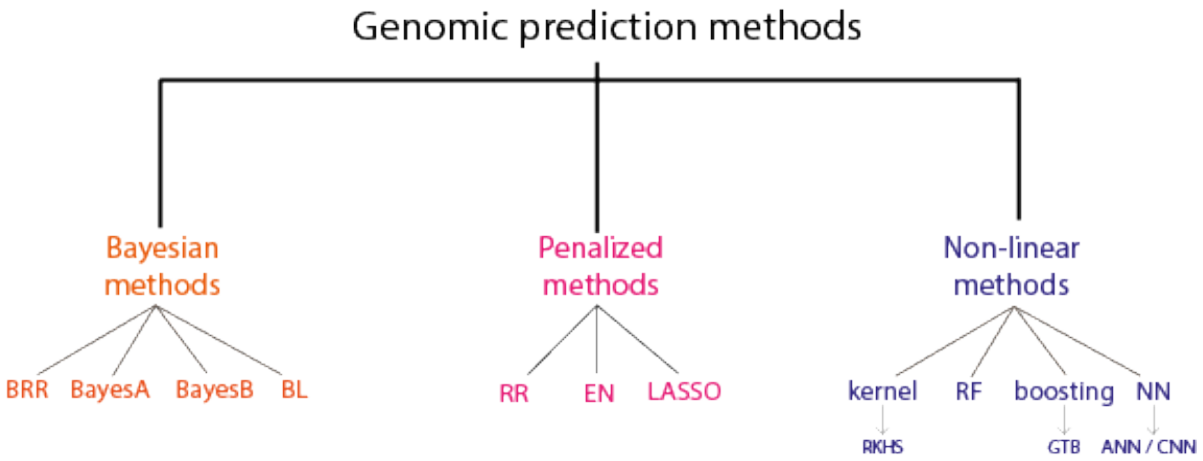


FIGURE 1.12 – Classification of GP methods. rrBLUP, ridge regression Best Linear Unbiased Predictor; B-RR, Bayesian Ridge Regression; BL, Bayesian LASSO; RKHS, reproducing kernel Hilbert space; RF, Random Forest; GTB, Gradient Tree Boosting; ANN, Artificial Neural Network; CNN, Convolutional Neural Network

Hastie, 2005). These three methods add shrinkage to the effects estimated by OLS. The amount of shrinkage is controlled by a parameter (called lambda) which is determined by cross-validation or maximum likelihood. For EN, the balance between RR and LASSO penalty is optimized with another parameter (called alpha). These three models are described in the Materials and Methods section from Brault et al. (2021a) in chapter 2.

In RR models, marker effects are shrunk towards zero but no marker effect is exactly equal to zero, whereas in LASSO and EN, marker effects can be set to zero (Figure 1.13). This sparsity property of LASSO and EN (marker selection) makes them useful also for QTL detection (Cho et al., 2009; Zou & Hastie, 2005).

Another model related to RR-BLUP is GBLUP (Genomic BLUP) which can be written as: $y = \mu + Za + \epsilon$, with $u \sim (0, \sigma_a^2 K)$, with σ_a^2 the additive genetic variance and K the additive genomic relationship matrix, as calculated by (VanRaden, 2008). It is worth mentioning that GBLUP model is equivalent to RR-BLUP when the RR parameter lambda is equal to σ_e^2 / σ_a^2 and K is the estimator (Habier et al., 2007).

In RR-BLUP, the genotypic matrix is directly entered in the model, whereas in GBLUP, genotypic data are used to compute the additive genomic relationship matrix K. Both RR-BLUP and GBLUP assume a fixed genetic variance for all marker effects.

Bayesian methods

In Bayesian methods, marker effects are considered as a random variable, with an associated distribution which can be entered in the model as a *prior* distribution, given the observed phenotypes. Then, a Gibbs sampler outputs a posterior distribution and best prediction of

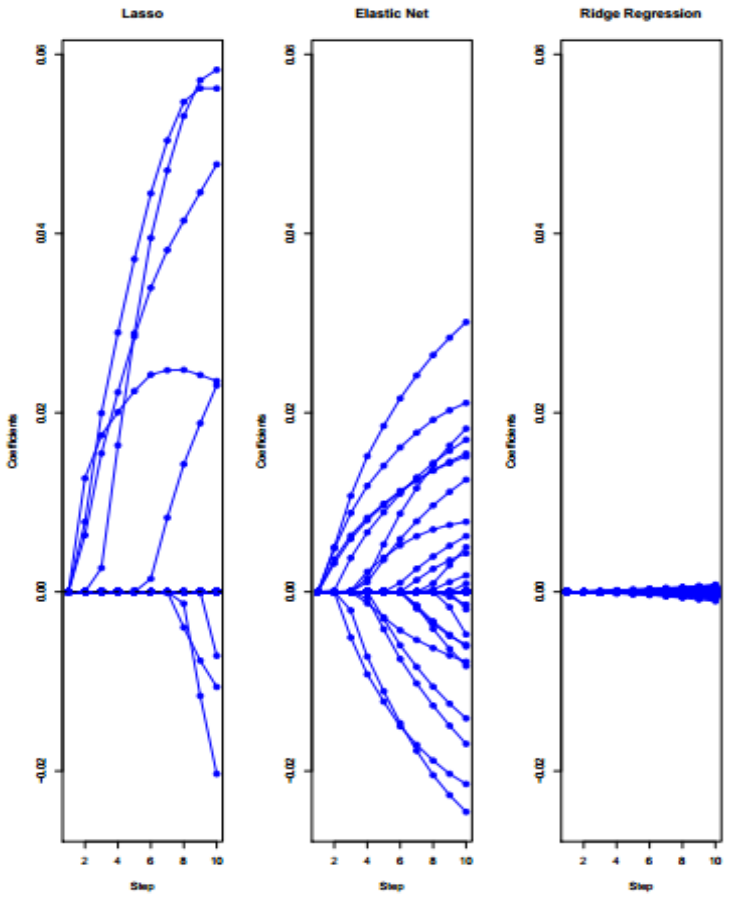


FIGURE 1.13 – Comparison of estimated variable effects. Example of estimated variable effects by RR, LASSO and EN on the same data. The first ten lambda values were used (x-axis) and variable effects were reported for all variables (each line represent a variable). Alpha was kept at 0.2 for EN (from Friedman et al., 2010)

.

marker effects, which can be seen from a frequentist point of view, as the values minimizing the difference between the true and estimated marker effects.

Frequentist methods assume a constant variance for marker effects. Conversely, for Bayesian methods, this variance follows a distribution that can be determined *a priori*. Thus, this distribution is not necessarily uniform.

GP methods from the Bayesian alphabet assume different distributions for marker effects: Student distribution for BayesA (Meuwissen et al., 2001), mixture of student distribution with a spike at 0 for BayesB (Meuwissen et al., 2001), normal distribution with unknown variances for BayesC (Habier et al., 2011), mixture of normal distribution with a spike at 0 and a proportion of markers with non-null effect for BayesCπ (Habier et al., 2011), a double exponential distribution for Bayesian Lasso (Park & Casella, 2008), a mixture of normal distribution with different variances for BayesR (Erbe et al., 2012) (Figure 1.14). The advantage of Bayesian methods is that all available phenotypes in the TS are used to train the

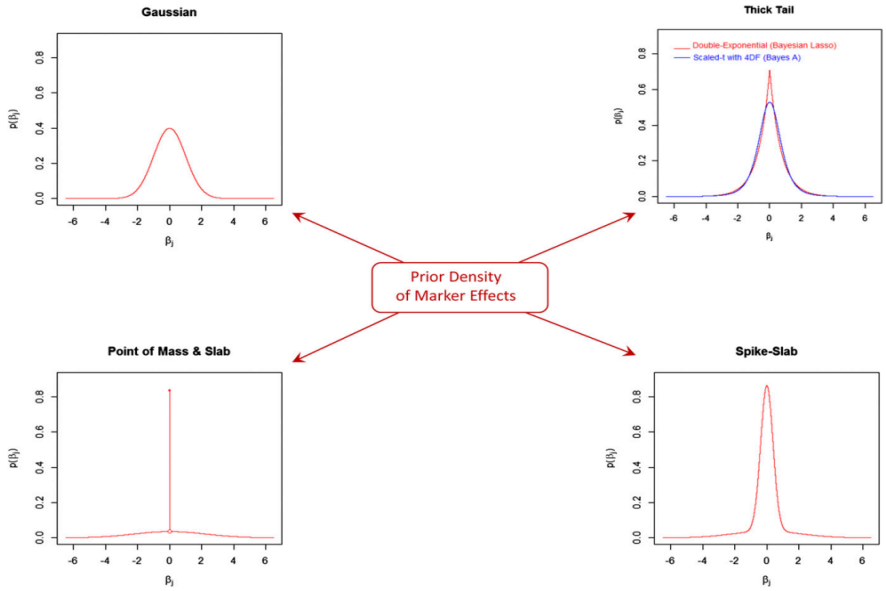


FIGURE 1.14 – Commonly used prior densities of marker effects (all with zero mean and unit variance). The densities are organized in a way that, starting from the Gaussian in the top left corner, as one moves clockwise, the amount of mass at zero increases and tails become thicker and flatter (from de los Campos et al., 2013).

model, whereas in frequentist models, a subset of phenotypes is left apart during the inner cross-validation (Gianola & Fernando, 1986). As with frequentist methods, some of Bayesian methods are performing variable selection.

Non-linear methods

Other types of GP methods such as non-parametric or semi-parametric methods could be better adapted to model non-additive gene effects (epistatic or dominance) (Desta & Ortiz, 2014; Jacquin et al., 2016). Unlike previously described methods, these methods do not make hypothesis on additivity of gene effects. Moreover, they provide a flexible manner of dealing with an always increasing p/n ratio, while former methods are prone to overfitting when p/n becomes large (more than 50) (González-Recio et al., 2014). I will provide a quick overview of the most used non-linear methods.

Some of these methods rely on kernel computation, such as Reproducing Kernel Hilbert Space (RKHS) (Gianola et al., 2006; Gianola & van Kaam, 2008), which is useful to deal with non-linear gene effects without explicitly including them into the model (Morota et al., 2013). Instead of using a genomic relationship matrix, this method computes a transformation of marker genotypes into a kernel matrix based on genetic (Euclidean) distance between genotypes, which is thereafter used in prediction. Gianola and van Kaam (2008) used a multivariate Gaussian kernel for this transformation (Sun et al., 2012). Support Vector Regression (SVR) also relies on a kernel and is a particular case of RKHS (Kadam & Lorenz, 2019). In kernel methods, using a Gaussian kernel comes to using GBLUP, thus the modeling of non-linear

gene effects relies on the chosen kernel.

Random Forest method (RF, Breiman (2001)) is a classifier that optimizes and updates marker effects at each step and thus provide a set of selected markers most likely to have an effect on the phenotype. However, marker selection is not consistent across repetitions.

Deep-learning methods such as Artificial or Convolutional Neural Network (ANN and CNN, respectively) are increasingly popular in statistics and in breeding. Briefly, marker genotypes or other information are entered in the model with an input layer, which is followed by hidden layers. Number of layers and links between them are parameters from the model. At each layer, neurons are weighted and this value is updated in order to minimize a loss function, so that model errors decrease at each layer. Finally, an activation function is applied on the last layer to provide the weighted outputs, i.e., the predicted values (González-Recio et al., 2014). ANN and CNN are underlined by many hyper-parameters that need to be optimized through cross-validation. Azodi et al. (2019) compared many GP methods, including ANN on multiple species and traits. These authors found that ANN did not performed better than RR-BLUP, confirming previous findings which pointed out that when gene effects are additive, non linear methods do not improve predictions over penalized or Bayesian ones (Desta & Ortiz, 2014).

Kadam and Lorenz (2019) suggested that a limited population size could hamper non-parametric methods to outperform classical parametric methods. This was confirmed on simulated data by Abdollahi-Arpanahi et al. (2020), who found a slight superiority of CNN over parametric methods when number of observations increased from 12k to 80k.

Genetic architecture

Genetic architecture can be defined as the number and type of QTL effects underlying a trait. Genetic architecture affects GP accuracy, since some methods are more adapted to a specific genetic architecture such as illustrated with the Bayesian alphabet (Figure 1.14).

Number of QTLs

When the trait under study is controlled by one major QTL, GP is not needed and QTL detection associated with MAS is better adapted. However, when traits are underlined by many QTLs with major or minor effects, GP is preferable (Brault et al., 2021a; Meuwissen et al., 2016).

The hypothesis behind RR-BLUP and GBLUP models is that all markers have a small effect on the phenotype (infinitesimal model). This assumption is most of the time strongly violated, with large-effect QTLs or oligogenic genetic architecture for some traits. However, these methods remain the most widely used in GP studies and give robust GP accuracy across many traits and species (Azodi et al., 2019).

For traits with simpler genetic architecture, i.e., when there are a few QTLs with large effects, sparse methods such as LASSO or BayesCπ give better results than RR (Wang et al., 2015).

The accuracy of different GP methods therefore can provide insights about genetic architecture of the trait.

Distribution of additive QTL effects

QTL detection studies were useful to provide major QTLs, explaining large part of phenotypic variance, such as resistance genes to powdery mildew in grapevine (Pauquet et al., 2001), but also QTLs with intermediate effects (as reviewed in Vezzulli et al., 2019 for grapevine). However, QTL detection studies apply a threshold and minor QTLs are not retained in the model. Thus, QTL mapping experiments cannot be a reliable source for determining QTL effects (Hayes & Goddard, 2001).

For a given number of QTLs, the distribution of QTL effects may differ. Studies hypothesize that the QTL distribution of many quantitative traits follow a gamma distribution with a few QTLs with major effects and many QTLs with small effects, estimation of gamma distribution parameters based on QTL detection studies are provided in (Hayes & Goddard, 2001). Indeed, these distributions are close to some prior distribution of marker effect from Bayesian methods (Figure 1.14). However, this does not necessarily imply that these methods would perform better, as RR or GBLUP remain the most stable method in terms of PA across many traits (Azodi et al., 2019). Finally, as with number of QTLs, GP methods relative performance gives us also insights about genetic architecture of the trait.

Genotype-by-environment interaction

When plants are grown in multi-environment trials (METs), they have different phenotypes, due to environmental effects and genotype-by-environment interaction (GxE). Such interaction is fundamental in plant breeding, as the ranking between genotypes may vary across environments. In particular, some genotypes will show higher performance for a broad range of environments, while others will have a high performance only in a subset of environments (Ceccarelli, 1989). From a statistical viewpoint, both residual and genotypic variances vary across environments and proper statistical models are needed to precisely estimate genotypic and GxE variances (Granato et al., 2018). In GP, the model can be trained with genotypic values from a specific environment and candidate genotypes can be predicted for different environments. Specific GP models allow to include environmental covariates such as solar radiation, water potential or temperatures to increase GP accuracy (Millet et al., 2019). Other authors include phenotypes from different environments in multivariate models (Burgueño et al., 2012). In general, incorporating GxE in GP models increased accuracy by 10% to 40% (Crossa et al., 2017).

1.2.2 Variables affecting genomic prediction accuracy

Many variables affect GP accuracy, namely GP statistical method, as mentioned above, in relation with the trait genetic architecture, the composition of TS relative to VS, its structure

and size, marker density and trait heritability. We will detail the last three variables in what follows.

Relatedness between training and validation sets

The relatedness between TS and VS, as well as TS structure are crucial for determining PA of GP. Indeed, this relies on different nested measures (e.g., additive relationship, LD), each of them having an impact on PA.

Measures of relatedness

Estimating genetic relatedness is useful in breeding as superior individuals transmit part of their genetic variance to the progeny. Moreover, relatedness is a fundamental part of the GP model. Before the advent of molecular markers, only pedigree could be used to define the degree of relatedness between individuals as the probability that a locus is 'identical by descent' (IBD), that is inherited from a common ancestor (Speed & Balding, 2015). However, the alleles inherited from both parents are variable among full-sibs due to recombination. Molecular markers allow determining relationship as identity by state (IBS), which is the *realized* proportion of shared alleles between individuals (Toro et al., 2011; VanRaden, 2008). In GP, additive relationship matrices based on either pedigree or molecular markers are used, depending on the species under study, but the second one is more precise for full-sibs, as Mendelian sampling variation is taken into account (Grattapaglia et al., 2018; Jannink et al., 2010; Nejati-Javaremi et al., 1997). However, using both IBS and IBD matrices could give a higher PA than using additive relationship matrix alone for population with large and complex pedigrees (Grattapaglia et al., 2018).

Impact of relatedness through LD and allelic frequency

Both allelic frequency and LD between SNPs and causal loci (partly due to distance) in the TS affect the estimation of additive statistical effects. Marker effects are estimated within TS, hence they rely on allelic frequencies in TS. Change in allelic frequency between TS and VS would decrease PA. Moreover, marker effects are also relying on physical association between a SNP and causal locus, but this association might be broken in VS and this probability increases with the decrease of genetic relatedness. Thus, differences in allelic frequencies and LD between training and target populations explain why the relationship between these populations is crucial to GP accuracy. Besides, Habier et al. (2007), Habier et al. (2013) distinguished effects of genetic relationship and LD on GP accuracy, the former changes at every generation while the latter is more persistent across generations. Then, using all individuals genotyped in the TS for model training may not be optimal for reaching good prediction accuracy. Indeed, Lorenz and Smith (2015–1) and Brandariz and Bernardo (2019) showed that adding TS individuals unrelated to VS decreased predictive ability.

Adaptation to breeding programs

Maximum GP accuracy is obtained when TS and VS are highly related, e.g., full-sibs (Habier

et al., 2010). However, this configuration is not feasible in a breeding program with hundreds of crosses every year nor for species with long juvenile period. Nevertheless, several adaptations of GP may increase effective relationship between TS and VS. First, TS should reflect the genetic structure (i.e., allelic frequencies) and diversity of the VS (de Roos et al., 2009; Rio et al., 2019). Second, the TS needs to be regularly updated with new released varieties, depending on generation time (Scutari et al., 2016). For that, training set optimization allows selecting the most related individuals from a TS to a given VS, without selecting close relatives in TS (Akdemir & Isidro-Sánchez, 2019; Rincent et al., 2012; Rincent et al., 2018). Finally, selected TS should be composed of enough genotypes depending on the genetic diversity, so that allelic effects could be correctly estimated (Norman et al., 2018; Zhang et al., 2017). TS design constitutes a resource allocation problem, as several variables are involved and the phenotyping budget is fixed (Ben-Sadoun et al., 2020).

Type and density of molecular markers

In grapevine, the most widely used molecular markers are SSR (single sequence repeat) and SNP (single nucleotide polymorphism) markers. Each SSR allele corresponds to a number of repetitions of short sequences, hence there might be dozens of different alleles (Kalia et al., 2011). Each SNP allele corresponds to a nucleotide, hence there are four possible alleles. Most often, only biallelic SNPs (by far the most frequent ones) are used in genetic analyses. That is why SNP genotyping requires at least two or three times more markers than SSR genotyping to give the same level of information in GP (Jannink et al., 2010; Solberg et al., 2008). For instance in grapevine, the partition of varieties into three subpopulations using 20 SSRs or 8840 SNPs led to close results (Flutre et al., 2020). But SNPs are much more abundant in the genome than SSRs, and nowadays, with the advence of sequencing, they can be typed at relatively low cost and thus constitute the main resource for QTL detection and GP studies. Moreover, SSRs are mostly located in non-coding regions, which makes them useful for linkage mapping, but less for GP (Kalia et al., 2011), while a fair proportion of SNPs an be identified in both genes and regulatory sequences.

Beyond the type of marker, required marker density depends mainly on the TS under study and its recombination history, as measured by effective population size (Ne) or LD extent. Many studies in plant and animal breeding measured the impact of SNP density on GP accuracy, by fitting the GP model after pruning the number of SNPs (Norman et al., 2018; Tayeh et al., 2015). Indeed, GP accuracy relies on the linkage between markers and causal polymorphisms. If marker density is too low compared to LD extent, this linkage may be broken in the VS, leading to a loss of PA. Thus, LD extent in the TS is a crucial parameter which will determine the number of SNPs required. In a bi-parental population, i.e., with long LD, the number of markers required to reach reasonable PA is quite low (de Bem Oliveira et al., 2020; Nsibi et al., 2020). These studies also showed that PA with an increasing number of markers quickly reached a plateau, and even a small decrease was observed with large numbers of markers. The level at which PA plateau takes place depends on the training set size and the span of LD (de los Campos et al., 2013). Marker density is no more a limiting factor

for PA. Still, sequencing technologies such as genotyping-by-sequencing (GBS) (Elshire et al., 2011) generate many missing data, thus marker imputation is needed. However, genotypic imputation is challenging for heterozygous genomes with low-coverage genotyping such as GBS (Swarts et al., 2014), but several software are available such as Beagle v4.1 (Browning & Browning, 2016) or FImpute3 (Sargolzaei et al., 2014).

The marker density parameter is also in interaction with the relatedness between training and validation sets, with a higher marker density being required in case of lower relatedness (Norman et al., 2018), since LD decrease dramatically with lower relatedness.

Phenotyping and population design

Besides genotyping and TS constitution, phenotyping has a substantial impact on GP PA, because marker effects will be estimated based on adjusted phenotypic values.

Plant trial design Genetic experiments are conceived so that the genetic value of individuals can be easily estimated. In order to estimate environmental effects, the same genotype can be replicated in several blocks and randomly distributed within each block.

Prediction of genetic value The genetic value is predicted based on a mixed model, including random and fixed effects. We first fit a global model, including genotype effects, all possible confounding effects such as block, year or field coordinates, and interactions.

Then, model selection can be done, on the basis of Fisher tests for fixed effects and log-likelihood ratio tests for random effects. Finally, Best Linear Unbiased Predictors (BLUP) for the random genotypic effect are predicted in the selected model for all individuals. These values can then be used in GP analysis.

Estimation of heritability value The heritability is the part of the phenotypic variance that is explained by the genetic variance. Variance components are extracted from the previously selected model.

There are two kinds of heritability: broad-sense heritability (H^2) is basically a measure of repeatability, that is, the total genotypic variance over the phenotypic variance; it takes into account both additive and dominance genetic variances. Narrow-sense heritability (h^2) measures the ratio of additive genetic variance on total phenotypic variance. This characterizes the part of variance that an individual may transmit to its progeny.

Broad-sense heritability for genotype-entry means can be estimated as in (Piepho & Möhring, 2007).

Broad-sense heritability value determines the part of genetic signal in phenotypic data. If the heritability value is low, then genetic gain will be limited, as environmental factors add noise to phenotypic data. Conversely, increasing the number of genotypes (and therefore

Brick/parameter of the breeder's equation	*i*	*r*	σ_G	*1/T*
(A) Add genotyping to increase selection accuracy	(+)	+		(+)
(B) Replace all or part of the phenotyping effort with genotyping	+			+
(C) Improve the choice of parents to optimize crossbreeding or preserve genetic diversity		+	+	

i, intensity of selection; r, accuracy of selection; σ_G, genetic standard deviation; T, duration of the breeding cycle. Impacts are denoted by "+" for direct impact and "(+)" for indirect impact.

FIGURE 1.15 – Impact of each brick on the parameters of the breeder's equation (from Consortium et al., 2021)

associated genetic variance), and the number of replicates and trials may help increasing the heritability value and getting more precise genotypic values.

Impact of phenotypic data on GP accuracy Heritability value is one of the main drivers of GP accuracy for a given trait (Zhang et al., 2017). Moreover, GP can only exploit the heritable part of a trait for genetic improvement; the rest being determined by environmental or residual variation. Many studies divide the correlation between observed and predicted genotypic value by the square root of heritability, so that traits can be compared to each other.

Consequences of GP on breeding programs

GP was successfully implemented in the breeding programs of many species, including cattle (Hickey et al., 2017; Meuwissen et al., 2016), cereals (Charmet et al., 2014; Krishnappa et al., 2021), also, but more rarely forest trees (Grattapaglia et al., 2018; Lebedev et al., 2020) and fruit trees (Cazenave et al., 2021; Nsibi et al., 2020; Roth et al., 2020). Impacts of GP implementation on breeding programs can be classified into three categories, depending on which components of the breeder's equation (genetic gain) are affected (Consortium et al., 2021, and Figure 1.15).

— Reduced cycle length (T)
Depending on the species under study, GS allows reducing cycle length, because genetic values of individuals can be predicted at early stage, without phenotyping. In animal breeding, GS was revolutionary because of the long generation time (Hickey et al., 2017). On annual species, the reduction was lower but enough to get higher genetic gain (Crossa et al., 2017). In grapevine, phenotypes on fruit and wine are available

only three to four years after crossing. With GP this phenotyping step is skipped and replaced by a much shorter genotyping step (Figure 1.6).

— Increased selection accuracy (r)
The increase of selection accuracy (r parameter in Figure 1.15) by using GS is allowed in several situations. First, if trait to be selected has high GxE interaction, low heritability or phenotype is costly to measure precisely. Thus, using GS allows maximizing phenotyping efforts on TS only and estimating BLUPs of genotypic value. Second, if trait is not available on selection candidates, for example female traits in cattle breeding. In this case, phenotypes from relatives are used but without genomic information, selection accuracy is limited because it is not possible to exploit the Mendelian sampling genetic variability (Consortium et al., 2021).

— Increased selection intensity
Within a breeding program at a given cost, more genotypes can be predicted with GP than would have been planted and phenotyped without GP, because only TS is phenotyped. Thus, selection intensity increases for a given number of released genotypes and budget.

— Decreased genetic variance (σ^2_g)
Compared to phenotypic selection, GP allows maximizing genetic gain at each breeding cycle. However, this implies a sharp decrease of genetic variance, which will hamper further genetic gain. To avoid that, GP needs to incorporate knowledge about inbreeding, in order to manage long-term genetic diversity (Colleau et al., 2017).

1.2.3 Phenomic prediction

The concept of phenomic prediction or selection (PP or PS) was originally defined by Rincent et al. (2018). The authors used a similarity matrix derived from near-infrared spectra (NIRS), instead of the genomic relationship matrix for GP. In some conditions, PP is similar to genomic-like omics based (GLOB) prediction, with the prediction model based on transcriptomics or metabolomics data (Feher et al., 2014; Fernandez et al., 2016; Guo et al., 2016; Riedelsheimer et al., 2012; Schrag et al., 2018; Xu et al., 2016).

Reflectance or absorbance at a given wavelength are determined by chemical bounds in molecules inside a tissue. This property has been extensively used to predict costly traits (Holroyd, 2013; Nicolaï et al., 2007; Osborne, 2006; Tsuchikawa & Kobori, 2015). However, this classical use of spectra in phenotyping is different from PP, as in PP NIRS replace genomic markers and are used to predict uncorrelated traits in a given environment (Rincent et al., 2018, Robert et al. 2021, in press). NIRS derived relationship matrix are expected to be more or less similar to the genomic relationship matrix, because reflectance at a given wavelength reflects the molecules present in a tissue, that are themselves determined by genetic and environmental factors. One step further, GLOB allows to predict uncorrelated traits in an environment for which NIRS are not available, while being measured in another environment (Robert et al. 2021, in press). This particular case is the closest one to GP, with a

relationship matrix coming from NIR measurements in one or several environments to predict traits in other environments. Another property of NIRS is that it integrates both genetic and GxE signals, compared to GP. That is why we expect PP where NIRS and traits are measured in a single environment, to be more accurate than GLOB prediction. Indeed, Rincent et al. (2018) found promising predictive abilities in wheat and poplar, both for PP and GLOB and with higher reliabilities compared to GP, especially in wheat. These results have been confirmed by Krause et al. (2019) in wheat, with spectra coming from unoccupied aerial vehicle (UAV). Besides PP and GLOB reliability, another advantage is that spectra are cheaper and higher-throughput than genotyping. Neither GLOB nor PP have been ever implemented in grapevine. Besides, because of its novelty, many questions remain unanswered about the reliability of GLOB prediction and the modalities under which GLOB should be implemented.

Theoretical, applied, and foreseen developments or GLOB and PP are reviewed in a book chapter entitled "Phenomic selection: a new and efficient alternative to genomic selection" that I have co-authored. This chapter is available in Appendix A.

1.3 Thesis objectives

Most grapevine cultivated varieties in France are susceptible to diseases such as PM and DM, thus requiring many pesticide treatments each year. Moreover, climate change affects grapevine growth, phenology, yield and the resulting wine quality. Grapevine breeding can contribute to mitigate these effects. In France, several winegrowing corporations are starting breeding programs which involve crossing genotypes resistant to diseases with emblematic varieties specific to each wine-growing region. Thus, methodologies for optimizing this selection and avoiding or limiting the time-consuming phenotyping stage are clearly needed. This thesis is a partnership between Inter-Rhône, a wine-growing corporation located in Orange, which work in *Vallée du Rhône* AOC, IFV (Institut Français de la Vigne et du Vin) and INRAE's team DAAV from UMR Agap Institute, which work on grapevine physiology, genetic diversity, quantitative genetics and breeding.

Indeed, grapevine breeding presents several peculiarities. First, the selection outcome is a heterozygous clone. Second, the ideotype is very complex, not clearly defined yet, and it involves many quantitative traits with various heritability values and genetic architectures. Thus, assessing PA of GP for a large number of traits, using populations typically encountered in breeding is a necessity.

It was demonstrated that GP was useful in animal and plant breeding programs mainly because it allows a precise selection of promising individuals, based on predicted genotypic values from molecular markers (Meuwissen et al., 2016). In grapevine, the main expected advantage of GP is time-saving, because of the long generation interval. However, in this crop, proof-of-concept is lacking for GS implementation in breeding programs. So far, only a few studies tested GP in grapevine. First, Fodor et al. (2014) used simulated data mimicking grapevine diversity to test GP first within a core-collection and then across–populations with

the core-collection as TS and bi-parental families as VSs. Second, Viana et al. (2016) implemented GP within an inter-specific bi-parental population of 143 individuals, on eight traits. Then, Migicovsky et al. (2017) predicted 33 traits in a set of 580 *V. vinifera* accessions. Finally, Flutre et al. (2020) used a diversity panel of 279 genotypes to implement GP on 127 traits, they also implemented across-population GP, but with a limited VS size (23 genotypes). These studies did not compare accuracy using different TS and VS, or different parameter affecting PA .

Strong evidence of the accuracy of GP is needed to implement it in current breeding programs in a practical way. The most crucial lacking knowledge is the accuracy - or predictive ability - which can be expected for a given trait with a given model. Several parameters are known to impact PA, such as GP method, heritability of the trait, relatedness between TS and VS, genetic architecture of the trait, marker density. Some of these parameters are specific to grapevine, as the LD, and cannot be extrapolated from other species. Then, a comprehensive study of GP implementation in grapevine, with the test of various parameters affecting PA in a breeding configuration is needed. Besides, new methodologies for genetic value prediction are continuously developed, such as PP, which gave promising results on wheat and poplar (Rincent et al., 2018).

In this context, the objective of my thesis was to implement GP and PP on many traits and in various configurations, in order to optimize the release of new grapevine varieties.

To reach this objective, I adopted the following strategy, relying on three different plant populations (Figure 1.16) already genotyped and phenotyped.

I first worked on a bi-parental population from a reciprocal cross between Syrah and Grenache, classical varieties from *Vallée du Rhône*. I implemented within-population GP on a genetically homogeneous bi-parental population because it was the most favorable configuration for GP and thus it was easier to compare multiple GP methods on many traits. SxG was already genotyped for 3,961 SNPs and phenotyped for 14 traits related to drought tolerance, measured in semi-controlled conditions after applying a differential water stress (chapter 2). Within this population, I tested several GP methods, including multivariate (or multi-trait) ones. Moreover, I used simulated data to truly assess parameters affecting methods relative performance, such as genetic architecture. Data simulation was also useful to study marker selection performed by some GP method or their extension. Finally, I applied these methods for marker selection on the experimental data.

Given the good PA in this ideal combination, I then investigated several configurations closer to those encountered in plant breeding programs in order to specifically test to what extent the relationship between TS and VS affects GP, and in turn whether SG could improve breeding in grapevine (chapter 3). Actually, across-population GP in grapevine has been little tested so far and the thorough characterization of parameters affecting PA has never been done in grapevine. I tested GP using as TS either a diversity panel or half-sib families to predict genotypic values in 10 bi-parental crosses of a half-diallel mating design. GP was

based on common data already available for 32,894 SNPs and 15 traits measured in the field. This allowed me to characterize which PA range could be obtained in across-population GP and which parameters were affecting these values. For this, I assessed PA in two different ways, corresponding to two breeding steps: 1) for predicting the mean genotypic value of each cross, which is useful to select parents to cross, also known as the mate allocation issue; 2) for predicting genotypic values within each cross, which is useful to select offspring after crossing, also known as Mendelian sampling prediction.

Finally, I tested the recently developed PP, which was shown to increase accuracy and reduce costs in breeding programs, and not yet been applied in grapevine. I thus compared its implementation comparatively to GP, based on the same genotypic and phenotypic data as in chapter 3, with NIRS data acquired during the thesis (chapter 4). For that, I first assessed the genetic component in NIRS variance in the two tested populations (half-diallel and diversity panel), in two tissues (wood and leaves), during two years (2020 and 2021). This allowed me to determine the conditions under which the genetic signal in NIRS could be best captured. Then, I applied PP within each population with various statistical models in order to increase PA.

After the presentation of the results of my work, I have included a general discussion and proposed several perspectives of the work (chapter 5).

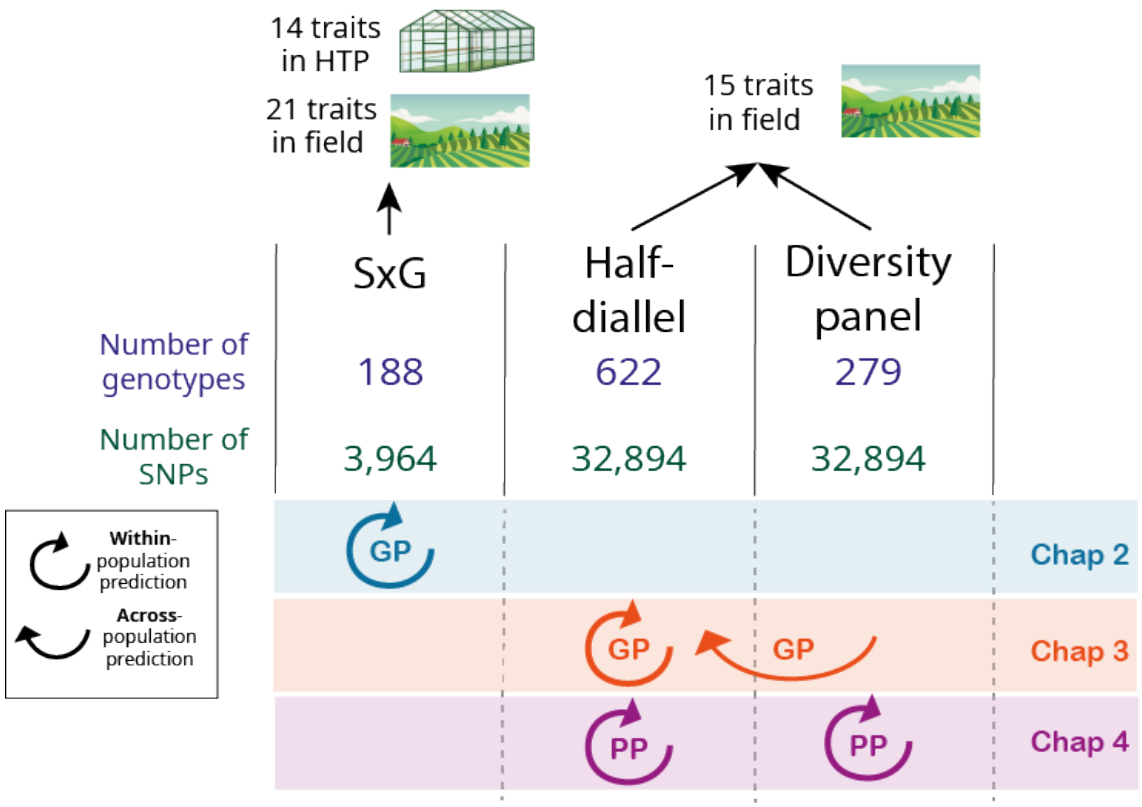


FIGURE 1.16 – Thesis outline.

Chapter 2

Multivariate genomic prediction

2.1 Summary of the chapter

The aim of this chapter was to provide a fine comparison between several GP methods on a bi-parental grapevine population. For that, I used a pseudo-F1 bi-parental population from a reciprocal cross between Syrah and Grenache varieties (SxG). The 188 offspring were phenotyped both in a high-throughput phenotyping platform and in the field, with a differential water stress applied. First, I constructed a new genetic map with 3,961 available SNP markers and compared it with a previous SSR map (Doligez et al., 2006). Second, I simulated phenotypic data based on existing genotypic data under various genetic architectures. Third, I compared univariate and multivariate statistical methods for both genomic prediction (GP) and marker selection (QTL detection); I applied these methods both on simulated and on previously studied phenotypic data from the phenotyping platform (Coupel-Ledru et al., 2016; Coupel-Ledru et al., 2014). Finally, I applied a subset of these methods on phenotypic data from the field.

2.2 Preliminary tests

In the article I (see below), we compared GP methods using a design matrix from a SNP genetic map. Before that, I compared two consensus genetic maps with SSR and SNP markers and their impact on predictive ability.

2.2.1 Quick material and methods

SSR map

We used the consensus genetic map with 153 multi-allelic SSR markers first published in Huang et al. (2012), and recalculated with the Kosambi mapping function by Doligez et al. (2013). The physical positions of SSR markers absent from the latest URGI JBrowse (https://urgi.versailles.inra.fr/Species/Vitis/Genome-Browser) were retrieved by aligning their forward primer with BLAST (Altschul et al., 1990) on the PN40024 12X.v2 reference sequence (Canaguier et al., 2017) using default parameter values, except for the Expect threshold which

was set to 1 or 10. Physical positions were still missing for six SSRs and one was uncertain (high Expect value).

SNP map

The construction of the SNP consensus genetic map including 3,961 markers is described in Article I (see below). This new map was compared to the SSR map based on genetic and physical marker positions. Since these 3,961 markers were ordered and imputed with LepMAP3 (Rastas, 2017), they had four alleles each.

Genotypic data format

In the following, we used the JoinMap version 3 format for genotypic data, according to which each marker genotype is encoded under the following segregation types: ab×cd, ef×eg, hk×hk, nn×np and lm×ll. Each of these comprises several allelic effects: e.g., for ab×cd, the additive effects are a, b, c and d, and the dominance effects are ac, ad, bc and bd. Among the 153 SSRs, 50 segregated ab×cd, 58 ef×eg, 10 hk×hk, 16 lm×ll and 19 nn×np.

Design matrices

The methods compared in this chapter for both GP and QTL detection are based on multiple linear regressions, as detailed in Article 1 (see below). They thus require a design matrix, built from the genotypic data. Here we compared several possible design matrices.

First, for each genetic map (153 SSRs and 3,961 SNPs), we derived two design matrices. In these matrices, a given marker encoded in the JoinMap v3 format corresponded to several columns, with one predictor per allelic effect. For each offspring, genotypes were then coded 0, 1 or 2 for additive allelic effects and 0 or 1 for dominance allelic effects. The first matrix included only additive effects (464 and 15,844 for the SSR and SNP map, respectively). The second one included both additive and dominance effects (996 and 31,688, respectively).

Then, we also recoded the 3,961 SNP markers into bi-allelic additive gene doses (i.e., 0, 1 or 2), which yielded an additional design matrix with 3,961 predictors (Supplementary Table 1 from Article I).

Finally, we used 17,298 SNP markers coded additively, available from the filtered vcf file, before genetic map construction.

Genomic prediction

We compared these six different design matrices, based on the mean predictive ability of seven penalized methods across the 14 traits measured on the phenotyping platform; for one additional penalized method, we compared only three matrices, and for IM methods only two (see detailed description of traits and methods in Article I below).

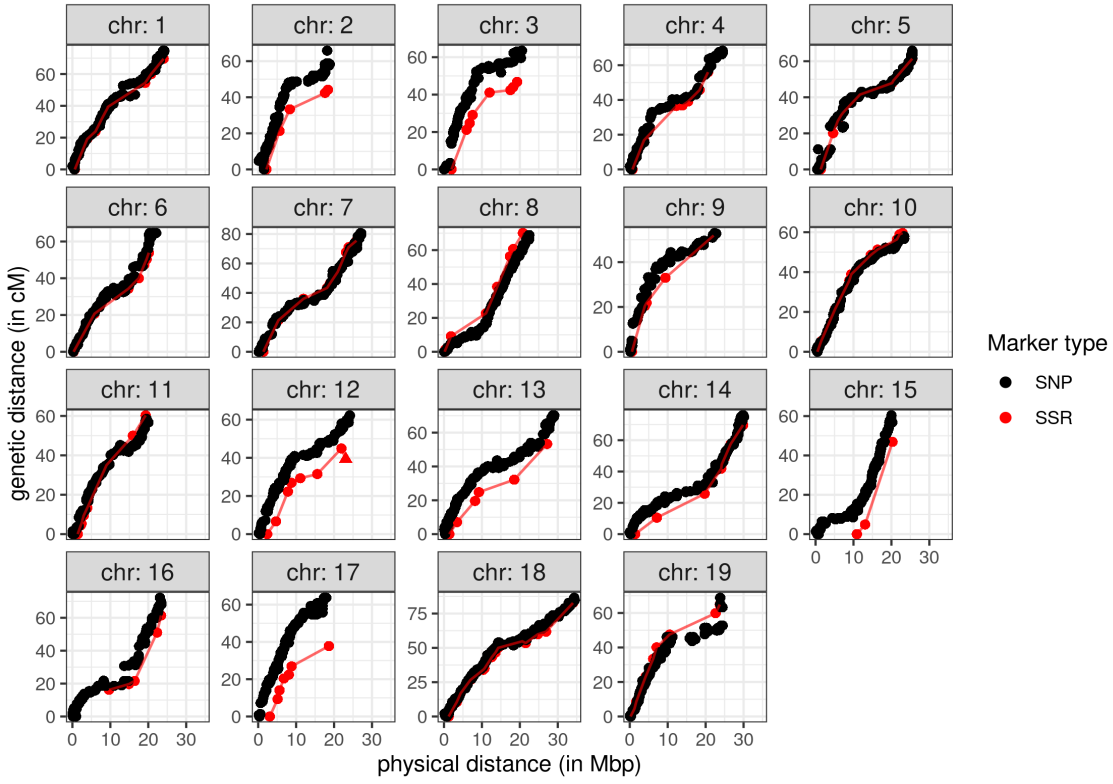


FIGURE 2.1 – Comparison of SSR and SNP consensus genetic maps of a pseudo-F1 *V. vinifera* population, obtained by plotting genetic positions as a function of physical positions for each chromosome. The position of the SSR marker indicated by a triangle on chromosome 12 was uncertain. The red line links the red dots together.

2.2.2 Results and discussion

Comparison between genetic maps

We constructed a saturated consensus genetic map with 3,961 SNP markers (Article I), covering 1,283 cM. It was essentially superimposed on the 1,116 cM SSR map (Figure 2.1). Chromosome 17 had the largest contribution to the 15% difference in genetic length, its length being 37.8 cM with SSRs and 63.7 cM with SNPs. Chromosomes 2, 3, 12, 13 and 15 were also longer on the SNP map. The SNP map was much denser, with an average distance between markers of 0.34 cM (compared to 9.0 cM for the SSR map) and a maximum distance between adjacent markers of 12.0 cM (compared to 29.4 cM for the SSR map). At most places along the genome, genetic order was consistent with physical order for both maps.

Comparison of design matrices

In the SNP genetic map, we observed that some genomic regions were less densely covered than others (e.g., a 10 cM gap on chromosome 19), which might lead to a decrease in predictive ability for traits with QTLs in these regions. We tested this hypothesis for penalized

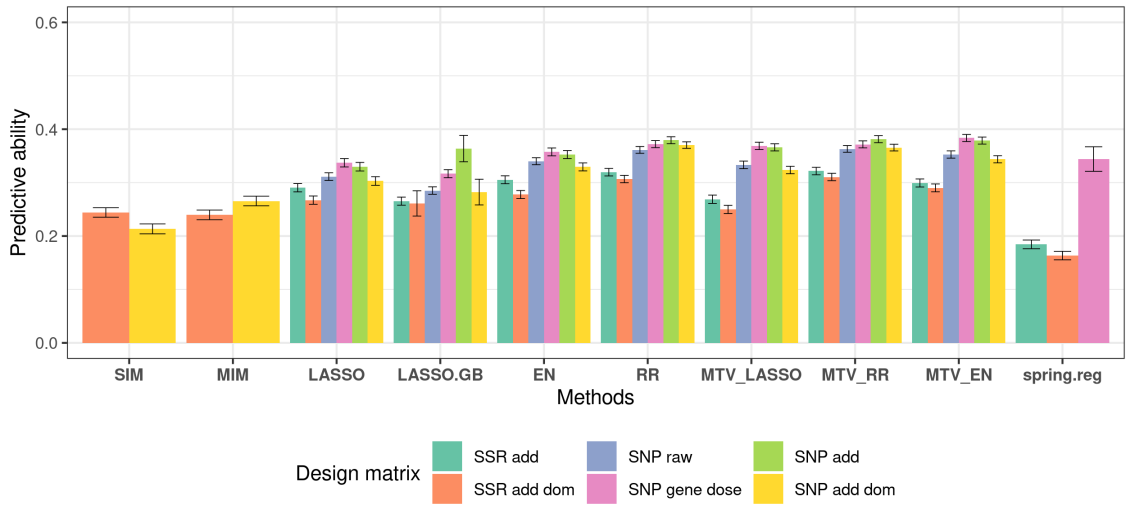


FIGURE 2.2 – Mean predictive ability across traits according to design matrix. Mean Pearson's correlation between predicted and observed genotypic values (BLUPs of phenotypic data) for different design matrices as input for 10 methods. Results were averaged across the 14 traits x 10 cross validation (CV) replicates x 5 CV-folds. Error bars correspond to the 95% confidence interval around the mean. **SSR add**: design matrix from multi-allelic SSR markers, with additive effects only. **SSR add dom**: design matrix from multi-allelic SSR markers, with both additive and dominance effects. **SNP raw**: gene dose design matrix from all bi-allelic SNP markers in the filtered vcf (before genetic mapping), imputed with the mean. **SNP gene dose**: gene dose design matrix from the multi-allelic SNP markers mapped and imputed with Lep-MAP3, and then recoded into bi-allelic markers. **SNP add**: design matrix from multi-allelic SNP markers mapped and imputed with Lep-MAP3, with additive effects only. **SNP add dom**: design matrix from multi-allelic SNP markers mapped and imputed with Lep-MAP3, with both additive and dominance effects.

methods, by using the raw genotypic data (all markers before mapping) imputed with the mean (SNP raw on Figure 2.2, 17,298 markers). On average, this design matrix gave worse prediction ability than other SNP ones with a lower dimension, except for TE.WW (transpiration efficiency for well-watered condition), for which the raw matrix gave the best predictive ability values (data not shown). This suggests that some QTLs for TE.WW were lost (markers not selected) when we predicted with sparser design matrices, whereas this was not the case for other traits. Filtering markers by genetic mapping before applying prediction thus proved to be useful for most traits.

Furthermore, we tested several design matrices for GP on grapevine data. The matrices derived from the SNP map most often led to better predictive ability than those derived from the SSR map, due to higher density, while the additive + dominant coding of allelic effects did not provide any increase in predictive ability over the additive-only coding (Figure 2.2). This could suggest that dominance effects have negligible impact on these traits. Moreover, the additive + dominant coding doubles the matrix dimension (up to 31,688 predictors), which might hamper allelic effect estimation and thus prediction.

We also found that SPRING can explicitly make use of a genetic map. We observed that SPRING had a larger increase in predictive ability from SSR to SNP design matrix than other methods (Figure 2.2). This was probably due to the fact that SPRING uses LD pattern for prediction, this pattern being better captured with a dense genetic map.

Finally, we showed that for most methods, the SNP genotypes recoded into gene doses gave the best predictive ability (Figure 2.2), tied with other SNP design matrices. For computational reasons, we hence chose to use this one for methods comparison in Article I.

2.3 Article I: Harnessing multivariate, penalized regression methods for genomic prediction and QTL detection of drought-related traits in grapevine

This article was submitted to G3: Genes, Genome, Genetics on 11^{th}, January 2021, was accepted on 2^{nd}, July 2021 and was published on $22^{th} July$, 2021. Most of supplementary materials can be found in Appendix B.

G3, 2021, 11(9), jkab248
DOI: 10.1093/g3journal/jkab248
Advance Access Publication Date: 22 July 2021
Genomic Prediction

Harnessing multivariate, penalized regression methods for genomic prediction and QTL detection of drought-related traits in grapevine

Charlotte Brault [1,2,3] Agnès Doligez,[2,3,*] Le Cunff,[1,2,3] Aude Coupel-Ledru [4] Thierry Simonneau,[4] Julien Chiquet,[5] Patrice This,[2,3] and Timothée Flutre[6,*]

[1]Institut Français de la Vigne et du Vin, Montpellier F-34398, France,
[2]UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier F-34398, France,
[3]UMT Geno-Vigne®, IFV-INRAE-Institut Agro, Montpellier F-34398, France,
[4]LEPSE, Univ Montpellier, INRAE, Institut Agro, Montpellier 34000, France,
[5]AgroParisTech, UMR MIA, Paris 75005, France, and
[6]Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE—Le Moulon, Gif-sur-Yvette 91190, France

*Corresponding authors: Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE—Le Moulon, Gif-sur-Yvette 91190, France. Email: timothee.flutre@inrae.fr (T.F.); UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier F-34398, France. Email: agnes.doligez@inrae.fr (A.D.)

Abstract

Viticulture has to cope with climate change and to decrease pesticide inputs, while maintaining yield and wine quality. Breeding is a key lever to meet this challenge, and genomic prediction a promising tool to accelerate breeding programs. Multivariate methods are potentially more accurate than univariate ones. Moreover, some prediction methods also provide marker selection, thus allowing quantitative trait loci (QTLs) detection and the identification of positional candidate genes. To study both genomic prediction and QTL detection for drought-related traits in grapevine, we applied several methods, interval mapping (IM) as well as univariate and multivariate penalized regression, in a bi-parental progeny. With a dense genetic map, we simulated two traits under four QTL configurations. The penalized regression method Elastic Net (EN) for genomic prediction, and controlling the marginal False Discovery Rate on EN selected markers to prioritize the QTLs. Indeed, penalized methods were more powerful than IM for QTL detection across various genetic architectures. Multivariate prediction did not perform better than its univariate counterpart, despite strong genetic correlation between traits. Using 14 traits measured in semi-controlled conditions under different watering conditions, penalized regression methods proved very efficient for intra-population prediction whatever the genetic architecture of the trait, with predictive abilities reaching 0.68. Compared to a previous study on the same traits, these methods applied on a denser map found new QTLs controlling traits linked to drought tolerance and provided relevant candidate genes. Overall, these findings provide a strong evidence base for implementing genomic prediction in grapevine breeding.

Keywords: genomic prediction; QTL detection; multi-trait; breeding; candidate gene; water stress; grapevine

Introduction

Viticulture is facing two major challenges, *i.e.*, coping with climate change and decreasing inputs such as pesticides, while maintaining high yield and quality. This requires understanding the physiological processes that determine adaptation to climate change, such as water use efficiency and their genetic basis (Condon *et al.* 2004). Breeding schemes could then use crosses between genotypes with high water use efficiency, and others resistant to downy and powdery mildews (Vezzulli *et al.* 2019b), to select offspring displaying the most favorable combinations. In crop species, the widespread use of molecular markers through marker-assisted selection (MAS) or genomic prediction (GP) substantially accelerates genetic gains as compared to the traditional phenotypic selection, by allowing early selection of promising genotypes, without the corresponding phenotypic information (Heffner *et al.* 2009). This is of acute interest in perennial fruit species because of the long juvenile period during which most traits of interest cannot be phenotyped. MAS and GP are now widely developed in many perennial species such as pear (Kumar *et al.* 2019), oil palm (Cros *et al.* 2015; Kwong *et al.* 2017), citrus (Gois *et al.* 2016), apple (Muranty *et al.* 2015), and *Coffea* (Ferrão *et al.* 2019). In grapevine, quantitative trait loci (QTL) detection in bi-parental populations led to the identification of major genes for traits with a simple genetic architecture such as resistance to downy and powdery mildews, berry color, seedlessness, and Muscat flavor (Fischer *et al.* 2004; Welter *et al.* 2007; Fournier-Level *et al.* 2009; Emanuelli *et al.* 2010; Mejía *et al.* 2011; Schwander *et al.* 2012). Based on these results, most breeding efforts in grapevine use MAS to improve disease resistance. However, genetic improvement is also needed for traits with a more complex genetic determinism, as well as for others, such as drought-related traits, that are difficult to phenotype. Many

Received: January 11, 2021. **Accepted:** July 02, 2021

minor QTLs have been found for tolerance to abiotic stresses (Marguerit *et al.* 2012; Coupel-Ledru *et al.* 2014, 2016), yield components (Doligez *et al.* 2010, 2013), and fruit quality (Huang *et al.* 2012), as reviewed in Vezzulli *et al.* (2019a). But MAS is not well suited to traits with many underlying minor QTLs (Bernardo 2008). Genomic prediction, which relies on high-density genotyping, is a promising tool for breeding for such complex traits, especially in perennial plants (Kumar *et al.* 2012). Nevertheless, in grapevine, GP has been used in three published papers, only once on experimental data (Viana *et al.* 2016a; Migicovsky *et al.* 2017) and once on simulated data (Fodor *et al.* 2014). Thus, before applying GP to this species, it has to be empirically validated by thoroughly investigating the efficiency of different methods on traits with various genetic architectures.

Both QTL detection and genomic prediction rely on finding statistical associations between genotypic and phenotypic variation. So far, QTL detection in grapevine has been mainly achieved by using interval mapping (IM) methods in bi-parental populations, or more recently genome-wide association studies (GWAS) in diversity panels [see Vezzulli *et al.* (2019a) for a comprehensive review of QTL detection studies in grapevine]. However, most quantitative traits are explained by many minor QTLs, which are difficult to detect either by IM methods or GWAS where each QTL has to exceed a significance threshold. In contrast, GP methods, by focusing on prediction, are less restrictive on the number of useful markers, sometimes resulting in all markers being retained as predictive with a nonzero effect. This is why GP methods are more efficient at predicting genotypic values (Goddard and Hayes 2007) and therefore increasingly popular with breeders (Heffner *et al.* 2010; Crossa *et al.* 2017; Kumar *et al.* 2020).

Widely used methods for GP are based on penalized regression (Hastie *et al.* 2009), notably RR [Ridge Regression, equivalent to Genomic BLUP, GBLUP, Habier *et al.* (2007)] and LASSO (Least Absolute Shrinkage and Selection Operator). Bayesian approaches are also commonly used (*e.g.*, de los Campos *et al.* 2013; Kemper *et al.* 2018), see Desta and Ortiz (2014) for a classification of GP methods. However, overall, Bayesian methods do not achieve better predictive ability than RR or LASSO, while they bear a heavy computational cost when fitted using Markov chain Monte-Carlo algorithms (Ferrão *et al.* 2019). Other methods based on nonparametric models (*e.g.*, Support Vector Machine, Reproducing Kernel Hilbert Space, Random Forest) have been shown to yield lower predictive ability than parametric models (frequentist or Bayesian) when the trait displayed an additive genetic architecture (Azodi *et al.* 2019).

Traits are often analyzed one by one in GP, using univariate methods. Nevertheless, breeders want to select the best genotypes that combine good performance for many favorable traits. Analyzing several traits jointly in GP allows taking into account any genetic correlation between traits (Henderson and Quaas 1976). Calus and Veerkamp (2011), Jia and Jannink (2012), Hayashi and Iwata (2013), and Guo *et al.* (2014) compared univariate *vs* multivariate models' performance. They found a slight advantage for multivariate analysis when heritability was low and data were missing. Predictive ability was particularly improved if the test set had been phenotyped for one trait while prediction was applied to another correlated trait (trait-assisted prediction) as in Thompson and Meyer (1986), Jia and Jannink (2012), Pszczola *et al.* (2013), Lado *et al.* (2018), Velazco *et al.* (2019), and Liu *et al.* (2020). However, this breaks independence between the training and test sets, leading to over-optimistic prediction accuracy (Runcie and Cheng 2019). Multivariate methods have also been proposed for QTL detection by Jiang and Zeng (1995), Korol *et al.* (1995), Meuwissen and Goddard (2004), notably for distinguishing between linkage and pleiotropy when a QTL is found common to several traits. Some methods of multivariate penalized regression, such as in Chiquet *et al.* (2017), were designed to allow both QTL detection and genotypic value prediction. Multivariate GP methods are expected to perform better if traits are genetically correlated, but this remains to be confirmed with additional data. We also hypothesize that these methods will have higher QTL detection power, by making better use of information on the genetic architecture of several intertwined traits.

Methods designed for QTL detection are rarely used for genotypic value prediction. As they select only the largest QTLs, we hypothesize that these methods will provide an accurate prediction so long as the genetic architecture is simple, but would result in poor prediction performance otherwise, as determined in several studies (Heffner *et al.* 2011; Wang *et al.* 2014; Arruda *et al.* 2016). Conversely, some methods for GP, such as the LASSO and its extensions, are also able to select markers with nonnull effects, and hence to perform QTL detection. Their accuracy in detecting QTLs has been partially investigated in an inbred species by Li and Sillanpää (2012) on a single trait and simulated data and by Cho *et al.* (2010) on human data and a binary trait. Additional analyses are thus clearly needed.

This article aims to compare the ability of various methods to predict genotypic values and to detect QTLs in a bi-parental grapevine progeny, by focusing on traits related to climate change adaptation. We first complemented the only available, low density, SSR genetic map (Huang *et al.* 2012) by restriction-assisted DNA sequencing, to construct a saturated SNP map. Using this map, we then simulated phenotypic data to compare several univariate and multivariate methods and assess the impact of simulation parameters. Finally, we reanalyzed the phenotypic data on water stress from Coupel-Ledru *et al.* (2014, 2016), obtained in semi-controlled conditions. The same genotyping data and methods as those applied to simulated data were compared, providing deeper insight into the genetic determinism of key traits underlying water use efficiency, by finding new QTLs and candidate genes.

Materials and methods

Plant material

This study was based on a pseudo-F1 progeny of 188 offspring from a reciprocal cross made in 1995 between *Vitis vinifera* L. cultivars Syrah and Grenache (Adam-Blondon *et al.* 2005).

GBS markers

Genotyping was done by sequencing was performed after genomic reduction, using RAD-sequencing technology, with *ApeKI* restriction enzyme (Elshire *et al.* 2011), as described in Flutre *et al.* (2020). Keygene N.V. owns patents and patent applications protecting its Sequence Based Genotyping technologies. This yielded a total number of 17,298 SNPs.

Consensus genetic map

The genetic map was built with Lep-MAP3 (Rastas 2017), as described in https://doi.org/10.15454/QEDX2V. The resulting map had 3961 fully-informative markers (abxcd segregation) without missing data (missing marker genotypes being automatically imputed in Lep-MAP3). These data were numerically recoded in biallelic doses (0,1,2) according to the initial biallelic segregation and phase (Supplementary Table S1).

Simulation

Phenotype simulations were carried out to (i) compare several methods for prediction accuracy, and (ii) assess the efficacy of these methods to select the markers most strongly associated with trait variation.

Two traits, $\mathbf{y}_1$ and $\mathbf{y}_2$ were jointly simulated according to the following bivariate linear regression model: $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$, where $\mathbf{Y}$ is the $n \times k$ matrix of traits, $\mathbf{X}$ the $n \times p$ design matrix of allelic effects, $\mathbf{B}$ the $p \times k$ matrix of allelic effects, and $\mathbf{E}$ the $n \times k$ matrix of errors. For $\mathbf{X}$, the 3961 SNP markers mapped for the SxG progeny were encoded in four additives and four dominance effects. Therefore $n = 188$, $k = 2$, and $p = 31{,}688$. For $\mathbf{B}$, allelic effects corresponding to s additive QTLs were drawn from a matrix-variate Normal distribution, $\mathbf{B} \sim MV(0, \mathbf{I}, \mathbf{V_B})$, with $\mathbf{I}$ the $p \times p$ identity matrix and $\mathbf{V_B}$ the $k \times k$ genetic variance-covariance matrix between traits such that $\mathbf{V_B} = \begin{bmatrix} \sigma^2_{B_1} & \rho_B \sigma_{B_1} \sigma_{B_2} \\ \rho_B \sigma_{B_1} \sigma_{B_2} & \sigma^2_{B_2} \end{bmatrix}$, where ρ_B is the genetic correlation among traits and $\sigma^2_{B_1}$ and $\sigma^2_{B_2}$ are the genetic variances for both traits $\mathbf{y}_1$ and $\mathbf{y}_2$. In the same way, $\mathbf{E} \sim MV(0, \mathbf{I}, \mathbf{V_E})$, with the $k \times k$ error variance-covariance matrix $\mathbf{V_E} = \begin{bmatrix} \sigma^2_{E_1} & \rho_E \sigma_{E_1} \sigma_{E_2} \\ \rho_E \sigma_{E_1} \sigma_{E_2} & \sigma^2_{E_2} \end{bmatrix}$, where ρ_E is the residual error correlation between traits, and σ^2_E the error variance. We set ρ_B to 0.8, $\sigma^2_{B_1}$ and $\sigma^2_{B_2}$ to 0.1, ρ_E to 0, and narrow-sense heritability to 0.1, 0.2, 0.4 or 0.8, and σ^2_E was deduced.

To explore different genetic architectures, we simulated $s = 2$ or 50 additive QTLs, located at s SNP markers, so that all corresponding additive allelic effects had nonzero values in $\mathbf{B}$. Because all allelic effects were drawn from the same distribution, all QTLs had "major" or "minor" effects for $s = 2$ and $s = 50$, respectively. All dominant allelic effects were set to zero. Two QTL distributions across traits were also simulated. For the first one, called "same," all QTLs were at the same markers for both traits. For the second one, called "diff," the two traits had no QTL in common. Thus, there was genetic correlation among traits only for the "same" QTL distribution.

For each configuration (2 or 50 QTLs, combined with "same" or "diff" distribution), the simulation procedure was replicated $t = 10$ times, each with a different seed for the pseudo-random number generator, resulting in different QTL positions and effects.

In a first simulation set, narrow-sense heritability was assumed equal for both traits and prediction was done with all methods described below. In a second set, we simulated two traits with different heritability values (0.1 and 0.5), for the "same" QTL distribution with $s = 20$ and $s = 200$ QTLs, with QTL effects drawn from a matrix-variate distribution with σ^2_B=1 and ρ_B = 0.5, in order to test the simulation parameters from Jia and Jannink (2012) with our genotyping data. For this second simulation set, prediction was done with a subset of methods only. Simulation parameters are summarized in Table 1.

Table 1 Parameter values in two sets of simulation of two traits in a bi-parental population

Simulation parameter	Same heritability values	Different heritability values
QTL number	2–50	20–200
Heritability value	0.8/0.8–0.4/0.4–0.2/0.2–0.1/0.1	0.1/0.5
Genetic variance	0.1/0.1	1/1
Genetic correlation	0.8	0.5
QTL distribution	Same-Diff	Same

Experimental design, phenotyping, and statistical analysis

Seven phenotypes related to drought tolerance had already been measured for 2 consecutive years on the Syrah × Grenache progeny (on 186 genotypes among the existing 188), in semi-controlled conditions on the PhenoArch platform (https://www6.montpellier.inrae.fr/lepse_eng/M3P, last accessed on 07-21-21) in Montpellier, France, as detailed in Coupel-Ledru *et al.* (2014, 2016). Briefly, of all replicates (six and five per genotype respectively in 2012 and 2013), three (in 2012), or two (in 2013) were maintained under well-watered conditions (well-watered condition, WW), whereas the other three were submitted to a moderate water deficit (water deficit condition, WD). Specific transpiration, *i.e.* transpiration rate per leaf area unit, was measured during daytime (*TrS*) and night-time (TrS_{night}). Midday leaf water potential (ψ_M, *PsiM*) was also measured and the difference between soil and leaf water potential ($\Delta\psi$, *DeltaPsi*) calculated. Soil-to-leaf hydraulic conductance on a leaf area basis (*KS*) was calculated as the ratio between *TrS* and *DeltaPsi*. Growth rate (*DeltaBiomass*) was estimated by image analysis. Transpiration efficiency (TE) was calculated over a period of 10 to 15 days as the ratio between growth and total water loss through transpiration during this period.

These seven phenotypes were studied under each watering condition (WW and WD). We thus considered 14 traits in this study, a trait being defined as a phenotype × watering condition combination, and used the raw data available online (https://doi.org/10.15454/YTRKV6). For each trait, a linear mixed model was fitted with R/lme4 version 1.1-21 (Bates *et al.* 2014) using data from both years:

$$y = \mu + Y + R + x_g + y_g + x_c + y_c + O + C + D + G + G{:}Y + G{:}D + \epsilon \quad (1)$$

First, model 1 with nine fixed effects (Y year, R replicate, x_g, y_g coordinates in the platform within the greenhouse, x_c, y_c coordinates in the controlled-environment chamber where *PsiM* and *TrS* were measured, O operator for *PsiM* measurements, C controlled-environment chamber and D date of measurement) and three random effects (G genotype, $G:Y$ genotype-year, and $G:D$ genotype-date interactions) was fitted with maximum likelihood (ML). The R/lme4 output was given to R/lmerTest version 3.1-2 (Kuznetsova *et al.* 2017) to use its function "step." Backward elimination of random-effect terms was performed using likelihood ratio test, followed by backward elimination of fixed-effect terms using F-test for all marginal terms, *i.e.*, terms that can be dropped from the model while respecting the hierarchy of terms in the model (Kuznetsova *et al.* 2017).

The final model after backward elimination was then fitted with restricted maximum likelihood (ReML) to obtain unbiased estimates of the variance components and empirical BLUPs (Best Linear Unbiased Predictions) of the genotypic values. The acceptability of underlying assumptions (homoscedasticity, normality, independence) was visually assessed by plotting residuals and BLUPs. Broad-sense heritability on a genotype-mean basis was computed assuming a balanced design [see the introduction of Piepho and Möhring (2007)], as:

$$H^2 = \frac{\sigma^2_G}{\sigma^2_G + \frac{\sigma^2_{G:Y}}{\overline{n}_{year}} + \frac{\sigma^2_e}{\overline{n}_{year} \times \overline{n}_{rep}}}, \quad (2)$$

with $\sigma^2_{G:Y}$ the genotype-year interaction variance, σ^2_e the residual variance, $\overline{n}_{year}$ the arithmetic mean number of trials (years) and $\overline{n}_{rep}$ the mean number of replicates per trial. Its coefficient of variation was estimated by bootstrapping with R/lme4 and R/boot packages.

Comparison of genotypic BLUPs

We first recomputed genotypic BLUPs from the raw phenotypic data of Coupel-Ledru *et al.* (2014, 2016) in order to control the model selection step in a reproducible way. These new BLUPs had a strong linear correlation (>0.9) with those used in Coupel-Ledru *et al.* (2014, 2016), as shown in Supplementary Figure S2. The estimates of broad-sense heritability followed the same trend as in Coupel-Ledru *et al.* (2014, 2016) (Supplementary Figure S3). They were higher in WD condition than in WW condition for all traits except *DeltaBiomass*.

Genetic correlation between traits varied widely, some absolute correlation values being very high (*e.g.*, up to 0.99 between *PsiM* and *DeltaPsi* in both conditions) when traits derived from others (Supplementary Figure S4).

Interval mapping methods

Two univariate IM methods were compared, using R/qtl version 1.46-2 (Broman *et al.* 2003). For both, the probability of each genotypic class was first inferred at markers and every 0.1 cM between markers along with the genetic map, using the R/qtl::calcgenoprob function.

Simple interval mapping:

Simple interval mapping (SIM, Lander and Botstein 1989) assumes that there is at most one QTL per chromosome. A LOD score was computed every 0.1 cM with R/qtl::scanone, then 1000 permutations were performed to determine the LOD threshold so that the family-wise (genome wide) error rate (FWER) was controlled at 5%.

Multiple interval mapping:

Multiple interval mapping (MIM, Kao *et al.* 1999) allows the simultaneous detection of several QTLs. It was performed with R/qtl::stepwiseqtl, using a forward/backward selection of Haley-Knott regression model (Haley and Knott 1992), with a maximum number of QTLs set to 4 (or 10 for ROC curve construction, see below), replicated 10 times to overcome occasional instability issues. Only main effects were included (no pairwise QTL × QTL interaction). The LOD threshold was computed with permutations (1000 for QTL detection and 10 for cross-validation of GP, see below) to determine the main penalty with R/qtl::scantwo. QTL positions and effects were determined with R/qtl::refineqtl and R/qtl::fitqtl, respectively. For both methods, QTL positions were determined as those of LOD peaks above the threshold, with LOD-1 confidence intervals (Lander and Botstein 1989).

Penalized regression methods:

Genomic prediction can be seen as a high-dimension regression problem with more allelic effects (in $\mathbf{B}$) to estimate than observations (in $\mathbf{Y}$), known as the "$n \ll p$" problem. The likelihood of such models must be regularized and various extensions, called penalized regression of the Ordinary Least Squares (OLS) algorithm were proposed. Such penalization generally induces a bias in the estimation of allelic effects.

Univariate methods

Ridge regression:

Ridge regression (RR, Hoerl and Kennard 1970) adds to the OLS a penalty on the effects using the *L2* norm and solves the following equation: $\hat{\beta}_{RR} = argmin_\lambda ||Y - X\beta||_2^2 + \lambda||\beta||_2^2$. As a result, all estimated allelic effects are shrunk toward zero, yet none is exactly zero. The amount of shrinkage is controlled by a regularization parameter (λ). We tuned it by cross-validation using the cv.glmnet function of the R/glmnet package version 3.0-2 (Friedman *et al.* 2010) with default parameters, except family ="gaussian" and $\alpha = 0$, keeping the λ value that minimizes the mean square error (MSE). Note that effects associated with correlated predictors are averaged so that they are close to identical, for a high level of regularization.

Least absolute shrinkage and selection operator:

Least absolute shrinkage and selection operator (LASSO, Tibshirani 1996) adds to the OLS a penalty on the effects using the *L1* norm and solves the following equation: $\hat{\beta}_{LASSO} = argmin_\lambda ||Y - X\beta||_2^2 + \lambda||\beta||_1$. As a result, some allelic effects are exactly equal to zero, while others are shrunk toward zero. Hence LASSO performs predictor selection, *i.e.*, provides a sparse solution of predictors included in the best model, in addition to estimating their allelic effect. The LASSO regularization parameter (λ) was tuned by cross-validation with cv.glmnet function (family ="gaussian," $\alpha = 1$), as above. If $n < p$, LASSO selects at most n predictors.

Extreme gradient boosting:

We first applied LASSO for dimension reduction and then Extreme Gradient Boosting, a popular machine learning method (Mason *et al.* 1999), to estimate marker effects. Hence, we called this method LASSO.GB. As gradient boosting is a nonlinear method, it can take into account any nonlinear interaction between markers, providing better prediction. Briefly, Extreme Gradient Boosting iteratively updates the estimation of weak predictors, in order to reduce the loss. This method requires an optimization of many parameters associated with the loss function (MSE). This optimization was done with train function from R/caret package version 6.0-86 (Kuhn 2008) using the "xgbTree" method. As the optimization of numerous parameters was computationally heavy, we fixed some of them (nrounds = max_depth = 2, colsample_bytree = 0.7, gamma = 0, min_child_weight = 1 and subsample = 0.5), while testing a grid of varying parameters (nrounds = 25, 50, 100, 150; eta = 0.07, 0.1, 0.2).

Elastic net:

Elastic net (EN, Zou and Hastie 2005) adds to the OLS both *L1* and *L2* penalties, the balance between them being controlled by a parameter (α); it solves the following equation: $\hat{\beta}_{EN} = argmin_\lambda ||Y - X\beta| \,|_2^2 + (1-\alpha)\lambda||\beta||_2^2 + \alpha\lambda||\beta||_1$. Both α and λ were tuned by nested cross-validation: 20 values of α were tested between 0 and 1 and, for each of them, we used cv.glmnet function to choose between 500 values of λ. EN performs predictor selection but is less sparse than LASSO.

Note that RR, LASSO, and EN all assume a common variance for all allelic effects.

Multivariate methods

Multi-task group-EN:

Multi-task group-EN (MTV_EN, Hastie and Qian 2016) is a multivariate extension of EN, it solves the following equation: $\hat{B_{MTV_EN}} = argmin_\lambda ||Y - X\beta||_F^2 + (1-\alpha)\lambda||\beta||_F^2 + \alpha\lambda||\beta||_2$, F being the Frobenius norm. It assumes that each predictor variable has either a zero or nonzero effect across all traits, allowing nonzero effects to have different values among traits. λ and α parameters were tuned using cv.glmnet (family = "mgaussian"). MTV_RR is the multivariate extension of RR, also tuned with cv.glmnet (family = "mgaussian," $\alpha = 0$). MTV_LASSO is the multivariate extension of LASSO, also tuned with cv.glmnet (family = "mgaussian," $\alpha = 1$). The implementation of these three methods is identical.

The multivariate structured penalized regression:

The multivariate structured penalized regression (called SPRING in Chiquet *et al.* 2017) applies a $L1 - penalty$ (λ_1 parameter) for controlling sparsity (like LASSO) and a smooth $L2 - penalty$ (λ_2 parameter) for controlling the amount of structure among predictor variables (L) to add in the model, *i.e.*, the correlation between markers according to their position on the genetic map. Both parameters λ_1 and λ_2 were tuned by cross-validation using cv.spring function (from R/spring package, version 0.1-0). The regression equation can be written as: $Y = XB + \epsilon$ with $\epsilon \sim \mathcal{N}(0, R)$, R is the covariance matrix of residuals (Gaussian noise). The allelic effects are: $B = -\Omega_{Xy}\Omega_{yy}^{-1}$ and they comprise both direct effects Ω_{Xy} and indirect ones Ω_{yy}.

SPRING solves the following equation: $(\hat{\Omega_{Xy}}, \hat{\Omega_{yy}}) = argmin - \frac{1}{n}\log\ell(\Omega_{Xy}, \Omega_{yy}) + \frac{\lambda_2}{2}tr(\Omega_{yX}L\Omega_{Xy}\Omega_{yy}^{-1}) + \lambda_1||\Omega_{Xy}||_1$. Unlike multi-task group-EN, SPRING selects specific predictors for each trait, *i.e.*, a selected predictor can have a nonzero effect for a subset of the traits. Moreover, SPRING allows the distinction between direct and indirect effects by using conditional Gaussian graphical modeling. These effects are due to covariance of the noise such as environmental effects affecting several traits simultaneously. This distinction results in two kinds of estimated allelic effects: the direct ones, re-estimated with OLS, which are best suited for QTL detection (we called the corresponding prediction method *spring.dir.ols*) and the regression ones, which involve both direct and indirect effects and are best suited for prediction (*spring.reg* method).

Robust extension for marker selection

To enhance the reliability of marker selection by penalized methods, we used two approaches: stability selection (SS) (Meinshausen and Buhlmann 2009) and marginal False Discovery Rate (Breheny 2019), both of which aim at controlling the number of false-positive QTLs. We did not use these methods for genomic prediction, as they are not designed for this purpose.

Stability selection:

SS is a method that controls FWER, it computes the empirical selection probability of each predictor by applying a high-dimensional variable selection procedure, *e.g.*, LASSO, to a different subset of half the observations for all λ values from a given set, and then retains only predictors with a selection probability above a user-chosen threshold. SS is implemented in R/stabs package version 0.6-3 (Hofner and Hothorn 2017) and can also be adapted to a multivariate framework. For QTL detection on experimental data, the probability threshold we applied was 0.6 for LASSO.SS and 0.7 for MTV_LASSO.SS.

Marginal false discovery rate:

Marginal false discovery rate (mFDR) has been defined by Breheny (2019) as a modified version of the FDR in which those variables correlated with the causative features are not considered as false discovery. This study provided an accurate estimation of mFDR for a given λ when using EN or LASSO, thus allowing the selection of a more conservative value of λ in order to remain below a given mFDR threshold. We applied mFDR with the R/nvcreg package version 3.12.0 (Breheny 2019). For QTL detection on experimental data, we set mFDR to 10% for LASSO.mFDR and EN.mFDR. To our knowledge, this approach had not been adapted yet to a multivariate framework.

Evaluation and comparison of methods

All methods were compared on two aspects: their ability to predict genotypic values, and their ability to select relevant markers, *i.e.*, to detect QTLs. To assess the prediction of genotypic values on simulated data, we used the Pearson's correlation coefficient between predicted and simulated genotypic values (prediction accuracy). On experimental data, we used the same criterion, but the actual genotypic values being unknown, we used their empirical BLUPs instead (predictive ability).

For QTL detection on simulated data, the methods were compared using criteria of binary classification based on the numbers of true positives and false negatives. On experimental data, because true QTLs are unknown, no such comparison could be performed; instead, we compared the outcome of the different methods and QTLs were deemed reliable when found with several methods.

Genomic prediction

A nested cross-validation (CV) was applied to assess prediction by the various methods.

- An outer $k1 - fold$ CV was performed to estimate the performance metrics, with an inner $k2 - fold$ CV applied to the training set of each outer fold to find the optimal tuning parameters for the method under study (Supplementary Figure S5). Both $k1$ and $k2$ were set to 5 (see Arlot and Lerasle (2016). The partitions of the outer CV were kept constant among traits and methods.
- For IM methods, the optimal tuning parameter was the LOD threshold obtained from permutations, and the effects for the four additive genotypic classes (ac, ad, bc, and bd) were estimated by fitting a multiple linear regression model with genotype probabilities at all QTL peak positions as predictors, using R/stat::lm. For penalized regression methods, parameters were optimized with specific functions such as cv.glmnet and cv.spring.
- As performance metrics, we used mainly Pearson's correlation (corP) but we also calculated the root mean square predicted error (RMSPE), Spearman correlation (corS), the model efficiency (Mayer and Butler 1993), and test statistics on bias and slope from the linear regression of observations on predictions (Piñeiro *et al.* 2008).

For experimental data, the whole nested cross-validation process was repeated 10 times ($r = 10$), whereas for simulated data it was performed only once, but on 10 different simulation replicates ($r = 1$ and $t = 10$). The 14 traits were analyzed jointly for MTV_RR, MTV_LASSO, and MTV_EN. But for SPRING, since analyzing all traits together was computationally too heavy, we split traits into three groups by hierarchical clustering (Supplementary Figure S6) performed with R/hclust applied to

Table 2 Parameter ranges for ROC curve computation, for comparing predictor selection performance of different methods

Method	SIM/MIM	LASSO/MTV_LASSO	Stability Selection	SPRING	EN	mFDR
Parameter name	*LOD*	λ	*probability threshold*	λ_1	λ	*mFDR*
Lowest constraint	0	10e-5	0.5	10e-8	10e-4	0.3
Highest constraint	14	0.25	0.9	0.25	8	0

genotypic BLUPs. All traits within a given cluster were analyzed together.

For simulated data with the same heritability values for both traits, performance results were averaged not only over simulation replicates and partitions of outer CV, but also over traits, because both traits were equivalent in terms of simulation parameters. For simulated data with different heritability values, performance results were averaged only over simulation replicates and partitions of outer CV. For experimental data, performance results were averaged over partitions of outer CV and outer CV repetitions.

QTL detection

Simulated data:

The quality of a predictor selection method is usually assessed through the relationship between statistical power (*i.e.*, the true positive rate, TPR) and type I error rate (*i.e.* the false positive rate, FPR). To compare methods, we thus used the ROC (receiver operating characteristic) curve (Swets *et al.* 1979), which is the plot of TPR as a function of FPR over a range of parameters (Table 2), and the pAUC (partial Area Under the Curve; McClish 1989; Dodd and Pepe 2003). Any marker selected at ±2 cM of a simulated QTL was counted as a True Positive.

For methods with two tuning parameters, one parameter was kept constant (α at 0.7 for EN and EN.mFDR, and λ_2 at 10e-8 for SPRING). We tested several α values for EN but it did not change much the results (data not shown). For MIM, the maximum number of QTLs that can be integrated into the model was set to 10.

Experimental data:

Comparison between methods was based on the number of detected QTLs, the magnitude of their effects, and the percentage of variance globally explained by all detected QTLs.

For MTV_LASSO and SPRING, we split traits into three groups as described above, for computational reasons (for SPRING) and to test whether such splitting evidenced more reliable QTLs (for MTV_LASSO). The parameters of penalized methods were tuned by cross-validation, with MSE as the cost function. We compared predictor selection between methods in terms of the number of common selected markers per trait, *i.e.*, the intersection between markers selected by penalized methods and markers inside confidence intervals found by IM methods. Then all markers in high LD with those selected were considered as selected too. The threshold was defined as the 95% quantile of LD value distribution, for all pairs of markers belonging to the same chromosome (Supplementary Figure S7), which gave a LD threshold of 0.84.

We deemed selected markers as highly reliable if they were either (i) selected by at least five methods, whatever the methods, (ii) or selected by both EN.mFDR and MIM (see Results). Then, we defined a highly reliable QTL as the interval of ± 3 cM around each highly reliable marker (Price 2006; Viana *et al.* 2016b), as predicted by polynomial local regression (loess) fitting of genetic positions to physical position. When several markers were selected inside the 6 cM interval, the QTL interval was extended accordingly. The genetic positions of this interval were then converted into physical positions, by fitting loess. QTLs overlapping for several traits on the SNP map were merged into a single QTL, by physical intervals' union.

Candidate genes exploration:

After merging the most highly reliable QTLs colocalized between traits, we searched for underlying candidate genes. We retrieved the list of genes overlapping the intervals of our QTLs from the reference *Vitis* genome 12X.v2 and the VCost.v3 annotation (Canaguier *et al.* 2017). We then used the correspondence between IGGP (International Grapevine Genome Program) and NCBI RefSeq gene model identifiers provided by URGI (https://urgi.versailles.inra.fr/Species/Vitis/Annotations, last accessed on 07-21-21) to identify putative functions from NCBI, when available. For those genes with a putative function, we then refined the analysis to retrieve additional information about their function and expression. We searched UniProt (https://www.uniprot.org/, last accessed on 07-21-21) and TAIR (https://www.arabidopsis.org/, last accessed on 07-21-21) databases based on homologies to access a complete description of gene function, name, and corresponding locus in *Arabidopsis*. In addition, we used the GREAT (GRape Expression Atlas) RNA-seq data analysis workflow (https://great.colmar.inrae.fr/app/GREAT, last accessed on 07-21-21), which gathers published expression data, to assess the level of expression of our candidate genes in grapevine leaves and shoots, relevant organs for the traits under study. RNA-seq data are normalized as detailed in the 'User manual' section of the GREAT platform: "from the raw read counts, the normalized counts (library size normalization) and the RPKM (gene size normalization) are calculated for each gene in each sample." Data were retrieved with all filters set to "Select All" except for the organ considered that was restricted to 'Leaves' and 'Shoot'.

Results

Comparison of methods with simulated data prediction: cross-validation results

Traits with the same heritability value:

Methods were compared for prediction accuracy by applying cross-validation on simulated data with four different configurations and four heritability values.

Mean Pearson's correlation coefficient varied from 0.16 to 0.98, with a strong effect of heritability on prediction accuracy in all configurations, for the seven main methods (Figure 1). As expected, MIM performed very well in the "major" configurations across all heritability values but yielded the least accurate prediction in the "minor" ones. On the opposite, RR performed very well in the "minor" configurations, but yielded the least accurate prediction in the "major" ones. EN prediction performance was always intermediate between those of RR and LASSO. QTL distribution among traits – "same" (for QTLs at the same positions) or "diff" (for QTLs at different positions) - had very little effect on prediction accuracy. Moreover, we did not observe any superiority of multivariate methods over univariate ones, despite

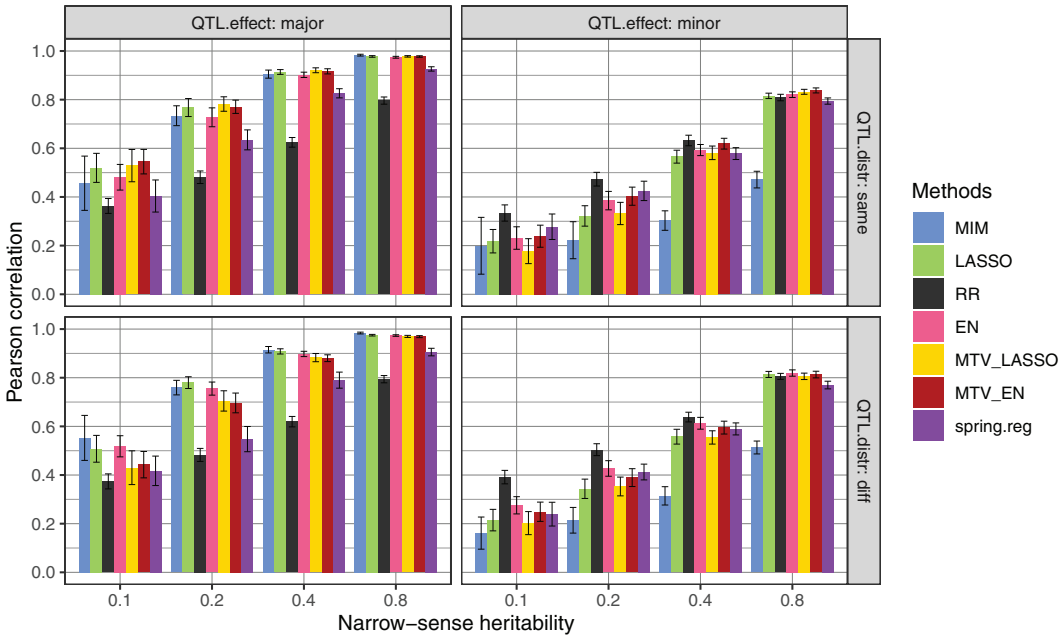


Figure 1 Genomic prediction accuracy (Pearson's correlation between predicted and true genotypic values) of seven methods applied to 3961 markers and two simulated traits in a bi-parental population with different heritability values and four QTL configurations (number × distribution among traits). major: 2 QTLs; minor: 50 QTLs; same: QTLs at the same positions for both traits; diff: QTLs at different positions between traits. For each heritability value and configuration, prediction accuracy was averaged over 100 values (2 traits × 10 simulation replicates × 5 cross-validation folds). The error bar corresponds to the 95% confidence interval around the mean.

the strong genetic correlation simulated between traits (ρ_B=0.8) and no correlation between errors.

The prediction accuracy of four additional methods is shown in Supplementary Figure S8 and prediction accuracy values. Other performance metrics (see Materials and Methods) are given for all methods in Supplementary Table S9. All IM methods yielded equivalent prediction accuracy. LASSO.GB did not improve performance compared to LASSO. MTV_RR showed equivalent performance to univariate RR. Prediction accuracy with *spring.dir.ols* was always lower than with *spring.reg*, and even very low for "minor" configurations. With 100 or 1000 simulated QTLs (under each QTL distribution) the ranking of methods based on prediction accuracy did not change compared to "minor" configurations (Supplementary Figure S10).

Traits with different heritability values:

To further compare prediction accuracy between univariate and multivariate methods, we simulated two correlated traits with different heritability values, 0.1 and 0.5. MTV_LASSO performed slightly better than univariate LASSO for the lowest heritability trait; however, differences were not significant (Supplementary Figure S11). On the opposite, prediction accuracy was lower with MTV_LASSO than with univariate LASSO for the highest heritability trait, reaching quite low values with 200 simulated QTLs. The same trends were also visible for MTV_EN and EN. MTV_RR never improved prediction over RR and *spring.reg* never performed better than RR.

Since these results were unexpected, we also compared prediction accuracy of the above methods using the simulated data published by Jia and Jannink (2012). We obtained very similar differences among methods to those with our own simulated data, even though prediction accuracy was higher in all cases (Supplementary Figure S12).

QTL detection: ROC curve results

We compared the main methods mentioned above (except RR that does not perform marker selection), as well as some robust extensions, for their marker selection performance, by means of ROC curves, using the same simulated data in the four configurations (Figure 2). The closer a ROC curve obtained through a given method approaches the optimum point (*i.e.*, FPR = 0 and TPR = 1), the better is the method's selection performance. As expected, IM methods (SIM and MIM) showed low selection performance when many minor QTLs were simulated and high selection performance when a few major QTLs were simulated. Note that the MIM curve was hardly visible; it roughly overlapped with the SIM curve but stopped at a low FPR because it could not select many QTLs by design.

The penalized regression methods always performed at least as well as the IM methods and even much better in the case of "minor" configurations. Among penalized methods, none was clearly better than the others in all configurations, except for a slight superiority of MTV_LASSO in the "same_minor" configuration. These methods, and particularly *spring.dir.ols*, displayed high variability in classification results with two simulated QTLs ("major" configurations). Indeed, when one QTL was not detected in one trait, impact on TPR was stronger than with 50 simulated QTLs.

The most interesting part of the ROC curve for QTL detection is the left most part, *i.e.*, that with a low FPR (*e.g.*, below 0.1). We thus calculated the partial Area Under the Curve (pAUC) for FPR between 0 and 0.1 for methods reaching that value (Supplementary Figure S13). EN resulted in constantly high pAUC across configurations and heritability values. In contrast, pAUC for SIM was quite high at low heritability values for the "same_major" configuration but dropped for other configurations and heritability values.

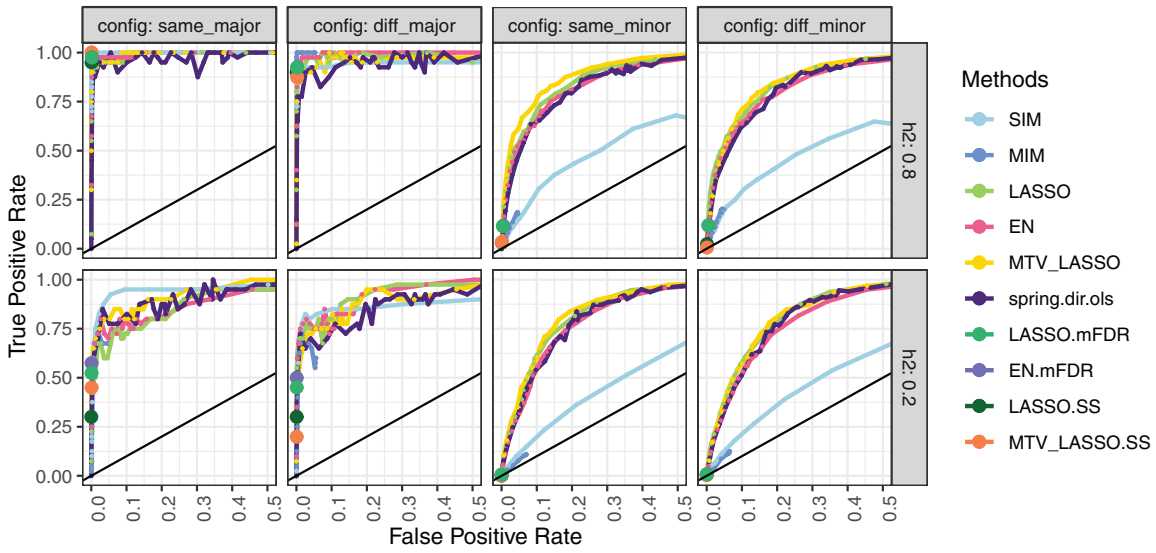


Figure 2 ROC curves for 10 methods applied to 3961 markers and two simulated traits in a bi-parental population with two heritability values and four QTL configurations (number × distribution among traits). major: 2 QTLs; minor: 50 QTLs; same: QTLs at the same positions for both traits; diff: QTLs at different positions between traits. Results are averaged over 2 traits × 10 simulation replicates. TPR: (number of correctly found QTLs/number of simulated QTLs), FPR: (number of falsely found QTLs/number of markers outside a QTL). For robust methods (mFDR and SS), as the FPR remained very low, we display only a single point corresponding to the lowest parameter constraint and thus to the highest TPR.

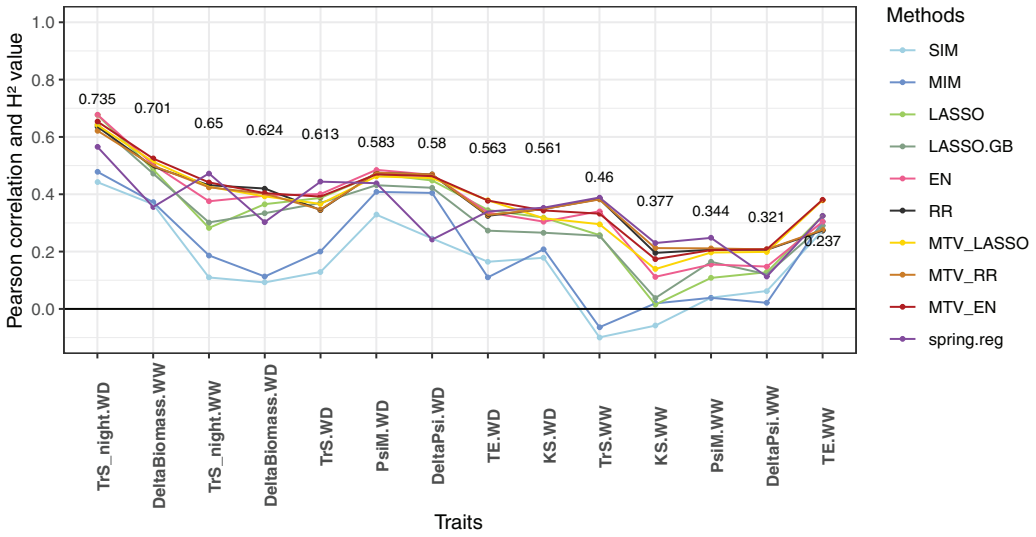


Figure 3 Mean genomic predictive ability (Pearson's correlation between genotypic BLUPs and their predicted values), obtained by cross-validation for 10 methods applied to 14 traits related to water deficit and GBS gene-dose data, within a grapevine bi-parental population. Broad-sense heritability values are reported for each trait (y-position of the number corresponds to heritability estimate). Traits are ordered by decreasing heritability. For each trait, predictive ability is averaged over 10 cross-validation replicates × 5 cross-validation folds).

Results on experimental data

Genomic predictive ability:

Mean genomic predictive ability per trait ranged from −0.10 to 0.68 (Figure 3 and Supplementary Table S14). It decreased with broad-sense heritability. IM methods (in blue) were always among the three poorest methods for prediction. Based on the mean predictive ability averaged across traits, MTV_EN yielded the highest correlation (0.384), followed by RR (0.3721), MTV_RR (0.3716), MTV_LASSO (0.369), EN (0.357), *spring.reg* (0.344), LASSO (0.329), LASSO.GB (0.313), MIM (0.200), and SIM (0.162). However, based on the number of traits for which each method gave the best prediction, *spring.reg* had the highest score, with 6 traits out of 14, followed by MTV_EN (3 out of 14) and EN (2 out of 14).

In a nutshell, MTV_EN and RR, tied with MTV_RR, provided the best mean predictive ability across traits. Even though *spring.reg* outperformed them for some traits, its performance was unstable and especially low for *DeltaBiomass.WW*, *DeltaBiomass.WD*, *DeltaPsi.WW*, and *DeltaPsi.WD*. For computational reasons, all

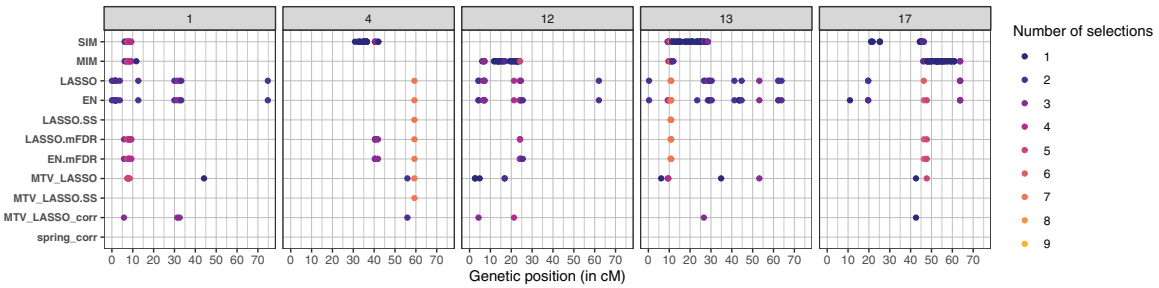


Figure 4 Marker selection by all methods for *TrS_night.WD* trait on chromosomes 1, 4, 12, 13, and 17. Each marker selected by a given method is represented by a colored point, the color indicating the number of methods that have selected that specific marker. The boxes correspond to chromosomes and the x-axis to the position along the genetic map (in cM).

traits could not be analyzed together with *spring.reg*, but were divided into three groups. These four traits with low predictive ability belonged to the same group. Yet, the effect of group membership on predictive ability was not significant at 5% (*P-value* = 0.30 and percentage of variance explained = 24%).

QTL detection:

To address the intersection of SNP selection by all methods, and determine the number of reliable intervals (QTLs) and their position, we examined in detail marker selection for each trait and chromosome. These results are given in Supplementary Table S15, together with genetic and physical positions and the percentage of variance explained. These results are plotted in Figure 4 for night-time transpiration under water deficit (*TrS_night.WD*) and in Supplementary Figure S16 for all traits.

Most of the time, more markers were selected for traits under water deficit than for traits in well-watered conditions, and they were most often selected by several methods. Penalized methods tended to select exactly the same markers, not only close ones; for example, for TrS_{night}.*WD* on chromosome 4, the same marker (at physical position 21,079,664 bp) was selected by seven different methods (Figure 4).

We considered markers selected by both MIM and EN.mFDR as highly reliable ones for three reasons: (1) markers selected by both MIM and EN were considered as reliable ones, because most markers selected by LASSO were also selected by EN, whereas MIM marker selection was quite different; (2) simulations showed that MIM and mFDR methods led to a very low FPR; (3) these methods belong to different method classes (IM *vs* penalized regression). We also considered as highly reliable those markers selected by at least five methods. These criteria resulted in a set of 59 highly reliable selected markers, which were converted to genetic intervals of ± 3 cM around each selected marker. Overlapping intervals per trait were merged, resulting in 25 highly reliable QTLs.

These 25 QTLs involved nine traits, mostly under water deficit, and were located on seven chromosomes (Supplementary Figure S17, Supplementary Table S18). Some QTLs colocalized for different traits, such as on chromosome 1, and had similar distributions of genotypic BLUPs according to genotypic classes (Supplementary Figure S19).

Among these 25 QTLs, we found eight new highly reliable QTLs compared to Coupel-Ledru *et al.* (2014, 2016), among which five were not detected by MIM. In particular, a completely new QTL for *TrS_night.WD* was found alone on chromosome 12. Most other new QTLs were colocalized with QTLs previously found in single-year analysis and/or for the other watering condition. Notably, we observed colocalization of *TrS_night.WD*, *TE.WD* and *DeltaBiomass.WD* QTLs on chromosomes 4 and 17.

In total, the percentage of variance explained (adjusted R^2) per trait was 51.3% for *TrS_night.WD* (36% in 2012 for Coupel-Ledru *et al.* 2016), 33.9% for *PsiM.WD*, 31.4% for *DeltaPsi.WD*, 26.9% for *DeltaBiomass.WW*, 19.4% for *TE.WD*, 18.6% for *TE.WW*, 17.0% for *KS.WD*, 14.9% for *DeltaBiomass.WD*, and 8.5% for *TrS.WD*.

Candidate genes

After merging the QTLs colocalized between traits, we obtained 12 intervals, located on chromosomes 1, 4, 10, 12, 13, 17, and 18, harboring a total of 3461 genes according to the VCost.v3 annotation (Canaguier *et al.* 2017). Among them, 2379 had an NCBI Refseq identifier and 1757 a putative function (Supplementary Table S20). We then focused our analysis on the eight "new" intervals, *i.e.*, those that were not overlapping with QTL intervals repeated over years by Coupel-Ledru *et al.* (2014, 2016). They encompassed 1155 genes, half of which were annotated. We were able to retrieve from TAIR and/or UniProt a more precise description of the gene function for 86% of the annotated genes (Supplementary Table S20). The remaining ones either did not have any homologous gene in *Arabidopsis* or were not described in the above-mentioned databases. RNA-seq data was available on the GREAT platform for 90% of the annotated genes. We further focused on the highly reliable QTL co-localized on chromosome 4 for *TE*, *TrS_night* and *DeltaBiomass* under various conditions. We proceeded to a functional classification of the 161 annotated genes underlying this QTL, based on the full description previously retrieved (Supplementary Tables S5 and S21). For 75 genes, an integrated function at the plant or organ level was explicitly quoted in the homologous-based description. We grouped these integrated functions into 12 major groups (Figure 5). For a substantial part of genes, functions consistently related to traits involved: 15 genes related to hydraulics (stomata, xylem, and trichomes), relevant for *TrS_night* and thus *TE*; 27 to growth or development and one to photosynthesis, both relevant to *DeltaBiomass* and thus *TE*. For the 86 genes for which no integrated function was explicitly quoted, we further built a classification based on their cellular or molecular function. Among them, we found six genes related to carbon metabolism, one to wall formation (both relevant for *DeltaBiomass*) and six to drought stress signaling and drought-related hormones (relevant for *TrS_night*).

Discussion

This study contributes to our knowledge of the complex genetic determinism of vegetative traits under different watering

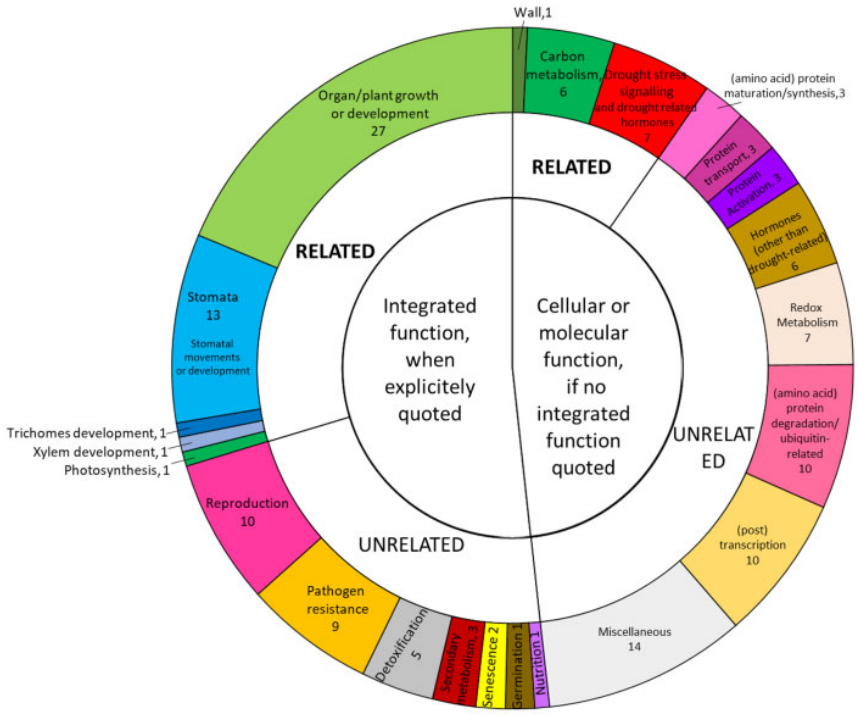


Figure 5 Functional classification of the annotated genes underlying the highly reliable QTL detected on chromosome 4 for night-time transpiration, growth, and TE. Hierarchical classification of the 161 genes based on their functions. See Supplementary Table S21 for the details of this classification. When an integrated function at the organ or plant level was explicitly quoted in the gene annotation, genes were classified on this basis. When no integrated function was explicitly quoted, they were classified based on their cellular or molecular function. In both cases, functions were then classified as "Related" if related to the traits of interest in this QTL, or "Unrelated" if not.

conditions in three different ways. We compared by simulation several univariate and multivariate methods for genomic prediction and QTL detection, and re-analyzed grapevine phenotypes obtained under semi-controlled conditions. In particular, we showed that penalized methods are valuable not only for prediction but also for QTL detection. Indeed, we found new QTLs using these methods and identified relevant candidate genes.

Methodological aspects: method comparison

Handling linkage disequilibrium:

IM methods estimate genotypic probabilities between markers according to a genetic map which is computationally costly to build. On the other hand, most penalized methods do not require any previous knowledge on LD.

The LASSO assumption that all predictor variables are independent is all the more violated that there are many markers. In the case of a group of correlated predictors (*e.g.*, SNPs in LD), EN selects either no or all predictors within the group with close estimated effects (Zou and Hastie 2005), whereas LASSO selects a single predictor. In that sense, EN aims at correcting the drawbacks of LASSO when predictor variables are highly correlated. By exploring a large number of configurations of the finite-sample high-dimensional regression problem, Wang *et al.* (2020) showed that EN is competitive for both prediction and selection in most cases with highly correlated predictors. In agreement with these results, we showed that EN performed well for both prediction and selection on our simulated data, and that multivariate EN performed best for prediction on grapevine experimental data.

It would be interesting to test whether EN still remains the main default method when applied to a population with a shorter LD, *e.g.*, a diversity panel as defined in Nicolas *et al.* (2016). Indeed, the ranking of methods is likely to depend not only on linkage disequilibrium and population size, but also on the genetic architecture of the traits of interest as well as the accuracy with which phenotypic values were obtained, and all these variables can interact with each other, but studying this was out of the scope of the current work.

Comparison between interval-mapping and penalized regression methods for genomic prediction:

As expected, IM methods performed poorly to predict accurate genotypic values when the QTL number was large. These conclusions are in agreement with previous studies (Figure 1 and Supplementary Figure S8), even though most implemented marker selection methods other than interval-mapping (Bernardo and Yu 2007; Lorenzana and Bernardo 2009; Mayor and Bernardo 2009; Olatoye *et al.* 2019). This confirms that for complex traits, genomic prediction should not be based only on QTLs detected by IM methods.

None of the penalized univariate methods performed optimally in all cases (Figures 1 and 3 and Supplementary Figure S8), as also found in the literature (Heslot *et al.* 2012; Riedelsheimer *et al.* 2012; Azodi *et al.* 2019). As shown by simulation, RR was better adapted to highly polygenic genetic architecture, whereas LASSO was better adapted to a few major QTLs. Moreover, in the case of many minor QTLs, RR was the most stable method across heritability values, as previously described for several traits and species (Heslot *et al.* 2012; Azodi *et al.* 2019). However, RR prediction accuracy dropped when the QTL number was too small, whereas EN still predicted as well as LASSO. EN was hence well adapted to various numbers and distributions of QTLs.

Multivariate vs univariate:

When the same heritability was simulated for both trait variables, no superiority of multivariate methods was observed, even when both traits had QTLs at the same positions (Figure 1 and Supplementary Figure S8). When different heritability values were simulated for the two traits, we observed a slight superiority of MTV_LASSO (resp. MTV_EN) over LASSO (resp. EN) only in the "same" and "major" configuration (with both traits sharing the same two QTLs) for the trait with small heritability (Supplementary Figure S11).

Other authors who tested multivariate GP on simulated data systematically applied different heritability values; they found a superiority of multivariate methods over univariate ones for the trait with the smallest heritability (Calus and Veerkamp 2011; Guo *et al.* 2014; Jiang *et al.* 2015; Dagnachew and Meuwissen 2019). However, all these studies were based on a smaller, more favorable, p/n ratio, a key component of high-dimensional models (Verzelen 2012). For example, in Jia and Jannink (2012), their 500 observations for 2020 predictors correspond to a ratio of ~4, compared to our 188 observations for 3961 predictors corresponding to a ratio of ~21. Indeed, parameters n and p are involved in the sample complexity function defined in Obozinski *et al.* (2011), which predicts the theoretical cases where MTV_LASSO is superior to its univariate counterpart in terms of variable selection. Accordingly, applying our methods to Jia and Jannink (2012) data allowed us to evidence a larger difference between univariate and multivariate LASSO than with our simulated data.

Unexpectedly, when reanalyzing the data simulated by Jia and Jannink (2012), we obtained lower prediction accuracy with our MTV_LASSO (Supplementary Figure S11) than they did with their multivariate BayesA (their Figure 1A). A similar result in a univariate setting was found by Guan and Stephens (2011), who compared BSVR (comparable to BayesA) and the LASSO. They found that BSVR had markedly higher power than LASSO. Moreover, the parameters of both BSVR [in Guan and Stephens (2011)] and BayesA [in Jia and Jannink (2012)] were estimated with a MCMC algorithm. No inner cross-validation was needed, hence the sample used to train the model was larger. This difference may explain why Figure 1A from Jia and Jannink (2012) shows better prediction accuracies for multi-trait models compared to their single-trait counterparts, although no confidence interval was displayed. Note that our RR prediction accuracies were close to those of their GBLUP (univariate and multivariate). In conclusion, prediction accuracy is affected both by the dimension of the problem (*i.e.*, n and p) and the method used (*i.e.*, Bayesian with MCMC or cross-validation).

For experimental data, we observed that MTV_LASSO (respectively MTV_EN) was superior to LASSO (resp. EN) for the five traits with the smallest heritability (Figure 3). Such this improvement suggests that MTV_LASSO (resp. MTV_EN) was able to borrow signals from the most heritable traits to the least heritable ones, likely because of a partially overlapping genetic architecture between these traits. This interpretation is reinforced by the fact that a QTL for low-H2 trait, *TE.WW*, colocalizes on chromosome 4 with QTLs for four high-to-moderate-H2 traits (*TrS_night.WD*, *DeltaBiomass.WW*, *DeltaBiomass.WD* and *TE.WD*). This improvement was not found in Jia and Jannink (2012), who also tested their methods on real pine data from Resende *et al.* (2012). These observations suggest that the number of traits analyzed (14 in our case and 2 in Jia and Jannink 2012 study) may also play a role in the gain in prediction accuracy of multivariate over univariate methods.

Furthermore, we simulated data with various levels of residual correlation among traits (0, 0.4, and 1) but this did not significantly change prediction results (data not shown). A more detailed methodological analysis is out of the scope of the current work.

Comparison between interval-mapping and penalized regression methods for QTL detection:

Comparison with the ROC curve between IM and penalized regression methods for marker selection has not been extensively studied before. As expected, we found that IM methods are well adapted to detect a few major QTLs but not many minor QTLs (Figure 2). Yi *et al.* (2015) similarly compared the FDR and TPR reached by single marker analysis and different penalized regression methods, some of which being adapted to control FDR; they found contrasting results, depending on the criteria studied (modified version of TPR or FDR). However, they focused only on an association panel whereas we worked on a bi-parental population. Other authors (Cho *et al.* 2010; Li and Sillanpää 2012; Waldmann *et al.* 2013) successfully applied LASSO or EN for performing GWAS, but without comparing IM and penalized methods for QTL detection. Moreover, we found that penalized methods could be as good at marker selection as IM methods, and even far better when there were many minor QTLs. Among the penalized methods we compared, none clearly outperformed the others for marker selection in all configurations.

Multivariate vs univariate:

As MTV_LASSO selects one predictor for all traits, its superiority over univariate LASSO depends on QTL distribution across traits, notably on the amount of genetic basis shared by the traits (Obozinski *et al.* 2011). However, as for prediction, we showed that MTV_LASSO performance was not different whether QTLs were at the same or at different positions across traits (Figure 2). Nevertheless, we observed that MTV_LASSO was slightly better than LASSO when many QTLs were simulated.

SPRING had never been evaluated before for its quality of predictor selection. As for prediction, SPRING showed unstable results across our simulation replicates and hyper-parameter values. However, for the ROC curve, we did not include predictor structure in the model, which may have hampered marker selection quality.

Efficient default method for both QTL detection and genomic prediction:

IM methods were designed for marker selection; hence they are not expected to be optimal for prediction, as confirmed in our study. Among penalized regression methods, some may be better at prediction than at marker selection, and vice versa. For example, our results showed that EN performed well across several configurations for both aims. Some methods such as SPRING are specially adapted to handle both purposes but this method produced too variable results for either prediction or QTL detection. However, SPRING is a recent method that still can be improved to correct this drawback.

New penalized regression methods are continuously being developed. In particular, graph structured sparse subset selection (Grass) recently proved to outperform existing methods for both prediction and predictor selection, thanks to a *L0* regularization that limits the number of nonzero coefficients in the model (Do *et al.* 2020). It could be tested on our data when available. Moreover, multivariate methods are presented as being more

efficient at using the whole signal in the data, whether for marker selection (Inouye *et al.* 2012) or prediction (Jia and Jannink 2012; Guo *et al.* 2014), but our results revealed no systematic advantage of multivariate methods over univariate ones for both aims.

Using penalized methods for both marker selection and genomic prediction requires adapted hyper-parameter values. For EN, LASSO, and SPRING, the λ value controls sparsity (*e.g.*, the number of selected markers). Thus, the optimal λ value might not be the same if the aim is to limit FPR or to maximize predictive ability (Li and Sillanpää 2012). For prediction, we traditionally use cross-validation to tune hyper-parameters by minimizing MSE. For marker selection, there is no direct equivalence. That is why we tested extensions of these methods (mFDR and SS) that control sparsity for robust marker selection; both proved efficient in selecting the most relevant markers.

In order to shed light on the link between prediction accuracy and marker selection, we plotted the prediction accuracy at each point of the ROC curve for EN and EN.mFDR against FPR for minor configurations (with 50 simulated QTLs) (Supplementary Figure S22). For EN, we showed that prediction accuracy reached its maximum when FPR was below 0.05. Then, FPR increased while prediction accuracy decreased until it reached a plateau. This means that prediction quality is intimately linked to selection quality, especially at low heritability. For EN.mFDR, FPR always stayed below 0.015 but prediction accuracy was lower.

As a consequence, as an efficient default method, we advise at this stage to apply EN for performing genomic prediction, and its extension EN.mFDR for performing sparser marker selection.

Genetic determinism and prediction of grapevine response to water deficit:

Based on experimental data on the Syrah × Grenache progeny (new genotypic data and already published phenotypic data), we compared the same methods as above for both prediction and marker selection. To the best of our knowledge, grapevine GP within a bi-parental family has only been applied to a limited number of traits, with very few methods and never multivariate GP. Fodor *et al.* (2014) studied GP in grapevine with simulated data on a diverse structured population; they tested RR-BLUP, Bayesian Lasso, and a combination of marker selection and RR. Viana *et al.* (2016a) used an inter-specific grapevine bi-parental population. They predicted cluster and berry phenotypes (number and length of clusters, number, and weight of berries, juice pH, titrable acidity) with RR-BLUP and Bayesian LASSO applied to table grape breeding. In addition to yielding further insights into method comparison beyond those obtained by simulation, our study brought valuable novel biological knowledge about grapevine water use under different watering conditions. Indeed, new methods and the new SNP genetic map allowed us to find novel QTLs, as compared to those previously detected with the same phenotypic data (Coupel-Ledru *et al.* 2014, 2016). Our study also provides novel results of practical interest to grapevine breeders. We showed what predictive ability they can expect for drought-related traits within a progeny: here, always higher than 0.3, and up to more than 0.65 for some traits. Even though these traits are difficult to phenotype, they correspond to crucial breeding targets in the context of climate change. Our results may help motivate their phenotyping in the training panels of breeding programs.

Predictive ability and genetic architecture:

Among univariate penalized methods, RR generally had equivalent or better predictive ability compared to LASSO. For the traits with the largest discrepancy between RR and LASSO, this suggests that trait variability was rather due to many minor QTLs than to a few major ones. On the other hand, for a few traits, *e.g.*, *PsiM.WD*, *DeltaPsi.WD*, and *TE.WW*, predictive abilities with sparse methods (s, LASSO, and IM methods) were better than with RR, suggesting a genetic architecture with few major QTLs rather than an infinitesimal one in those cases.

Finally, while not considered by the penalized methods used, nonadditive genetic effects such as epistasis could be involved. We, therefore, tested the superiority of LASSO.GB over LASSO. Extreme Gradient Boosting methods are indeed among the best machine learning methods (Chen and Guestrin 2016). LASSO.GB did not markedly increase predictive ability on experimental data (Figure 3). However, we cannot exclude that this might be due to a poor optimization of Extreme Gradient Boosting parameters or to an insufficient number of observations to correctly fit the model. We also tested if coding differently the design matrix to estimate dominance genetic effects improved predictive ability and it was not the case (data not shown).

Candidate gene analysis:

The thorough methodology deployed for candidate genes analysis allowed us not only to retrieve a list of genes potentially underlying the QTLs of interest, but also to classify them based on their function and expression to point at the most likely candidates. We focused on the highly reliable QTL detected on chromosome 4 for *TrS_night*, *TE*, and *DeltaBiomass*. *TrS_night* QTL was previously described as a promising target for marker-assisted selection, as alleles limiting night-time transpiration also favor plant growth, resulting in a doubly, beneficial impact on improving TE (Coupel-Ledru *et al.* 2016). Moreover, this QTL was found using seven methods. Among the plethora of integrated functions represented within the list of annotated genes underlying this QTL, we show here that a subset of more likely candidates can be defined as possibly related to the traits of interest. On the one hand, these include genes related to broad-sense hydraulics and water loss, with a possible direct impact on *TrS_night*: seven genes involved in stomatal development, nine involved in stomatal opening—sometimes through the abscicic acid signaling pathway—, one related to xylem development and one to trichome development (Supplementary Table S21). One of these genes, the trihelix transcription factor GT-2 (*Vitvi04g01604*), was specifically shown to impact transpiration and TE in *Arabidopsis* by acting as a negative regulator of stomatal density. On the other hand, 27 genes in the list are directly related to growth, development, or photosynthesis, hinting to a possible direct impact on *DeltaBiomass*. A histidine kinase 1 (*Vitvi04g01483*) may be a particularly interesting candidate for its multiple roles in *Arabidopsis* in ABA signaling, stomatal development, and plant growth, hence potentially simultaneously acting on both components of *TE*. Both these likely candidates were often highly expressed in grapevine leaves. More precise analyses of these candidate genes, including functional genomics work and possibly gene editing, will now be necessary to identify the causative polymorphisms under these new QTLs.

Conclusions

Rather than decoupling genomic prediction from the identification of major QTLs, we argue for the need to pursue both goals jointly. Indeed, they provide complementary information on the genetic architecture of the target traits, as well the key

underlying functions. Our study provided encouraging findings for further implementing genomic prediction in grapevine breeding programs. Applied to both simulated and 14 experimental traits, univariate and multivariate Elastic Net proved to be efficient for both goals, followed by mFDR control for the robust identification of QTLs. Moreover, of interest to plant biologists seeking to understand the response to water stress, our results highlighted several candidate genes underlying integrated traits such as night-time transpiration, TE, and biomass production. For some, their putative functions suggest causal links with stomatal functioning, trichome development, or the ABA pathway.

Code availability

All software we used was free and open-source and most analyses were done with R (R Core Team 2020), notably graphs that were created using the ggplot2 package (Wickham 2016), version 3.3.2. All R scripts used for analysis, *i.e.* genetic mapping, simulation, phenotypic analysis, prediction, and QTL detection, are available online at https://doi.org/10.15454/NOUQY2. Many of the custom functions we used are available in a package for reproducibility purposes, R/rutilstimflutre (Flutre 2019).

Author contribution statement

T.F., A.D., and L.L.C. conceived the idea of the study and contributed to funding acquisition; A.C.L. and T.S. obtained the phenotypic data used in this work; C.B. and A.D. carried out the analysis of the genotypic data, with input data previously analyzed by T.F.; T.F. and J.C. performed preliminary multivariate genomic prediction analysis; A.C.L. and T.S. interpreted the results from the candidate gene analysis and wrote the corresponding parts in the manuscript; PT is the PhD supervisor of CB; CB wrote the original draft, which was reviewed and edited by all authors. All authors read and approved the final manuscript.

Ethical standards

The authors declare that the experiments comply with the current laws of the country in which they were carried out.

Data availability

Genotypic data, as well as the genetic map, are available online at https://doi.org/10.15454/QEDX2V. Raw phenotypic data are available in a second online repository at https://doi.org/10.15454/YTRKV6.

Supplementary material is available at G3 online.

Acknowledgments

We thank Marie Perrot-Dockès for her help concerning penalized regression methods, the South Green platform for computing facilities, notably Bertrand Pitollat, the GREAT platform, notably Amandine Velt and Camille Rustenholz, and Philippe Châtelet for his useful suggestions on the text.

Funding

We acknowledge the funding of the SelVi and FruitSelGen projects (BAP department and Selgen metaprogram of INRAE). Partial funding of the PhD was provided by ANRT (Association nationale recherche technologie, grant number 2018/0577), IFV and Inter-Rhône.

Conflicts of interest

None declared.

Literature cited

Adam-Blondon A-F, Bernole A, Faes G, Lamoureux D, Pateyron S, *et al.* 2005. Construction and characterization of BAC libraries from major grapevine cultivars. Theor Appl Genet. 110:1363–1371.

Arlot S, Lerasle M. 2016. Choice of V for V-fold cross-validation in least-squares density estimation. J Mach Learn Res. 17:7256–7305.

Arruda MP, Lipka AE, Brown PJ, Krill AM, Thurber C, *et al.* 2016. Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). Mol Breed. 36:84.

Azodi CB, Bolger E, McCarren A, Roantree M, Campos G. D L, *et al.* 2019. Benchmarking parametric and machine learning models for genomic prediction of complex traits. G3 (Bethesda)). 9: 3691–3702.

Bates D, Mächler M, Bolker B, Walker S. 2014. Fitting linear mixed-effects models using lme4. arXiv. 1406:5823.[stat] arXiv: 1406.5823.

Bernardo R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci. 48: 1649–1664.

Bernardo R, Yu J. 2007. Prospects for genomewide selection for quantitative traits in maize. Crop Sci. 47:1082–1090.

Breheny PJ. 2019. Marginal false discovery rates for penalized regression models. Biostatistics. 20:299–314.

Broman KW, Wu H, Sen C, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics. 19:889–890.

Calus MP, Veerkamp RF. 2011. Accuracy of multi-trait genomic selection using different methods. Genet Sel Evol. 43:26.

Canaguier A, Grimplet J, Gaspero GD, Scalabrin S, Duchêne E, *et al.* 2017. A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). Genom Data. 14:56–62.

Chen T, Guestrin C. 2016. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16. p. 785–794. arXiv: 1603.02754.

Chiquet J, Mary-Huard T, Robin S. 2017. Structured regularization for conditional Gaussian graphical models. Stat Comput. 27: 789–804.

Cho S, Kim K, Kim YJ, Lee J-K, Cho YS, *et al.* 2010. Joint identification of multiple genetic variants via elastic-net variable selection in a genome-wide association analysis: identifying multiple variants via EN. Ann Hum Genet. 74:416–428.

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD. 2004. Breeding for high water-use efficiency. J Exp Bot. 55:2447–2460.

Coupel-Ledru A, Lebon C, Christophe A, Doligez A, Cabrera-Bosquet L, *et al.* 2014. Genetic variation in a grapevine progeny (*Vitis vinifera* L. cvs Grenache×Syrah) reveals inconsistencies between maintenance of daytime leaf water potential and response of transpiration rate under drought. J Exp Bot. 65:6205–6218.

Coupel-Ledru A, Lebon E, Christophe A, Gallo A, Gago P, *et al.* 2016. Reduced nighttime transpiration is a relevant breeding target for high water-use efficiency in grapevine. Proc Natl Acad Sci USA. 113:8963–8968.

Cros D, Denis M, Sánchez L, Cochard B, Flori A, *et al.* 2015. Genomic selection prediction accuracy in a perennial crop: case study of oil palm (*Elaeis guineensis* Jacq.). Theor Appl Genet. 128:397–410.,

Crossa J, Pérez-Rodrìguez P, Cuevas J, Montesinos-López O, Jarquìn D, *et al.* 2017. Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci. 22:961–975.

Dagnachew B, Meuwissen T. 2019. Accuracy of within-family multi--trait genomic selection models in a sib-based aquaculture breeding scheme. Aquaculture. 505:27–33.

de los Campos G, Hickey JM, Pong-Wong R, Daetwyler HD, Calus MPL. 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. Genetics. 193:327–345.

Desta ZA, Ortiz R. 2014. Genomic selection: genome-wide prediction in plant improvement. Trends Plant Sci. 19:592–601.

Do H, Cheon M-S, Kim SB. 2020. Graph structured sparse subset selection. Inform Sci. 518:71–94.

Dodd LE, Pepe MS. 2003. Partial AUC estimation and regression. Biometrics. 59:614–623.

Doligez A, Bertrand Y, Dias S, Grolier M, Ballester J-F, *et al.* 2010. QTLs for fertility in table grape (*Vitis vinifera* L.) Tree Genet Genomes. 6: 413–422.

Doligez A, Bertrand Y, Farnos M, Grolier M, Romieu C, *et al.* 2013. New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). BMC Plant Biol. 13:217.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, *et al.* 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One. 6:e19379.

Emanuelli F, Battilana J, Costantini L, Cunff LL, Boursiquot J-M, *et al.* 2010. A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.). BMC Plant Biol. 10:241.

Ferrão LFV, Ferrão RG, Ferrão MAG, Fonseca A, Carbonetto P, *et al.* 2019. Accurate genomic prediction of *Coffea canephora* in multiple environments using whole-genome statistical models. Heredity (Edinb). 122:261–275.

Fischer BM, Salakhutdinov I, Akkurt M, Eibach R, Edwards KJ, *et al.* 2004. Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. Theor Appl Genet. 108:501–515.

Flutre T. 2019. rutilstimflutre: Timothee Flutre's personal R package. Zenodo. 10.5281/zenodo.3580267.

Flutre T, Cunff L, Fodor A, Launay A, Romieu C, *et al.* 2020. Genome-wide association and prediction studies using a grapevine diversity panel give insights into the genetic architecture of several traits of interest. bioRxiv. 2020.09.10.290890.

Fodor A, Segura V, Denis M, Neuenschwander S, Fournier-Level A, *et al.* 2014. Genome-wide prediction methods in highly diverse and heterozygous species: proof-of-concept through simulation in Grapevine. PLoS One. 9:e110436.

Fournier-Level A, Cunff LL, Gomez C, Doligez A, Ageorges A, *et al.* 2009. Quantitative genetic bases of anthocyanin variation in Grape (*Vitis vinifera* L. ssp. sativa) Berry: a quantitative trait locus to quantitative trait nucleotide integrated study. Genetics. 183: 1127–1139.

Friedman J, Hastie T, Tibshirani R. 2010. Regularization paths for generalized linear models via coordinate descent. J Stat Softw. 33:1–22.

Goddard ME, Hayes BJ. 2007. Genomic selection. J Anim Breed Genet. 124:323–330. _eprint:https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1439-0388.2007.00702.x

Gois I, Borém A, Cristofani-Yaly M, de Resende M, Azevedo C, *et al.* 2016. Genome wide selection in Citrus breeding. Genet Mol Res. 15.

Guan Y, Stephens M. 2011. Bayesian variable selection regression for genome-wide association studies and other large-scale problems. Ann Appl Stat. 5:1780–1815.

Guo G, Zhao F, Wang Y, Zhang Y, Du L, *et al.* 2014. Comparison of single-trait and multiple-trait genomic prediction models. BMC Genet. 15:30.

Habier D, Fernando RL, Dekkers JCM. 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics. 177:2389–2397.

Haley CS, Knott SA. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity (Edinb). 69:315–324.

Hastie T, Qian J. 2016. Glmnet vignette.

Hastie T, Tibshirani R, Friedman J. 2009. The Elements of Statistical Learning. Springer New York, New York, NY: Springer Series in Statistics.

Hayashi T, Iwata H. 2013. A Bayesian method and its variational approximation for prediction of genomic breeding values in multiple traits. BMC Bioinformatics. 14:34.

Heffner EL, Jannink J-L, Sorrells ME. 2011. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. Plant Genome. 4:65–75.

Heffner EL, Lorenz AJ, Jannink J-L, Sorrells ME. 2010. Plant breeding with genomic selection: gain per unit time and cost. Crop Sci. 50: 1681–1690.

Heffner EL, Sorrells ME, Jannink J-L. 2009. Genomic selection for crop improvement. Crop Sci. 49:1.

Henderson CR, Quaas RL. 1976. Multiple trait evaluation using relatives' records. J Anim Sci. 43:1188–1197.

Heslot N, Yang H-P, Sorrells ME, Jannink J-L. 2012. Genomic selection in plant breeding: a comparison of models. Crop Sci. 52: 146–160.

Hoerl AE, Kennard RW. 1970. Ridge regression: biased estimation for nonorthogonal problems. Technometrics. 12:55–67.

Hofner B, Hothorn T. 2017. stabs: Stability Selection with Error Control. R package version R package version 0.6-3, https://CRAN.R-project.org/package=stabs.

Huang Y-F, Doligez A, Fournier-Level A, Cunff LL, Bertrand Y, *et al.* 2012. Dissecting genetic architecture of grape proanthocyanidin composition through quantitative trait locus mapping. BMC Plant Biol. 12:30.

Inouye M, Ripatti S, Kettunen J, Lyytikäinen L-P, Oksala N, *et al.* 2012. Novel loci for metabolic networks and multi-tissue expression studies reveal genes for atherosclerosis. PLoS Genet. 8: e1002907.

Jia Y, Jannink J-L. 2012. Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics. 192:1513–1522.

Jiang C, Zeng ZB. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics. 140:1111–1127.

Jiang J, Zhang Q, Ma L, Li J, Wang Z, *et al.* 2015. Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. Heredity (Edinb). 115:29–36.

Kao C-H, Zeng Z-B, Teasdale RD. 1999. Multiple interval mapping for quantitative trait loci. Genetics. 152:1203–1216.

Kemper KE, Bowman PJ, Hayes BJ, Visscher PM, Goddard ME. 2018. A multi-trait Bayesian method for mapping QTL and genomic prediction. Genet Sel Evol. 50:10.

Korol AB, Ronin YI, Kirzhner VM. 1995. Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics. 140:1137–1147.

Kuhn M. 2008. Building predictive models in R using the caret package. J Stat Soft. 28:1–26.

Kumar S, Chagné D, Bink MCAM, Volz RK, Whitworth C, *et al.* 2012. Genomic selection for fruit quality traits in apple (Malus×domestica Borkh.). PLoS One. 7:e36674.

Kumar S, Hilario E, Deng CH, Molloy C. 2020. Turbocharging introgression breeding of perennial fruit crops: a case study on apple. Hortic Res. 7:47.

Kumar S, Kirk C, Deng CH, Shirtliff A, Wiedow C, *et al.* 2019. Marker-trait associations and genomic predictions of interspecific pear (Pyrus) fruit characteristics. Sci Rep. 9:9072.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: tests in linear mixed effects models. J Stat Soft. 82: 1–26.

Kwong QB, Ong AL, Teh CK, Chew FT, Tammi M, *et al.* 2017. Genomic selection in commercial perennial crops: applicability and improvement in oil palm (*Elaeis guineensis* Jacq). Sci Rep. 7: 2872.

Lado B, Vázquez D, Quincke M, Silva P, Aguilar I, *et al.* 2018. Resource allocation optimization with multi-trait genomic prediction for bread wheat (*Triticum aestivum* L.) baking quality. Theor Appl Genet. 131:2719–2731.

Lander ES, Botstein D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 121: 185–199.

Li Z, Sillanpää MJ. 2012. Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection. Theor Appl Genet. 125:419–435.

Liu X, Hu X, Li K, Liu Z, Wu Y, *et al.* 2020. Genetic mapping and genomic selection for maize stalk strength. BMC Plant Biol. 20: 196.

Lorenzana RE, Bernardo R. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor Appl Genet. 120:151–161.

Marguerit E, Brendel O, Lebon E, Leeuwen CV, Ollat N. 2012. Rootstock control of scion transpiration and its acclimation to water deficit are controlled by different genes. New Phytol. 194: 416–429.

Mason L, Baxter J, Bartlett PL, Frean MR. 1999. Boosting algorithms as gradient descent. Adv Neural Inform Process Syst. 12:7.

Mayer DG, Butler DG. 1993. Statistical validation. Ecol Model. 68: 21–32.

Mayor PJ, Bernardo R. 2009. Genomewide selection and marker-assisted recurrent selection in doubled Haploid versus F2 populations. Crop Sci. 49:1719–1725.

McClish DK. 1989. Analyzing a portion of the ROC curve. Med Decis Making. 9:190–195.

Meinshausen N, Buhlmann P. 2009. Stability selection. J R Statist Soc B. 72:417-473.

Mejìa N, Soto B, Guerrero M, Casanueva X, Houel C, *et al.* 2011. Molecular, genetic and transcriptional evidence for a role of *VvAGL11* in stenospermocarpic seedlessness in grapevine. BMC Plant Biol. 11:57.

Meuwissen TH, Goddard ME. 2004. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. Genet Sel Evol. 36:261–279.

Migicovsky Z, Sawler J, Gardner KM, Aradhya MK, Prins BH, *et al.* 2017. Patterns of genomic and phenomic diversity in wine and table grapes. Hortic Res. 4:17035.

Muranty H, Troggio M, Sadok IB, Rifaï MA, Auwerkerken A, *et al.* 2015. Accuracy and responses of genomic selection on key traits in apple breeding. Hortic Res. 2:15060.

Nicolas SD, Péros J-P, Lacombe T, Launay A, Paslier M-CL, *et al.* 2016. Genetic diversity, linkage disequilibrium and power of a large grapevine (*Vitis vinifera* L) diversity panel newly designed for association studies. BMC Plant Biol. 16:74.

Obozinski G, Wainwright MJ, Jordan MI. 2011. Support union recovery in high-dimensional multivariate regression. Ann Statist. 39: 1–47. arXiv: 0808.0711.

Olatoye MO, Clark LV, Wang J, Yang X, Yamada T, *et al.* 2019. Evaluation of genomic selection and marker-assisted selection in Miscanthus and energycane. Mol Breed. 39:171.

Piñeiro G, Perelman S, Guerschman JP, Paruelo JM. 2008. How to evaluate models: observed vs. predicted or predicted vs. observed? Ecol Model. 216:316–322.

Piepho H-P, Möhring J. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics. 177: 1881–1888.

Price AH. 2006. Believe it or not, QTLs are accurate!. Trends Plant Sci. 11:213–216.

Pszczola M, Veerkamp RF, de Haas Y, Wall E, Strabel T, *et al.* 2013. Effect of predictor traits on accuracy of genomic breeding values for feed intake based on a limited cow reference population. Animal. 7:1759–1768.

R Core Team 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.r-project.org/.

Rastas P. 2017. Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. Bioinformatics. 33: 3726–3732.

Resende MFR, Muñoz P, Resende MDV, Garrick DJ, Fernando RL, *et al.* 2012. Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). Genetics. 190: 1503–1510.

Riedelsheimer C, Technow F, Melchinger AE. 2012. Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. BMC Genomics. 13:452.

Runcie D, Cheng H. 2019. Pitfalls and remedies for cross validation with multi-trait genomic prediction methods. G3 (Bethesda). 9: 3727–3741.

Schwander F, Eibach R, Fechter I, Hausmann L, Zyprian E, *et al.* 2012. Rpv10: a new locus from the Asian Vitis gene pool for pyramiding downy mildew resistance loci in grapevine. Theor Appl Genet. 124:163–176.

Swets JA, Pickett RM, Whitehead SF, Getty DJ, Schnur JA, *et al.* 1979. Assessment of diagnostic technologies. Science. 205: 753–759.

Thompson R, Meyer K. 1986. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. Livestock Prod Sci. 15:299–313.

Tibshirani R. 1996. Regression shrinkage and selection via the lasso. J R Statist Soc B. 58:267–288.

Velazco JG, Jordan DR, Mace ES, Hunt CH, Malosetti M, *et al.* 2019. Genomic prediction of grain yield and drought-adaptation capacity in sorghum is enhanced by multi-trait analysis. Front Plant Sci. 10:997.

Verzelen N. 2012. Minimax risks for sparse regressions: ultra-high dimensional phenomenons. Electron J Statist. 6:38–90.

Vezzulli S, Doligez A, Bellin D. 2019a. Molecular mapping of grapevine genes. In: D Cantu, MA Walker, editors. The Grape Genome. Cham: Compendium of Plant Genomes Springer International Publishing. p. 103–136.

Vezzulli S, Zulini L, Stefanini M. 2019b. Genetics-assisted breeding for downy/powdery mildew and phylloxera resistance at fem. BIO Web Conf. 12:01020.

Viana AP, Resende M. D V D, Riaz S, Walker MA. 2016a. Genome selection in fruit breeding: application to table grapes. Sci Agric (Piracicaba, Braz). 73:142–149.

Viana JMS, e Silva FF, Mundim GB, Azevedo CF, Jan HU. 2016b. Efficiency of low heritability QTL mapping under high SNP density. Euphytica. 213:13.

Waldmann P, Mészáros G, Gredler B, Fuerst C, Sölkner J. 2013. Evaluation of the lasso and the elastic net in genome-wide association studies. Front Genet. 4:270.

Wang F, Mukherjee S, Richardson S, Hill SM. 2020. High-dimensional regression in practice: an empirical study of finite-sample prediction, variable selection and ranking. Stat Comput. 30:697–719.

Wang Y, Mette MF, Miedaner T, Gottwald M, Wilde P, *et al.* 2014. The accuracy of prediction of genomic selection in elite hybrid rye populations surpasses the accuracy of marker-assisted selection and is equally augmented by multiple field evaluation locations and test years. BMC Genomics. 15:556.

Welter LJ, Göktürk-Baydar N, Akkurt M, Maul E, Eibach R, *et al.* 2007. Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). Mol Breed. 20:359–374.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag.

Yi H, Breheny P, Imam N, Liu Y, Hoeschele I. 2015. Penalized multimarker vs. single-marker regression methods for genome-wide association studies of quantitative traits. Genet Sec Invest. 199: 205–222.

Zou H, Hastie T. 2005. Regularization and variable selection via the elastic net. J R Statist Soc B. 67:301–320.

Communicating editor: J-L. Jannink

2.4 Application of some GP methods on field data

2.4.1 Material and methods

Experimental design and phenotypes

Field trial was located in Villeneuve-lès-Maguelone (Occitanie, South of France), with two complete randomized blocks planted in 2003 (Doligez et al., 2013). Within a plot, five replicates of the same genotype were planted in a row. A mild water stress was applied during two years (2010 and 2011) on one block with inter-row grassing, while the other block was irrigated and thus underwent no water stress.

Several phenotypes related to development and production were measured in 2010 and 2011: pruning weight (**pruw**), the total weight of pruned canes per vine (in kg); **verday**, the véraison date (onset of ripening, in number of days since January 1st); **nbclu**, the number of clusters per vine at maturity (defined as véraison date + 35 days); **yield**, the total fruit production per vine at maturity (in kg); mbw, the mean berry weight (over 200 random berries), **fertility**, the number of clusters divided by the number of shoots. One phenotype related to drought tolerance was also measured the same years: $\delta 13C$ (**d13C**), which is the relative amount of carbon 13 over carbon 12 isotopes, measured in the must at maturity. This measure reflects the intrinsic water-use efficiency (WUE) during the growing season (Farquhar & Richards, 1984).

Processing of phenotypes

For each of the seven phenotypes, I derived three traits from raw phenotypic data, corresponding to three modalities: without water stress (well-watered modality, referred to as *WW*), with water stress (water deficit modality, referred to as *WD*) and response to stress (normalized difference $\frac{WW-WD}{WW}$, referred to as *diff.norm*).

Then, for each of the 21 resulting traits, I applied a transformation to raw data when distribution was skewed. Then, I fitted a mixed model, derived broad-sense heritability from variance components and used the best linear unbiased predictor (BLUP) of offspring genotypic values for implementing GP, following the same procedure as in Article I. The fitted model was: $y_{ij} = \mu + \underline{geno_i} + year_j + rank_i + location_i + \epsilon_{ij}$. For *diff.norm* modality, *rank* and *location* couldn't be estimated as the difference is computed from vines at different field coordinates.

Genomic prediction

The same genotypic data as in Article I were used (3,961 SNP markers). The six GP methods applied and predictive ability measures are described in Article I. We chose a subset of GP methods which gave promising results in Article I.

2.4.2 Results and discussion

Broad-sense heritability

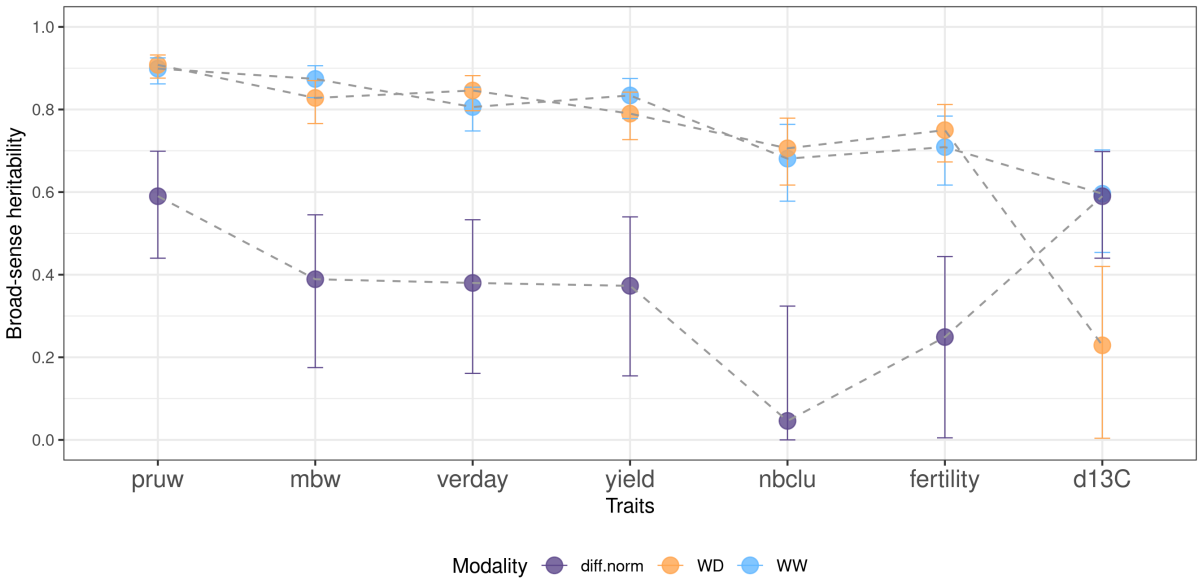


FIGURE 2.3 – Broad-sense heritability for the 21 traits.

For *WW* and *WD* modalities, broad-sense heritability was around 0.8 for all traits, except **d13C** for which it was about 0.6 and 0.2 for *WW* and *WD*, respectively (Figure 2.3). For all traits except **d13C**, computing the normalized difference between WW and WD led to a sharp decrease in heritability, which nearly reached 0 for **nbclu**. **d13C** was the only trait for which there was no difference in heritability between *diff.norm* and *WW*.

The genetic correlation between these traits under the two modalities was 0.06 for **d13C**, 0.69 for **fertility**, 0.83 for **mbw**, 0.71 for **nbclu**, 0.76 for **pruw**, 0.78 for **verday**and 0.76 for **yield**.

These results show that signal/noise ratio is lower for *diff.norm* modality. One reason may be that the the WW and WD phenotypic data came from two different vine plots, from two blocks and with no common coordinate in the field, which might make it more difficult to capture the genetic variance, or to correct for spatial heterogeneity. Another hypothesis could be that there was a low genotype-by-environment variance, as the genetic correlation was quite high between the two modalities, except for **d13C**. This is consistent with the low genetic correlation between *d13C.WW* and *d13C.WD*, which suggest specific response to drought, which is modelled with *diff.norm* modality. The fact that *d13C.diff.norm*, the trait most tightly related to drought tolerance, had a heritability value of 0.6 is encouraging for studying response to drought in the field.

Predictive ability

Strikingly, PA for *diff.norm* modality and **d13C** trait were close to zero, while PA values for other traits were all close to 0.6 (Figure 2.4). PA values for *WW* and *WD* modalities were mostly consistent with heritability values from Figure 2.3, except for **d13C**.

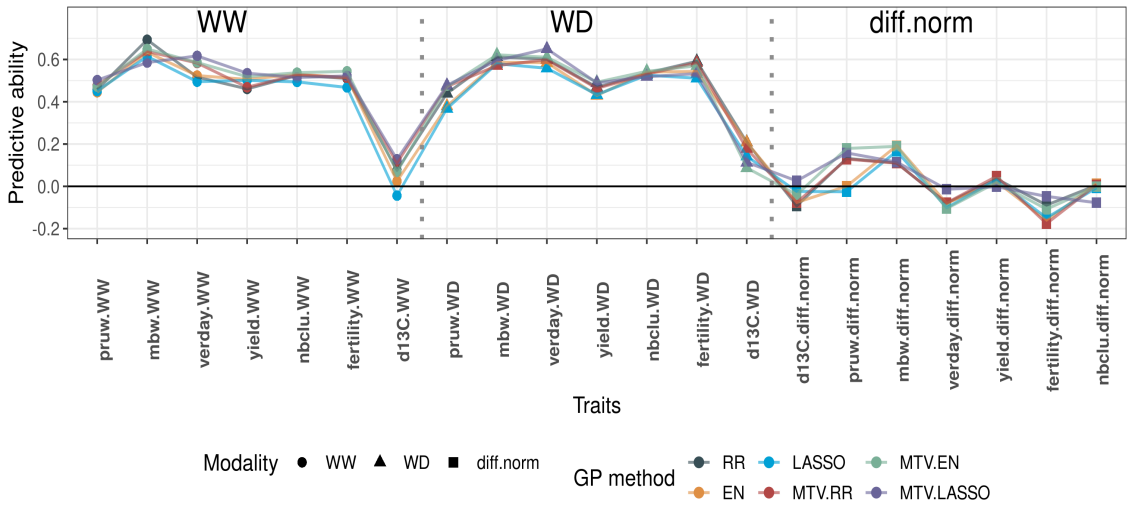


FIGURE 2.4 – Predictive ability of 6 GP methods for 21 traits (7 phenotypes x 3 modalities) in the bi-parental grapevine population SxG. WW: well-watered; WD: water deficit; diff.norm: normalized difference between WW and WD. pruw: pruning weight, mbw: mean berry weight, verday: véraison date, nbclu: number of clusters, d13C: carbon isotope discrimination.

The six GP methods had very close PA for the 21 traits studied (Figure 2.4). Closer inspection of the figure shows that EN and LASSO were the worst performing methods for several traits such as *d13C.WW*, *verday.WW*, *pruw.WD*, *verday.WD* or *pruw.diff.norm*. Mean PA per method was 0.331 for MTV_LASSO, 0.329 for MTV_EN, 0.320 for RR, 0.318 for MTV_RR, 0.303 for EN and 0.285 for LASSO. Overall, multivariate methods thus gave slightly better results than univariate ones.

Such a discrepancy for some traits or modalities between broad-sense heritability values and PA could be explained by non-additive effects that were not taken into account in our GP models. To assess this hypothesis, we could compute the narrow-sense heritability using molecular marker, in order to characterize the relative proportion of genetic additive variance.

Studying grapevine response to drought on field remains a challenge. This situation is closer to what is encountered by wine-growers but drought stress intensity is not controlled. Moreover, drought stress likely affects the plant for several years, and phenotypes were only measured during two years. Besides, using two vines at different field coordinates makes impossible the estimation of field heterogeneity. Instead, we could use more blocks as repetition within each modality.

2.5 Chapter general discussion

Interestingly, for single-modality trait, heritability values measured in PhenoArch phenotypic platform and in the field could be as high as 0.7. Moreover, range of PA values were mostly consistent with heritability in both experiments. Then, the ranking of methods, when compared based on PA, was nearly the same in the field as in the high-throughput phenotypic platform PhenoArch, with a slight superiority of multivariate methods over univariate ones.

Overall, heritability values for drought-related traits measured in PhenoArch were lower than those measured in the field. A possible explanation for this difference is that traits measured in PhenoArch are highly variable and difficult to measure and thus subjected to micro-environmental heterogeneity, despite a strong control of environmental conditions, which may hinder genetic signal. This heterogeneity has probably a smaller impact on integrative traits, such as those measured in the field.

However, results with data from field were not completely comparable with those obtained in Article I because we did not compute *diff.norm* modality. Genotypic correlations between water scenarios were around 0.4 for all traits in Figure S4 from Appendix B, we thus hypothesize that some GxE could be captured by computing *diff.norm* modality.

Overall, these findings suggest that genetic variance and heritability for phenotypes facing water deficit is moderate to high and that these traits could be quite accurately predicted using GP. As a perspective, we could compare genotype ranking between **d13C** measured in the field and other phenotypes measured in PhenoArch.

Nevertheless, these results were focused on a single bi-parental population, with a narrow genetic diversity explored.

Chapter 3

Across-population genomic prediction

3.1 Summary of the chapter

In this chapter, we aimed at implementing GP in a configuration more applicable in a breeding context, i.e., across-population. For that, we used two other populations: a diversity panel (N=277) as TS and a half-diallel (N=622) as TS and VS. PA was decomposed between predicting cross mean and predicting individuals within a cross. We also implemented the prediction of cross variance. Finally, we applied TS optimization and we quantified the relative effect of different parameters on PA, both for cross mean and individual prediction.

3.2 Article II: Across-population genomic prediction in grapevine opens up promising prospects for breeding

This article was submitted to Horticulture Research on 12^{th}, August 2021 and is available in bioRxiv. Supplementary materials can be found in Appendix C.

Across-population genomic prediction in grapevine opens up promising prospects for breeding

Charlotte Brault[1,2,3]; Vincent Segura[1,2]; Patrice This[1,2]; Loïc Le Cunff[1,2,3]; Timothée Flutre[4]; Pierre François[1,2]; Thierry Pons[1,2]; Jean-Pierre Péros[1,2]; Agnès Doligez*[1,2]

1: UMT Geno-Vigne® , IFV-INRAE-Institut Agro, F-34398 Montpellier, France

2: UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, F-34398 Montpellier, France

3: Institut Français de la Vigne et du Vin, F-34398 Montpellier, France

4: Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE – Le Moulon, 91190, Gif-sur-Yvette, France

Corresponding author: Agnès Doligez, UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, F-34398 Montpellier, France. E-mail: agnes.doligez@inrae.fr.

Abstract

Crop breeding involves two selection steps: choosing progenitors and selecting offspring within progenies. Genomic prediction, based on genome-wide marker estimation of genetic values, could facilitate these steps. However, its potential usefulness in grapevine (*Vitis vinifera* L.) has only been evaluated in non-breeding contexts mainly through cross-validation within a single population. We tested across-population genomic prediction in a more realistic breeding configuration, from a diversity panel to ten bi-parental crosses connected within a half-diallel mating design. Prediction quality was evaluated over 15 traits of interest (related to yield, berry composition, phenology and vigour), for both the average genetic value of each cross (cross mean) and the genetic values of individuals within each cross (individual values). Genomic prediction in these conditions was found useful: for cross mean, average per-trait predictive ability was 0.6, while per-cross predictive ability was halved on average, but reached a maximum of 0.7. Mean predictive ability for individual values within crosses was 0.26, about half the within-half-diallel value taken as a reference. For some traits and/or crosses, these across-population predictive ability values are promising for implementing genomic selection in grapevine breeding. This study also provided key insights on variables affecting predictive ability. Per-cross predictive ability was well predicted by genetic distance between parents and when this predictive ability was below 0.6, it was improved by training set optimization. For individual values, predictive ability mostly depended on trait-related variables (magnitude of the cross effect and heritability). These results will greatly help designing grapevine breeding programs assisted by genomic prediction.

Keywords: genomic prediction, grapevine, half-diallel, multi-parental population, diversity panel, across-population

Introduction

Breeding for perennial species is mostly based on phenotypic selection and is hindered by cumbersome field trials and the long generation time. Genomic prediction (GP), based on genome-wide prediction of genetic values [1], has been widely adopted in modern plant and animal breeding programs, for its superiority in terms of cost and time saved compared to traditional phenotypic selection, but also because it allows handling traits with complex genetic determinism. GP requires a model training step within a reference population, prior to model application to a target population of selection candidates [2]. In perennial crops, a universal population encompassing most of the species' genetic diversity could be particularly interesting as a training population to reduce phenotyping effort, since breeding cycle and juvenile phase are long.

Breeding schemes typically involve first the choice of parents (individuals to be crossed) and then the selection of offspring within crosses. GP is adapted both for predicting cross mean and for ranking genotypes within a cross (Mendelian sampling). These steps correspond to the components of the predictive ability (PA) of GP. It is indeed defined as the sum of cross mean and Mendelian sampling terms, as detailed in Werner et al. [3].

Under an additive framework, cross mean is expected to be the sum of the breeding values of parents, but some deviation may result from dominance or epistasis [4]. So far, a few studies only have investigated cross mean PA [5, 6, 7, 8], although none of them clearly investigated its influencing parameters.

In contrast, the prediction of genetic values within a cross (Mendelian sampling), has been widely studied, both with simulated and real data. Various parameters affecting PA have been pointed out, including the statistical method used [9], the composition and size of training and validation populations [10, 11], the trait genetic architecture and heritability [12, 13] and marker density [14]. Genetic relationship between the training and validation sets is known to strongly affect PA [15], with low or even sometimes negative accuracies for across-breed GP in animals [16]. This can be explained by the loss of linkage phase between the marker and QTL or by differences in linkage disequilibrium among populations [17]. Another explanation is the presence of specific allelic effects and allele frequencies, due to the genetic background [18]. All these effects are linked to genetic relationship. Some studies specifically derived deterministic equations to predict PA for across-population GP, based on genetic relationship and heritability (e.g., 19, 20, 21).

In grapevine (*Vitis vinifera* subsp. *vinifera*), very few authors have assessed the potential interest of GP. Viana et al. [22] investigated GP within a bi-parental population from a cross between an interspecific hybrid and a seedless table grape. Later, Migicovsky et al. [23] used a panel of 580 *V.vinifera* accessions to perform both GP and genome-wide association study (GWAS) for 33 phenotypes. More recently, Brault et al. [24] investigated GP within a bi-parental population from a cross between Syrah and Grenache. In a related study, Fodor et al. [25] had simulated a structured and highly diverse grapevine panel and bi-parental populations with parents originating from the panel. They applied GP and found little difference between PA values estimated within the panel or across populations. Finally, Flutre et al. [26] studied 127 traits with GWAS and GP within a diversity panel; they also applied across-population GP, but with 23 test offspring and for one trait only. Before genomic selection can be deployed in grapevine, evaluating PA across populations is thus crucially needed. In particular, PA should be evaluated with a diversity panel and a bi-parental progeny as training and validation sets, respectively, a

configuration much more likely to occur in actual breeding schemes than GP within the same population. As in grape, studies investigating across-population GP are also lacking in most clonally propagated crops.

The aim of this study was to assess across-population genomic PA and to provide a more thorough understanding of parameters affecting PA in a situation close to the one typically encountered in a breeding context, i.e. across populations, for a clonally propagated crop such as grapevine. Our study was based on phenotypic data for 15 traits, related to yield, berry composition, phenology and vigour, measured both in a diversity panel [27], and in a half-diallel with 10 bi-parental crosses. We assessed PA under three scenarios, first for cross mean, and then for Mendelian sampling term; the results provided keys to understand PA determinants in both cases. Finally, we implemented training population optimization to investigate under which conditions PA can be improved.

Results

Extent of genetic diversity within the half-diallel population

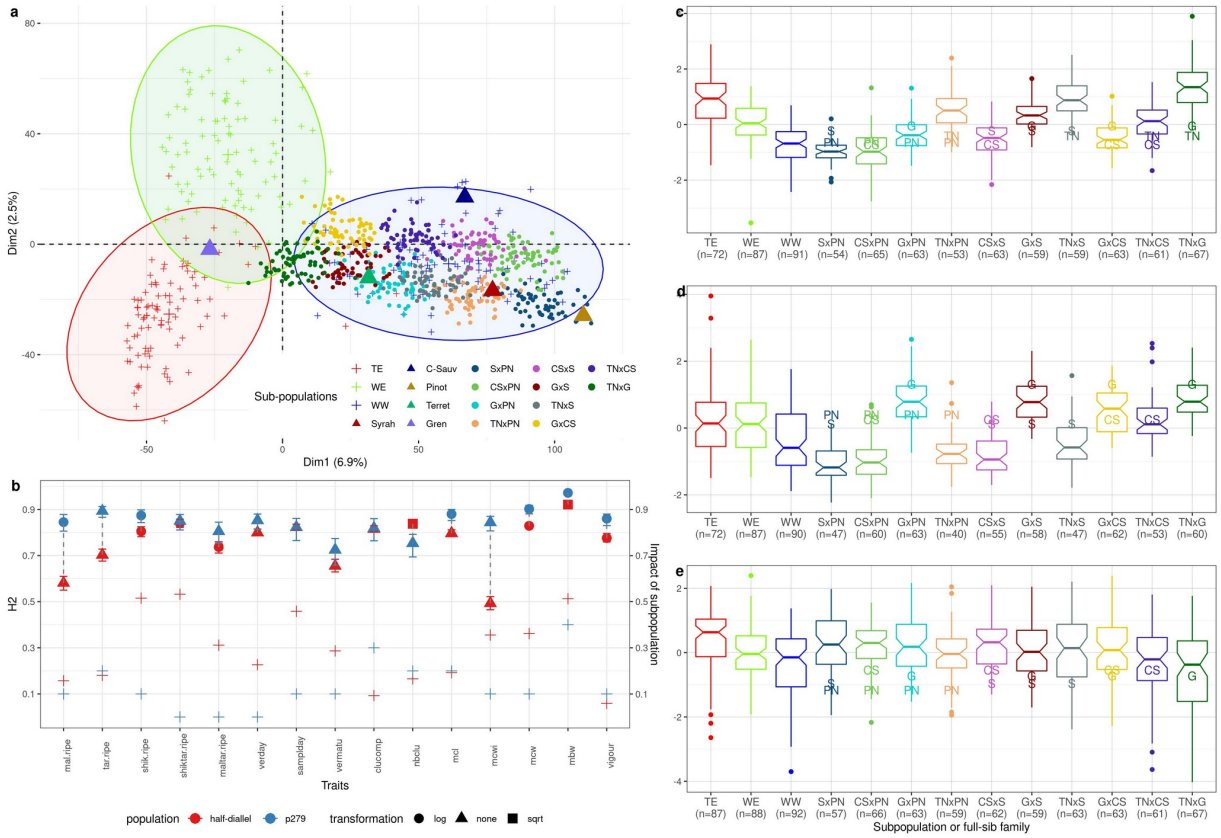


Figure 1: Description of the half-diallel, relative to the diversity panel. a: PCA of the diversity panel based on 32,894 SNPs with the 3 sub-populations distinguished by different colors, on which half-diallel progenies (dots) and parents (triangles) were projected. b: Broad-sense heritability estimates in the whole half-diallel (red) and in the diversity panel (blue) for the 15 traits studied (left axis), with shape corresponding to the transformation applied to raw data; the relative variance due to the cross effect and the R^2 of the subpopulation effect, for the half-diallel (red) and the diversity panel (blue), respectively, are also reported with '+' (right axis). c, d, e: genotypic value BLUP distribution in each subpopulation or progeny, for mean berry weight, mean cluster width and vigour, respectively; BLUPs for parents are indicated by their initial letters (Table S4). Number of genotypes per subpopulation/progeny is indicated below the subpopulation/progeny name. These traits were chosen to represent various levels of H^2 and relative importance of cross effect. BLUP distributions for all traits are presented in Figure S3.

We first evaluated the genetic variability of half-diallel crosses with respect to the diversity panel, through their projection on the first plane of a PCA based on genotypic data at 32,894 SNPs within the diversity panel. The half-diallel crosses were genetically close to the wine west (WW) subpopulation from the diversity panel (Figure 1a), which was expected, given that all half-diallel parents except Grenache are wine varieties from western Europe (Figure 1a, Figure S1). The half-diallel diversity covered the whole range of WW diversity, and progenies, all located exactly between their respective parents, were well separated from each other along the first two PCA axes (Figure 1a).

We then investigated broad-sense heritability values (H^2) for 15 traits related to yield, berry composition, phenology and vigour. Overall H^2 values were medium to high, ranging from 0.49 for **mcwi** in the half-diallel to 0.92 for **mbw** in the panel (Figure 1b; Table S1). Correlation between half-diallel and diversity panel heritability values was 0.31. Per-cross H^2 values for each trait varied among half-diallel crosses (Figure S2), which might result from the fairly small number of offspring per cross (from 64 to 70). Nevertheless, we observed a 0.68 correlation between overall and per-cross H^2. Mean cluster width displayed extreme variation in H^2 per cross (from 0.02 to 0.67). This might be due to the difficulty to phenotype this specific trait because of the presence of lateral wings in some individuals.

Within the half-diallel and for all traits, the *cross* effect was retained in the mixed model for genetic value estimation, but its magnitude with respect to the total genetic variance varied depending on the trait, ranging from less than 10% to ca. 50% (Figure 1b; Table S1). Depending on the trait or cross, the distribution of genotypic BLUPs varied widely (Figure 1c-e; Figure S3), some traits such as **vigour** being quite comparable among crosses, while others such as **mbw** or **mcwi** varied greatly. We also observed transgressive segregation within the half-diallel progenies (Figure 1c-e; Figure S3) for most traits and subpopulations. The 15 traits studied represented a large phenotypic diversity, structured among crosses (Figure S4).

Across-population genomic prediction in grapevine

Prediction of cross mean and Mendelian sampling within- and across-populations

Prediction of cross mean

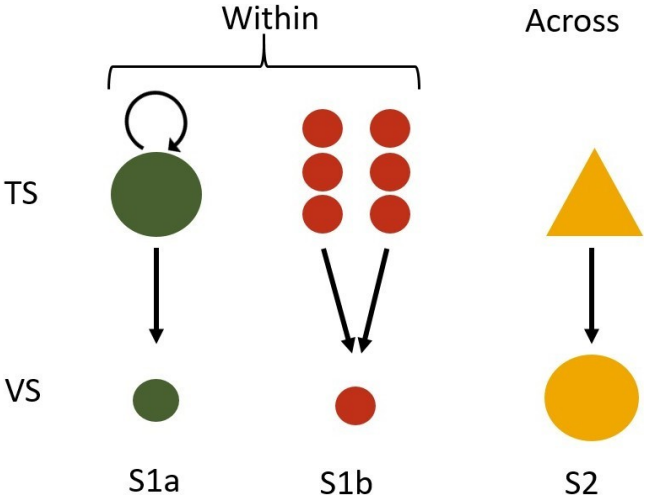


Figure 2: Schematic description of the three scenarios tested. TS: training set, VS: validation set. In scenario 1a, GP was applied within the half-diallel population with 10-fold cross-validation repeated 10 times. In scenario 1b, half-sib families from each parent were used separately as TS. In scenario 2, TS was the diversity panel. See details in Table S5.

We first implemented cross mean prediction, as if aiming to select parents for future crosses, selecting the method best adapted to genetic architecture between RR and LASSO (see Material and Methods). Predictive ability (PA) was assessed as Pearson's correlation between the observed mean genotypic value per half-diallel cross and the one predicted based on parental average genotypes (Table S2). Three scenarios were tested (Material and Methods, Figure 2): allelic effects estimated within the whole half-diallel (scenario 1a), in families with one parent in common (scenario 1b), or within the whole diversity panel (scenario 2).

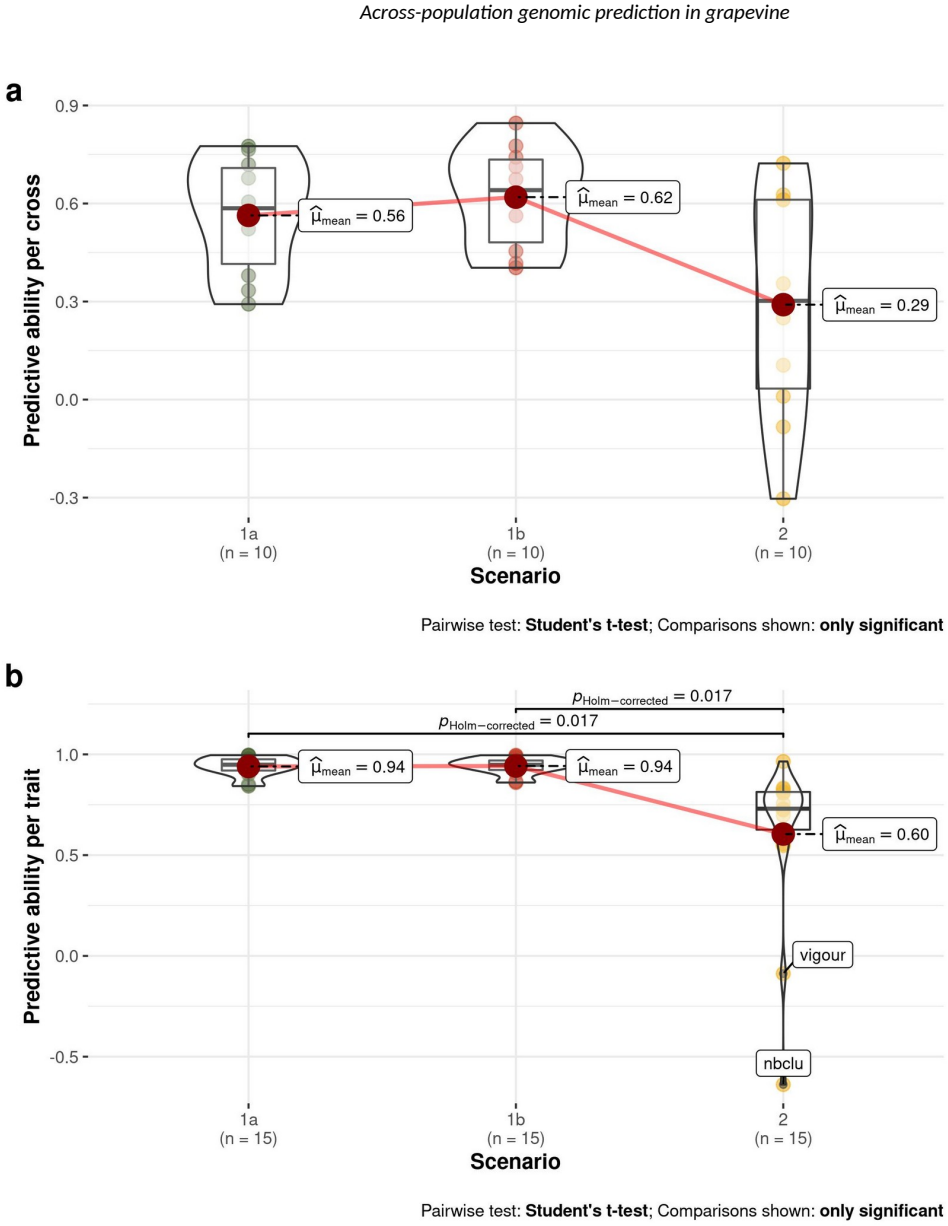


Figure 3: Boxplots of PA values for the three scenarios (1a: within whole half-diallel prediction; 1b: half-sib prediction within half-diallel; 2: across-population prediction with diversity panel as training set and each half-diallel cross as validation set). Each PA value was the best one obtained between RR and LASSO methods. Average PA is indicated next to each boxplot.
a: per-cross PA, b: per-trait PA.

In scenario 2, per-trait and per-cross predictive ability was lower and more variable than in scenarios 1a and 1b (Figure 3). Average per-cross PA was 0.56, 0.62 and 0.29 in scenarios 1a, 1b and 2, respectively (Figure 3a). Average per-trait PA was close to 1 for most traits in scenarios 1a and 1b (Figure 3b), and still high (around 0.75) in scenario 2, when excluding **nbclu** and **vigour** (Table S3). Overall PA (over the 150 cross x trait combinations) was 0.32. There was upward or downward bias for some traits, scenarios or methods, and in scenario 1a, LASSO resulted in larger bias (Figure S5).

Prediction of Mendelian sampling

We then measured PA for individual offspring within each half-diallel cross, thus considering separately the Mendelian sampling component. For each cross and trait, we compared the observed and predicted genotypic values in the three scenarios (Figure 2; Figure S6)

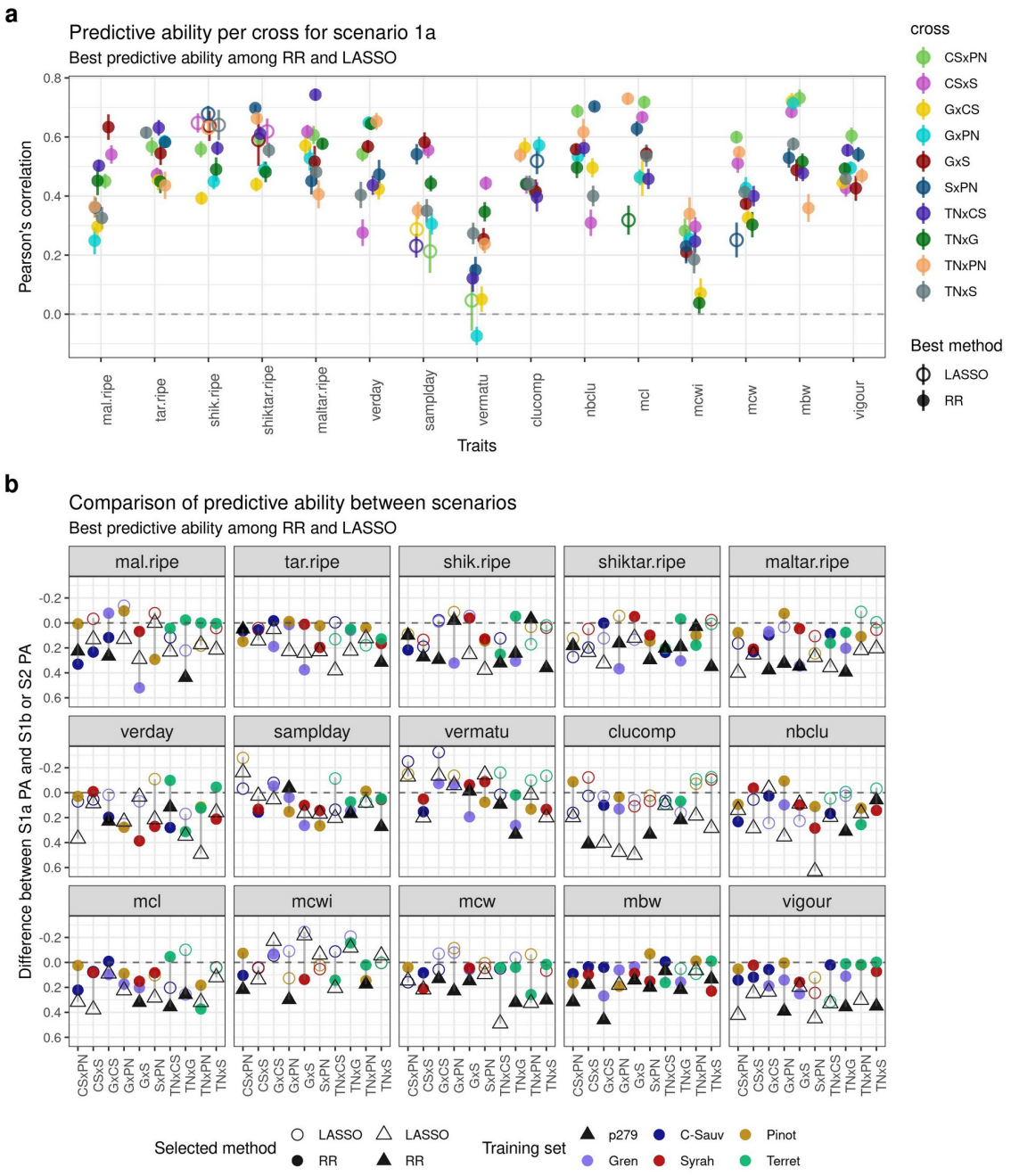


Figure 4: a: Mendelian sampling PA per trait and cross for scenario 1a with the best method between RR and LASSO. Vertical bars represent the standard error around the mean (95 % of the confidence interval), based on the outer cross-validation replicates. PA corresponds to the Pearson's correlation between the BLUPs of the genotypic value and the predicted genotypic values.
b: Difference between PA of scenario 1a and of the other scenarios. S2 is displayed with a triangle, and S1b by circles, colored according to the parental training set and filled if the best method was RR and empty otherwise.

In scenario 1a (Figure 4a), average PA per trait ranged from 0.18 for **vermatu** to 0.58 for **mbw**, with a 0.47 overall average (Figure S7a). The extent of PA variation among crosses depended on the trait and could be very large, as for **vermatu** (from -0.074 to 0.443). Unlike for traits, no cross had constantly high or low PA (Figure S7b). RR method yielded the highest PA values in most cases (91% of the 150 trait x cross combinations).

In scenario 1b (Figure 4b), there were two PA values per cross, one for each parental training set (TS). The difference between these two values varied widely, depending on the cross and trait (up to about 0.5 for **mal.ripe** in GxS), with an overall average of 0.39. Most often, PA was lower in scenario 1b than in scenario 1a, likely because no full-sibs were included in the training set. However, there were several cases with PA values similar or higher in scenario 1b for one parental TS compared to scenario 1a. RR method produced the best PA in 61% of the 300 combinations (2 parents x 15 traits x 10 crosses).

In scenario 2 (Figure 4b), overall average PA (0.26) was nearly halved compared to scenario 1a, with trait dependent differences in PA between both scenarios. Some traits such as **vigour**, **clucomp** and **maltar.ripe** displayed a particularly marked decrease. On the opposite, **mcwi** and **vermatu** reached equivalent PA values in both scenarios. RR provided the best PA in 61% of the 150 combinations.

Exploring factors affecting predictive ability, and training set optimization

We sought those variables affecting the PA values observed above, both for prediction of cross mean and Mendelian sampling. We then implemented training set (TS) optimization in an attempt to increase PA.

Variables affecting the prediction of cross mean

In scenario 2, per-cross PA was highly negatively correlated (-0.9) with the cross parents' pairwise distance on the first axis of the diversity panel PCA (Figure 5a, Figure S8a). Correlation with the additive relationship between parents was slightly lower (0.75) and non-significant at 5% (Figure S8a). No such strong correlation was found for per-cross PA in scenarios 1a or 1b (Figure S8a). The proportion of non-segregating markers showed low correlation with per-cross PA in all scenarios (Figure S8a).

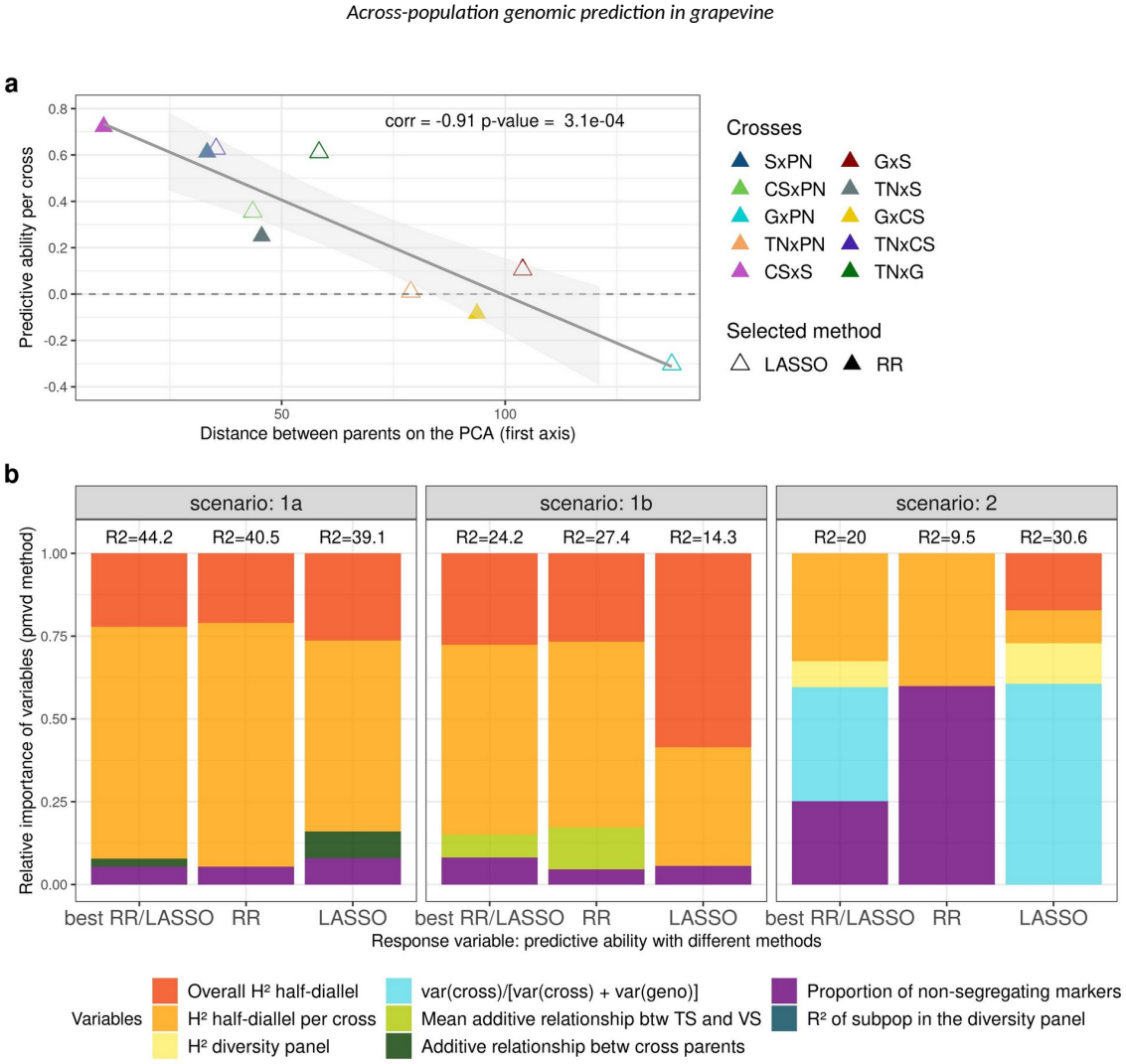


Figure 5: a: Plot of per-cross PA for cross mean in scenario 2, obtained with the best method between RR and LASSO for each cross, against the distance between cross parents on the first axis of the diversity panel PCA (Figure 1a). Best method is indicated with the triangle filling and cross with the color. b: Relative importance of variables affecting PA for Mendelian sampling in the three scenarios tested. Variables were selected from an overall model, after a model selection step. Response individual PA values were obtained either as the best one between RR and LASSO, with RR or with LASSO. Relative importance was estimated with pmvd method, from relaimpo R-package version 2.2-5.

Since variation in per-cross PA for scenario 2 was extremely large, from -0.3 for GxPN to 0.72 for CSxS (Figure 3a), we implemented TS optimization for each cross, to try and increase low PA values. Optimization actually improved PA for crosses with PA initially below 0.6, for TS sizes between 50 and 150 (Figure 6). The largest improvement, from -0.29 to 0.62, was observed for GxPN cross.

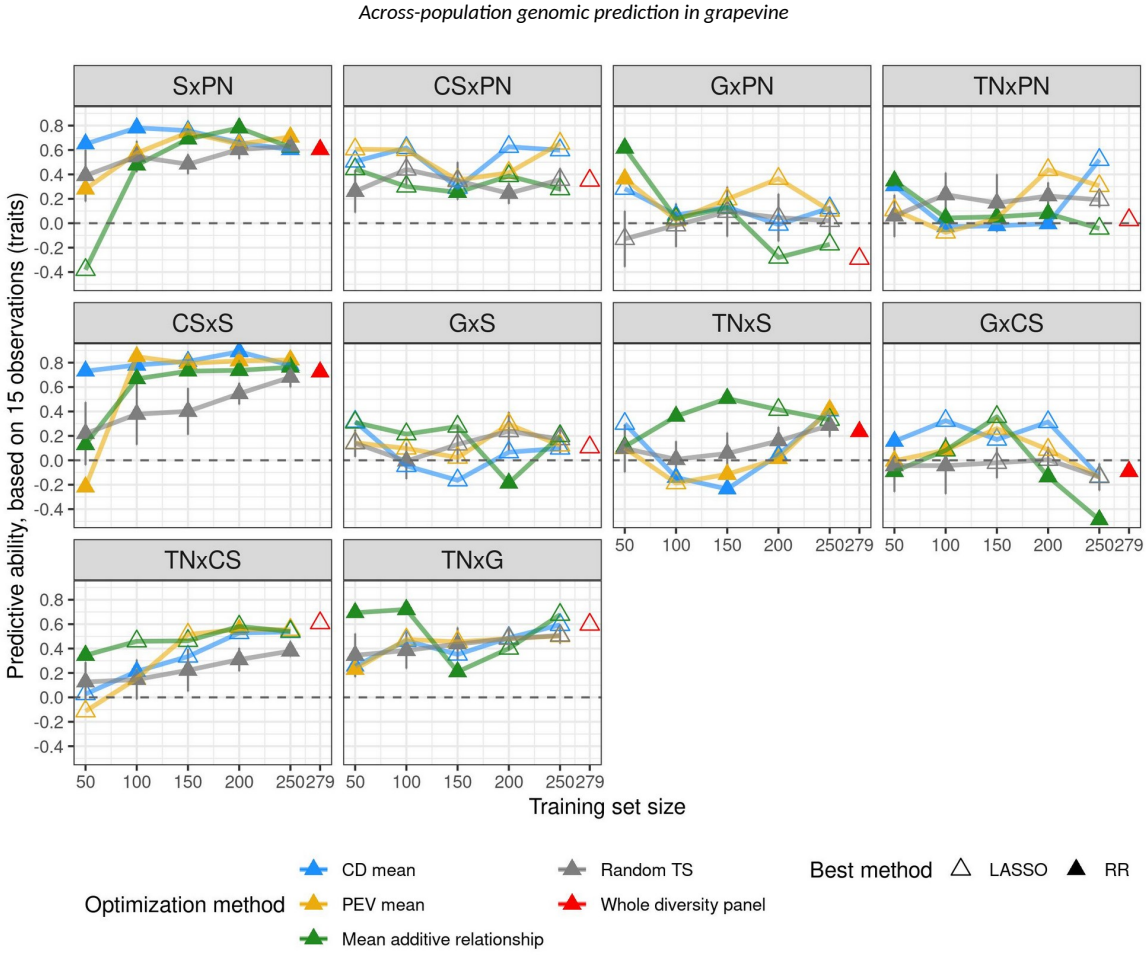


Figure 6: PA for cross mean predicion after training set optimization and with the best method between RR and LASSO, for each cross. Best method is indicated with the triangle filling and TS optimization method with the color. For comparison, random selection of TS genotypes (in grey) was performed and repeated ten times, error bars correspond to 95% of the confidence interval around the mean. We also report per-cross PA with the whole diversity panel (in red), with a maximum TS size of 279 which may vary depending on traits.

The variable that most affected per-trait PA was the $\sigma_C^2/\left(\sigma_C^2+\sigma_G^2\right)$ ratio (relative variance of cross effect). It was strongly correlated with PA in scenarios 1a and 1b (0.82 and 0.88, respectively), but not in scenario 2 (Figure S8b).

No other explanatory variable displayed any significant impact despite a fairly high correlation with per-trait or per-cross PA, which could be due to low sample sizes (15 and 10 for per-trait and per-cross PA, respectively).

Factors affecting Mendelian sampling prediction

To model Mendelian sampling PA for each scenario and method selected for each trait (RR, LASSO or best), we applied multiple linear regression on six to nine variables depending on the scenario, as detailed in Material and Methods. The highest coefficient of determination (44.2%) was obtained in scenario 1a with the best method (Figure 5b). Coefficients of determination were equivalent, lower and higher for LASSO compared to RR in scenarios 1a, 1b and 2, respectively. Three variables were found to impact PA in all scenarios: half-diallel overall

H^2, per-cross H^2 and the proportion of non-segregating markers. Surprisingly, half-diallel overall H^2 was not selected in scenario 2 with either RR or best method, while it had a strong effect in other modalities.

The selected variables were quite similar between scenarios 1a and 1b, with a high effect of half-diallel overall and per-cross H^2, but differed in scenario 2 in which more variables were selected. Overall, most of the relative importance came from variables related to the trait and not to the genetic composition of TS or validation set (VS).

We also calculated individual PA with optimized TSs derived from the diversity panel (Figure S9). However, we did not observe any improvement compared to using the whole diversity panel. This is consistent with the fact that genetic relationship seemed not to impact PA (Figure 5b).

Discussion

Our study allowed us to thoroughly explore GP potential in grapevine breeding, by scanning a large range of potentially useful configurations: (i) with 15 weakly related traits with variable levels of H^2 and phenotypic structure (subpopulation or cross effects on phenotypic data) (Figure S4), (ii) in across-population scenarios with TS ranging from half-sibs (scenario 1b) to a diversity panel (scenario 2), (iii) with 10 balanced VS crosses. Moreover, we decomposed PA into cross mean and Mendelian sampling components, each being useful in breeding to select parental genotypes and offspring within crosses, respectively. All these results allowed us to get insight into main factors affecting PA. We will focus our discussion on prediction with the diversity panel as TS, since this is the most sought-after configuration in perennial species breeding.

Range of PA values

For the prediction of cross mean, overall PA was 0.32 in scenario 2, equivalent to the average per-cross PA (0.29), while the average per-trait PA was twice as high (0.6) (Figure 3). In other studies concerning other plant crops, the average per-cross PA was not reported [5, 6, 7, 8], probably because, in most cases, there were not enough traits to estimate it. Bernardo et al. [5] and Osthushenrich et al. [6] also reported a high-average per-trait PA, above 0.9, while Yamamoto et al. [8] reported PA values from 0.21 to 0.57 depending on the trait.

For the prediction of Mendelian sampling, overall average PA was slightly lower than overall PA for cross mean in scenario 2 (0.26 and 0.32, respectively). Yet, Mendelian sampling PA was still quite high, considering that TS was essentially unrelated to VS, i.e., with no first-degree relationship with predicted progenies. The same diversity panel was previously used in Flutre et al. [26] for predicting individual genotypic values of 23 additional Syrah x Grenache offspring. The reported PA for **mbw** was 0.56, whereas in the present study, we obtained 0.35 in the Grenache x Syrah progeny (n=59). We further investigated such discrepancy, and found it related to a sampling bias due to the small VS size in Flutre et al. [26] (data not shown).

The range of average per-trait Mendelian sampling PA observed in scenario 2 (from 0.15 to 0.38) was consistent with those described on fruit perennial species where individual prediction was performed with a TS not directly related to the VS (neither half-sib nor full-sib). In *Coffea*, Ferrao et al. [28] reported differences in per-trait PA, from slightly negative values up to ca. 0.60. But, in this study, overall PA was calculated for all crosses of the VS, thus encompassing both cross mean and Mendelian sampling predictions, making comparison with our Mendelian sampling results alone impossible. In contrast, some studies in apple yielded within cross individual

PA values. For instance, Muranty et al. [29] reported average per-trait PA ranging from -0.14 to 0.37, and Roth et al. [30] found PA values from -0.29 to 0.72 for fruit texture, highly dependent on the cross for all traits. Conversely, our PA values were mainly stable over crosses and variable over traits, in the three scenarios (Figure S7). This difference might partly be due to the larger trait diversity we explored as compared to Roth et al. [30], as suggested by comparing our Figure S4 with their Figure 1A. A complementary explanation could be that progeny size varied from 15 to 80 in Roth et al. [30], while here progeny sizes were very close and thus less likely prone to sampling variability and to upward or downward bias.

Several factors may influence Mendelian sampling PA in our study compared to others. Among potential inflating factors, we can mention a slight over-representation of phenotyped individuals from the WW panel subpopulation, to which four out of the five parents of the half-diallel belong, leading to a higher genetic relationship between effective TS and VS. Factors potentially decreasing PA could be differences between TS and VS experimental designs since the diversity panel and the half-diallel were not phenotyped on the same years, had different plant management systems (overgrafting or simple grafting, respectively) and were planted a few kilometers apart. Nevertheless, for most studied traits, two years of phenotyping were used to compute genotypic BLUPs, which could at least compensate for differences between years, usually referred to as the millesime effect.

Variables affecting PA in across-population genomic prediction

We focused on PA obtained with the best method between RR and LASSO, to take into account the part of variability among traits associated with genetic architecture. Indeed, LASSO is supposed to be better adapted to traits underlined by few QTLs, while RR would yield better PA for highly polygenic traits. However, we showed that for a given trait x cross combination, i.e., for a given genetic architecture, the best method selected changed depending on the scenario: LASSO was more often selected for scenario 2 than for scenario 1a, both for cross mean and individual prediction. This means that the best method choice also depends on the relationship between TS and VS. This was also suggested in cattle breeding by MacLeod et al. [31], who found that BayesRC method (comparable to LASSO) yielded better results than GBLUP (comparable to RR) for across-population GP.

Regarding the other factors affecting PA, for cross mean prediction in scenario 2, no tested variable significantly affected per-trait PA. Conversely, per-cross PA was strongly affected by the genetic distance between parents (Figure 5a, Figure S8a). To our knowledge, such correlation has never been reported before, most probably because previous works investigated too few traits to afford per-cross PA calculation. We could hypothesize that when one parent is farthest from WW -the most represented panel subpopulation in TS- (e.g., Grenache, Figure 1a, Figure S1), marker effects for this parent might reflect different QTLs or allelic frequencies, compared to WW ones, thereby explaining the decrease in PA for crosses related to Grenache. Such differences underlying marker effects were already described in maize [32]. Simultaneously, some QTLs in this parent might be less genetically linked to causal polymorphisms due to more recombinations. However, this cannot be the only explanation for the large correlation of per-cross PA with pairwise parent distance, because the correlation between PA and genetic distance between TS and VS was much lower (Figure S8a).

For the prediction of Mendelian sampling, the variables explaining individual PA in scenario 2 were quite different from those explaining cross mean PA. Trait-related variables had a large impact on individual PA: half-

diallel overall and per-cross heritability, but also the relative variance of cross effect (Figure 5b). Surprisingly, genetic relationship between TS and VS had little to no impact on PA, although this factor has often been reported to affect PA [17,15]. Most studies reported separately the effects of different variables on individual PA. Riedelsheimer et al. [33] also performed multiple linear regression of individual PA on several factors to study their impact. They found that TS composition (number of crosses and their relationship with VS) explained most of the variance (41.7 %), followed by trait (27.6%) and VS composition (4.8%). The variance in genetic relationship between TS and VS may be smaller in our study.

Practical consequences on breeding programs

Across-population GP with model training in a diversity panel appeared to be promising in grapevine breeding for some traits and crosses, particularly for parent choice (Figure 3; Figure 4; Table S3; Figure S7).

The usefulness of GP for better selecting parents for future crosses can be at first assessed by the low overall correlation between mean parental genotypic values (BLUPs) and mean offspring BLUPs (0.28; see also Figure S10). This correlation was much lower than overall PA for cross mean in scenario 1b (0.66) and slightly lower than overall PA for cross mean in scenario 2 (0.32). In strawberry, Yamamoto et al. [8] also evidenced the interest of GP for predicting cross mean, with no additional benefit from including dominance effects into GP models, even if cross means were not equal to parental means. Moreover, in some cases, GP could provide other advantages over mean parental genetic values, for instance when parents are not phenotyped for some reasons, because too young or without representative phenotypes (e.g., using microvine [34], in a new environment, etc). This was actually the case, in our half-diallel trial, for the Terret Noir parent, which suffered from mortality probably due to rootstock incompatibility and consequently had no phenotypic record for most studied traits.

Even though PA was quite high for some traits and crosses in scenario 2, on average it remained moderate both for cross mean and individual prediction. Both PAs were much higher in scenario 1a, due to increased relationship between training and validation sets. Nevertheless, such an extreme configuration is rarely used in plant breeding programs, especially in perennial species, because it requires to partly phenotype the cross to be predicted. An intermediate configuration, scenario 1b, could be implemented in breeding programs when PA from scenario 2 is not sufficient and half-sib families are available, because in this scenario, cross mean PA was similar as in scenario 1a and individual PA intermediate between scenarios 1a and 2.

We found TS optimization useful mostly for cross mean prediction for crosses with low PA. The advantage of TS optimization was less clear for individual prediction. This was consistent with the fact that genetic parameters more strongly affected cross mean PA than individual PA. In contrast, Roth et al. [30] observed in apple a systematic increase of individual PA with an optimized TS in the same context (i.e., with a diversity panel as TS and bi-parental families as VS, and common optimization methods). To our knowledge, only a single study tested TS optimization for cross mean prediction, by Heslot and Feoktistov [35], who implemented optimization of parent selection for hybrid crossing in sunflower while selecting individuals to phenotype, but did not calculate cross mean PA.

Since our results show that prediction of cross mean can be quite accurate and useful in scenario 2, we decided to go one step further and implemented cross mean prediction for all 38,781 possible crosses between the 279 genotypes of the diversity panel, based on parental average genotypes (Table S2) and on marker effects estimated with RR in this population. As predicted cross mean were biased for some traits in the ten half-diallel

crosses (Figure S5), we estimated the bias for each trait from these data to correct the predicted mean in the possible diversity panel crosses. Figure S11 shows the large potential diversity to be explored through crossing in grape, for all the traits considered in the present study, illustrating the finding of Myles et al. [36] that genetic diversity in grapevine was largely unexploited. Such an example opens many prospects for the use of GP to design future crosses. Indeed, we limited here our prediction to the 279 panel genotypes representing the *Vitis vinifera* diversity, but potentially any other (unphenotyped) genotype of interest with dense genotypic data could be used for this purpose as exemplified with the half-diallel, since its five parents were not part of the diversity panel.

Prospects

Based on our results, the following improvements could be tested: i) increase SNP density [25, 37] and include structural variants ii) implement non-additive effects in GP models such as dominance or epistatic effects and iii) add crosses from other panel subpopulations as VSs. Indeed, since all our half-diallel crosses had at least one parent belonging to the WW subpopulation, it would be beneficial to include crosses with parents from the WE and TE subpopulations too. Specific GP models that include genetic structure in marker effect estimation [38, 39] could also be tested.

Predicting cross variance could also prove useful to design the offspring selection step, more specifically for choosing the number of offspring to test or produce for a given cross. Depending on the available funds and breeding program, a breeder may want to select crosses with high genetic variance, in order to maximize the probability to generate top-ranking genotypes. Conversely, choosing a cross with low variance could limit the risk of breeding poor genotypes.

Conclusion

We implemented GP in grapevine in a breeding context, i.e., across populations, on 15 traits, in ten related crosses, and obtained moderate to high PA values for some crosses and traits, thus showing GP usefulness in grapevine. Never before had genomic prediction been implemented for so many traits and crosses simultaneously in this species. We showed that per-cross PA was strongly correlated with the genetic distance between parents, whereas Mendelian sampling PA was largely determined by trait-related variables, such as heritability and the magnitude of the cross effect.

Material and Methods

Plant material

The half-diallel consists of 10 pseudo-F_1 bi-parental families obtained by crossing five *Vitis vinifera* cultivars: Cabernet-Sauvignon (CS), Pinot Noir (PN), Terret Noir (TN), Grenache (G) and Syrah (S) [40]. Each family comprised between 64 and 70 offspring, with a total of 676 individuals including parents.

The diversity panel consists of 279 cultivars selected as maximizing genetic diversity and minimizing kinship among cultivated grapevine. Grapevine genetic diversity is highly heterozygous and weakly structured into three subpopulations: WW (Wine West), WE (Wine East) and TE (Table East) [27].

Across-population genomic prediction in grapevine

Field experiments

Field design

The half-diallel was created in 1998 at INRAE Montpellier, grafted on Richter 110, and planted in 2005, at the Institut Agro experimental vineyard "Le Chapitre" in Villeneuve-lès-Maguelone (Southern France). The progenies were planted in two randomized complete blocks, with plots of two consecutive plants per offspring per block.

The field design for the diversity panel was previously described in Flutre et al. [26]. Briefly, cultivars were overgrafted on 6-year-old Marselan in 2009, itself originally grafted on Fercal rootstock, a few kilometers away from the diversity panel. They were planted in five randomized complete blocks, with one plant per cultivar per block.

Phenotyping

We studied 15 traits in both trials: berry composition with malic (**mal.ripe**), tartaric (**tar.ripe**) and shikimic acid (**shik.ripe**) concentrations in $\mu_{eq}.L^{-1}$ measured at ripe stage (20° Brix) (according to Rienth et al. [41]), from which two ratios were derived, shikimic / tartaric acid (**shiktar.ripe**) and malic / tartaric acid (**maltar.ripe**); morphological traits with mean berry weight (**mbw**, in g) measured on 100 random berries, mean cluster weight (**mcw**, in g), mean cluster length (**mcl**, in cm) and mean cluster width (**mcwi**, in cm), measured on 3 clusters, number of clusters (**nbclu**) and cluster compactness (**clucomp**) measured on the OIV semi-quantitative scale; phenology traits with veraison date (onset of ripening; **verday,** in days since January 1st), maturity date corresponding to berries reaching 20° Brix (**samplday**, in days since January 1st) and the interval between veraison and maturity (**vermatu**, in days); vigour (**vigour**, in kg), derived as the ratio between pruning weight and the number of canes. Phenotypic data were collected between 2013 and 2017 for the half-diallel and in 2011-2012 for the diversity panel. There was a slight over-representation of phenotypes from the WW subpopulation because of fertility issues in WE and TE subpopulations

SNP genotyping

For the half-diallel, we used genotyping-by-sequencing (GBS) SNP markers derived by Tello et al. [40], 622 of the 676 individuals being successfully genotyped, as well as the five parents. Raw GBS data were processed separately for each cross, and then markers from all crosses were merged together (390,722 SNPs), thus generating many missing data (85% of missing data per marker on average), since all markers did not segregate in all progenies. Markers with more than 80% of missing data were removed and remaining markers were imputed with FImpute3 [42] (86,017 SNPs). Some parental cultivars were used either as female or male, depending on the cross, a configuration not allowed by FImpute3. We thus declared only a partial pedigree maximizing the number of crosses defined with both parents (Table S4). For the diversity panel, we used the same SNP markers as in Flutre et al. [26], except that we applied a filter on minor allelic frequency (5%) and no filter on linkage disequilibrium, which yielded 83,264 SNPs.

Finally, we only retained the 32,894 SNPs common to both populations.

Phenotypic data analyses

Half-diallel

- **Statistical modeling for estimating genotypic values**

For each trait, we excluded outlier values by visual inspection of raw phenotypic data and computed a log or square-root transformation if its distribution looked skewed. Then, we fitted the following linear mixed full model by Maximum Likelihood:

$$y_{ijkl} = \mu + \underline{G_i} + \underline{C_j} + B_k + Y_l + \underline{(B:Y)_{kl}} + \underline{(G:Y)_{il}} + \underline{(C:Y)_{jl}} + \underline{x} + \underline{y} + \underline{x:y} + \underline{(x:Y)_l} + \underline{(y:Y)_l} + \epsilon_{ijkl}$$

with y_{ijkl} the phenotype of genotype i from cross j in block k and year l. Among the fixed terms, μ was the overall mean, and B_k and Y_l the effects of block k and year l. Among the random terms, G_i and C_j were the effects of genotype i nested within cross j, and x and y the field coordinates. Interactions are indicated with ":". ϵ_{ijkl} was the random residual term, assumed to be normally distributed.

Sub-model selection was based on Fisher tests for fixed effects and log-likelihood ratio tests for random effects. It was performed with the step function from lmerTest R-package [43]. Variance components were estimated after re-fitting the selected model by Restricted Maximum Likelihood, and diagnostic plots were drawn to visually check the acceptability of model hypotheses such as homoscedasticity or normality. Best Linear Unbiased Predictors (BLUPs) of cross (C) and genotype (G) values were computed. For genomic predictions, we used their sum (C+G) as total genotypic values for both training and validation data. Variance component estimates were used to compute the proportion of genetic variance due to differences between crosses as: $\sigma^2_C / (\sigma^2_C + \sigma^2_G)$.

- **Heritability estimation**

We estimated overall (for the whole half-diallel) broad-sense heritability for genotype-entry means [44] as:

$$H^2 = \frac{\sigma^2_C + \sigma^2_G}{\sigma^2_C + \sigma^2_G + \frac{\sigma^2_{C:Y} + \sigma^2_{G:Y} + \sigma^2_{x:Y} + \sigma^2_{y:Y}}{n_{year}} + \frac{\sigma^2_x + \sigma^2_y + \sigma^2_{x:y} + \sigma^2_\epsilon}{n_{year} \times n_{rep.\,year}}}$$

with genotype (G) and cross (C) variances at the numerator. Random variance components involving year (Y) were divided by the mean number of years (n_{year}). Other random variance components involving spatial effects or residuals were divided by the mean number of years times the mean number of replicates per year ($n_{rep.\,year}$).

We also estimated broad-sense heritability per cross (thereafter used to name half-diallel full-sib family). For that, we applied the same selected model, but removed all effects involving cross. Then, we estimated variance components within each cross, and heritability with the same formula, after removing variances involving cross.

All information on analyses of phenotypic data and heritability of the half-diallel is detailed in Table S1.

Diversity panel

We used the genotypic values previously estimated in Flutre et al. [26] with a similar statistical procedure to the one described above for the half-diallel. All phenotypic analysis information is provided in Table S3 of Flutre et al. [26].

For each of the two populations, genotypic BLUPs were scaled, allowing comparison among traits.

Genomic prediction statistical methods

Marker effects were estimated using two methods to take into account varying genetic architecture among the traits studied. Ridge regression (RR) [45], best adapted to many minor QTLs, shrinks marker effects towards 0. Least Absolute Shrinkage and Selection Operator (LASSO) [46], best adapted to a few major QTLs, applies a L1 norm on allelic effects, thus forcing some to be exactly 0. Both methods were implemented with R/glmnet package [47] and the amount of shrinkage, controlled by λ parameter, was calibrated by five-fold inner cross-validation within each training set, using cv.glmnet function.

Genomic prediction scenarios

We assessed prediction within half-diallel crosses under three different training scenarios (Figure 2; Table S5):

- **Scenario 1a:** whole half-diallel prediction. We applied random outer 10-fold cross-validation over the whole half-diallel population. In each fold, 90% of the phenotyped offspring were used as the training set (TS) and the remaining 10% as the validation set (VS). Cross-validation was replicated ten times.
- **Scenario 1b**: half-sib prediction. For each half-diallel cross used as VS, we trained the model with the three half-sib crosses of each parent in turn, thus predicting each cross twice.
- **Scenario 2**: across-population prediction. We used the whole diversity panel as TS and each half-diallel cross as VS.

Predictive ability assessment

In order to account for the effect of genetic architecture, we applied both RR and LASSO methods for each trait and cross and kept the best PA, for both cross mean and within cross individual prediction.

Prediction of cross mean

Cross mean PA was assessed as Pearson's correlation between the average value of observed total genotypic values (sum of genotype and cross BLUPs for each offspring) for each cross, and the mean predicted genotypic value per cross, calculated in two ways, as:

- average predicted value over all offspring of the cross. In scenario 1a, each offspring was predicted 10 times, thus we also averaged the predicted value over the 10 replicates.
- predicted value for the parental average genotype, defined at each locus and for each cross as the mean allelic dosage according to the expected segregation pattern based on parents' genotypes (Table S2).
- genotypic values predicted with these two modalities were highly correlated (above 0.98) in the three scenarios and for the two methods (partly shown in Figure S12). Therefore, in subsequent analyses, we used only prediction with parental average genotypes.

Pearson’s correlation between observed and predicted values was calculated on all cross x trait combinations (overall PA), for each trait (per-trait PA) and for each cross (per-cross PA).

Within-cross individual prediction

We measured PA within each cross in each scenario as Pearson's correlation between observed total genotypic values and predicted genotypic values.

Test of variables affecting predictive ability

We tested the effect of several variables on within-cross individual PA, in each scenario. We built a multiple linear regression model with PA per trait x cross combination as the response variable and as predictors, a set of variables common to all three scenarios plus specific variables for scenarios 1b and 2. Common variables were: the proportion of non-segregating markers in the cross, overall and per-cross broad-sense heritability, the distance between the parents of the cross measured either as the additive relationship or as the distance on the first or first two axes of the panel PCA (Figure 1a) and the proportion of genetic variance due to differences between crosses ($\sigma^2_C / \left(\sigma^2_C + \sigma^2_G \right)$ ratio). A specific variable for scenarios 1b and 2 was the mean additive relationship between training and validation sets. In scenario 2, it was calculated for each trait only with phenotyped individuals. Specific variables for scenario 2 were: broad-sense heritability in the diversity panel (retrieved from Flutre et al. [26] and Table S1) and the percentage of trait variance explained by the subpopulation factor (see below). After fitting the overall model, we applied a forward-backward stepwise regression, with the AIC criterion to select the best explanatory model. Then, we estimated the relative importance of each variable selected in this model with the pmvd method [48], which allows to decompose the R^2 of correlated regressors with the R-package relaimpo [49].

The percentage of trait variance within the diversity panel explained by subpopulation (WW, WE or TE) was evaluated by fitting for each trait the following linear model: $G = P + \epsilon$, where G is the genotypic (BLUP) value within the diversity panel, P is a fixed subpopulation effect, and ϵ a random residual term. The percentage of variance due to differences between subpopulations was then estimated as the coefficient of determination (R^2) of the model.

Training set optimization

We tested three methods for optimizing TS in scenario 2, for both cross mean and within-cross individual prediction. We used the STPGA R-package[50] to implement Prediction Error Variance (*PEVmean*) and *CDmean* (based on the coefficient of determination)[10]. Moreover, we computed the mean relationship criterion (*MeanRel*), as the mean additive relationship between each genotype in TS and all genotypes in VS. Each optimized TS was specific to a cross. The realized additive relationship based on marker data was estimated using the rrBLUP R-package[51] with the A.mat function implementing the formula from VanRaden et al. [52]. For each of these three optimization methods, we tested five TS sizes (50, 100, 150, 200, 250). PA values obtained with each optimized TS were compared with those obtained with a random sample of genotypes of the same size, repeated 10 times.

Data availability

All analyses were conducted using free and open-source software, mostly R. Phenotypic and genotypic data, R scripts and result tables are available at https://data.inrae.fr/privateurl.xhtml?token=1925c973-a11b-45ad-b297-69db8ec2c270 .

Acknowledgments

We thank Charles Romieu for his help on the acquisition of phenotypic data as well as Philippe Châtelet and Morgane Roth for their suggestions on the text. This work has been realized with the support of the SouthGreen platform and MESO@LR-Platform at the University of Montpellier.

Author contribution statement

VS, AD, PT, TF, JPP and LLC conceived the idea of the study and contributed to funding acquisition; TP, PF, AD and JPP obtained the phenotypic data used in this work; VS, AD, PT, LLC, TF and CB performed and interpreted results; AD conceived the half-diallel population; PT is the PhD supervisor of CB; CB wrote the original draft, which was reviewed and edited by all authors. All authors read and approved the final manuscript.

Funding

Partial funding of the PhD was provided by ANRT (grant number 2018/0577), IFV and Inter-Rhône. The authors declare that they have no conflict of interest. The authors declare that the experiments comply with the current laws of the country in which they were carried out.

Bibliography

1. Meuwissen, T. H. E., Hayes, B. J. & Goddard, M. E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 11 (2001).

2. Heffner, E. L., Sorrells, M. E. & Jannink, J.-L. Genomic Selection for Crop Improvement. *Crop Science* **49**, 1 (2009).

3. Werner, C. R. *et al.* How Population Structure Impacts Genomic Selection Accuracy in Cross-Validation: Implications for Practical Breeding. *Frontiers in Plant Science* **11**, 592977 (2020).

4. Falconer, D. S. & Mackay, T. F. C. *Introduction to quantitative genetics*. (Pearson, Prentice Hall, 2009).

5. Bernardo, R. Genomewide Selection of Parental Inbreds: Classes of Loci and Virtual Biparental Populations. *Crop Science* **54**, 2586–2595 (2014).

6. Osthushenrich, T., Frisch, M. & Herzog, E. Genomic selection of crossing partners on basis of the expected mean and variance of their derived lines. *PLOS ONE* **12**, e0188839 (2017).

7. Imai, A. *et al.* Predicting segregation of multiple fruit-quality traits by using accumulated phenotypic records in citrus breeding. *PLOS ONE* **13**, e0202341 (2018).

8. Yamamoto, E., Kataoka, S., Shirasawa, K., Noguchi, Y. & Isobe, S. Genomic Selection for F1 Hybrid Breeding in Strawberry (Fragaria × ananassa). *Frontiers in Plant Science* **12**, (2021).

9. Heslot, N., Yang, H.-P., Sorrells, M. E. & Jannink, J.-L. Genomic Selection in Plant Breeding: A Comparison of Models. *Crop Science* **52**, 146–160 (2012).

10. Rincent, R. *et al.* Maximizing the Reliability of Genomic Selection by Optimizing the Calibration Set of Reference Individuals: Comparison of Methods in Two Diverse Groups of Maize Inbreds (Zea mays L.). *Genetics* **192**, 715–728 (2012).

11. Charmet, G. *et al.* Genome-wide prediction of three important traits in bread wheat. *Molecular Breeding* **34**, 1843–1852 (2014).

12. Riedelsheimer, C., Technow, F. & Melchinger, A. E. Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. *BMC Genomics* **13**, 452 (2012).

13. Ornella, L. *et al.* Genomic Prediction of Genetic Values for Resistance to Wheat Rusts. *The Plant Genome* **5**, (2012).

14. Heffner, E. L., Jannink, J.-L. & Sorrells, M. E. Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. *The Plant Genome* **4**, 65–75 (2011).

15. Lund, M. S., Su, G., Janss, L., Guldbrandtsen, B. & Brøndum, R. F. Genomic evaluation of cattle in a multi-breed context. *Livestock Science* **166**, 101–110 (2014).

16. Hayes, B. J., Bowman, P. J., Chamberlain, A. C., Verbyla, K. & Goddard, M. E. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics Selection Evolution* **41**, 51 (2009).

17. de Roos, A. P. W., Hayes, B. J. & Goddard, M. E. Reliability of Genomic Predictions Across Multiple Populations. *Genetics* **183**, 1545–1553 (2009).

18. Rio, S. *et al.* Disentangling group specific QTL allele effects from genetic background epistasis using admixed individuals in GWAS: An application to maize flowering. *PLOS Genetics* **16**, e1008241 (2020).

19. Wientjes, Y. C. *et al.* Empirical and deterministic accuracies of across-population genomic prediction. *Genetics Selection Evolution* **47**, 5 (2015).

20. Schopp, P., Müller, D., Wientjes, Y. C. J. & Melchinger, A. E. Genomic Prediction Within and Across Biparental Families: Means and Variances of Prediction Accuracy and Usefulness of Deterministic Equations. *G3: Genes, Genomes, Genetics* **7**, 3571–3586 (2017).

21. Raymond, B., Wientjes, Y. C. J., Bouwman, A. C., Schrooten, C. & Veerkamp, R. F. A deterministic equation to predict the accuracy of multi-population genomic prediction with multiple genomic relationship matrices. *Genetics Selection Evolution* **52**, 21 (2020).

22. Viana, A. P., Resende, M. D. V. de, Riaz, S. & Walker, M. A. Genome selection in fruit breeding: application to table grapes. *Scientia Agricola* **73**, 142–149 (2016).

23. Migicovsky, Z. *et al.* Patterns of genomic and phenomic diversity in wine and table grapes. *Horticulture Research* **4**, 17035 (2017).

24. Brault, C. *et al.* Harnessing multivariate, penalized regression methods for genomic prediction and QTL detection to cope with climate change affecting grapevine. *bioRxiv* 2020.10.26.355420 (2020) doi:10.1101/2020.10.26.355420.

25. Fodor, A. *et al.* Genome-Wide Prediction Methods in Highly Diverse and Heterozygous Species: Proof-of-Concept through Simulation in Grapevine. *PLoS ONE* **9**, e110436 (2014).

26. Flutre, T. *et al.* Genome-wide association and prediction studies using a grapevine diversity panel give insights into the genetic architecture of several traits of interest. *bioRxiv* 2020.09.10.290890 (2020) doi:10.1101/2020.09.10.290890.

27. Nicolas, S. D. *et al.* Genetic diversity, linkage disequilibrium and power of a large grapevine (Vitis vinifera L) diversity panel newly designed for association studies. *BMC Plant Biol* **16**, 74 (2016).

28. Ferrão, L. F. V. *et al.* Accurate genomic prediction of Coffea canephora in multiple environments using whole-genome statistical models. *Heredity* **122**, 261–275 (2019).

29. Muranty, H. *et al.* Accuracy and responses of genomic selection on key traits in apple breeding. *Horticulture Research* **2**, (2015).

30. Roth, M. *et al.* Genomic prediction of fruit texture and training population optimization towards the application of genomic selection in apple. *Horticulture Research* **7**, 148 (2020).

31. MacLeod, I. M. *et al.* Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* **17**, 144 (2016).

32. Rio, S., Mary-Huard, T., Moreau, L. & Charcosset, A. Genomic selection efficiency and a priori estimation of accuracy in a structured dent maize panel. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **132**, 81–96 (2019).

33. Riedelsheimer, C. *et al.* Genomic Predictability of Interconnected Biparental Maize Populations. *Genetics* **194**, 493–503 (2013).

34. Chaïb, J. *et al.* The grape microvine – a model system for rapid forward and reverse genetics of grapevines. *The Plant Journal* **62**, 1083–1092 (2010).

35. Heslot, N. & Feoktistov, V. Optimization of Selective Phenotyping and Population Design for Genomic Prediction. *Journal of Agricultural, Biological and Environmental Statistics* **25**, 579–600 (2020).

36. Myles, S. *et al.* Genetic structure and domestication history of the grape. *Proceedings of the National Academy of Sciences* **108**, 3530–3535 (2011).

37. Dai, Z., Long, N. & Huang, W. Influence of Genetic Interactions on Polygenic Prediction. *G3 GenesGenomesGenetics* **10**, 109–115 (2020).

38. Heslot, N. & Jannink, J.-L. An alternative covariance estimator to investigate genetic heterogeneity in populations. *Genetics Selection Evolution* **47**, 93 (2015).

39. Ramstein, G. P. & Casler, M. D. Extensions of BLUP Models for Genomic Prediction in Heterogeneous Populations: Application in a Diverse Switchgrass Sample. *G3 GenesGenomesGenetics* **9**, 789–805 (2019).

40. Tello, J. *et al.* A novel high-density grapevine (Vitis vinifera L.) integrated linkage map using GBS in a half-diallel population. *Theoretical and Applied Genetics* **132**, 2237–2252 (2019).

41. Rienth, M. *et al.* Temperature desynchronizes sugar and organic acid metabolism in ripening grapevine fruits and remodels their transcriptome. *BMC Plant Biology* **16**, 164 (2016).

42. Sargolzaei, M., Chesnais, J. P. & Schenkel, F. S. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* **15**, 478 (2014).

43. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models using lme4. *arXiv:1406.5823 [stat]* (2014).

44. Piepho, H.-P. & Möhring, J. Computing Heritability and Selection Response From Unbalanced Plant Breeding Trials. *Genetics* **177**, 1881–1888 (2007).

45. Hoerl, A. E. & Kennard, R. W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **12**, 55–67 (1970).

46. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58**, 267–288 (1996).

47. Hastie, T. & Qian, J. Glmnet vignette. (2016).

48. Feldman, B. E. *Relative Importance and Value*. https://papers.ssrn.com/abstract=2255827 (2005) doi:10.2139/ssrn.2255827.

49. Groemping, U. Relative Importance for Linear Regression in R: The Package relaimpo. *Journal of Statistical Software* **17**, 1–27 (2006).

50. Akdemir, D. STPGA. (2018).

51. Endelman, J. B. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome* **4**, (2011).

52. VanRaden, P. M. Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* **91**, 4414–4423 (2008).

3.3 Variance prediction

Article II describes the prediction of cross means in three different prediction scenarios. Here I describe the prediction of cross variances, which could be useful for choosing new crosses in breeding. But this exploratory study was done only in one scenario, for which estimates of the required parameters were available.

3.3.1 Material and methods

I predicted cross variances in scenario 1a (within half-diallel prediction), using the **predCrossVar** R-package (Wolfe et al., 2021), based on parental haplotypes, recombination frequency and additive marker effects in the whole half-diallel. Parental haplotypes were recovered from FImpute3 outputs (see Article II). Recombination frequency was estimated from the half-diallel consensus genetic map (Tello et al., 2019). The genetic position of markers that were not located in the genetic map was predicted from their genomic location with local polynomial regression (**loess** R-package). Marker effects were averaged over 10 cross-validation folds repeated 10 times in the half-diallel.

3.3.2 Results and discussion

Observed variances were calculated from the genotypic BLUPs. Predicted variances were computed for each cross and trait within the half-diallel. Predicted cross variances were fairly well correlated with observed values (overall correlation of 0.58, Figure 3.1), even though systematically lower, i.e., with a downward bias. Some traits, such as **mcwi** and **mcw** were predicted with a lower variance compared to other traits, although genotypic BLUPs were scaled for all traits.

Predicting cross variance could also prove useful to design the offspring selection step, more specifically for choosing the number of offspring to test or produce within a given cross. Depending on the budget and breeding program, a breeder may want to select crosses with high genetic variance, in order to maximize chances to get top-ranking genotypes. Conversely, choosing a cross with low variance could limit the risk of getting poor genotypes. As a preliminary test, I implemented cross variance prediction in scenario 1a, and the results showed that cross variance prediction was quite accurate, almost as much as cross mean prediction (respectively 0.58 and 0.61 with RR). Conversely, Wolfe et al. (2021) found that per trait cross variance was less accurately predicted than cross mean. Implementing cross variance prediction in scenario 2 was unfortunately not possible, because I did not have haplotypes nor recombination fraction within the TS. If haplotypes could be retrieved using a dedicated phasing software, further studies are needed to determine recombination fraction in a diversity panel without related genotypes.

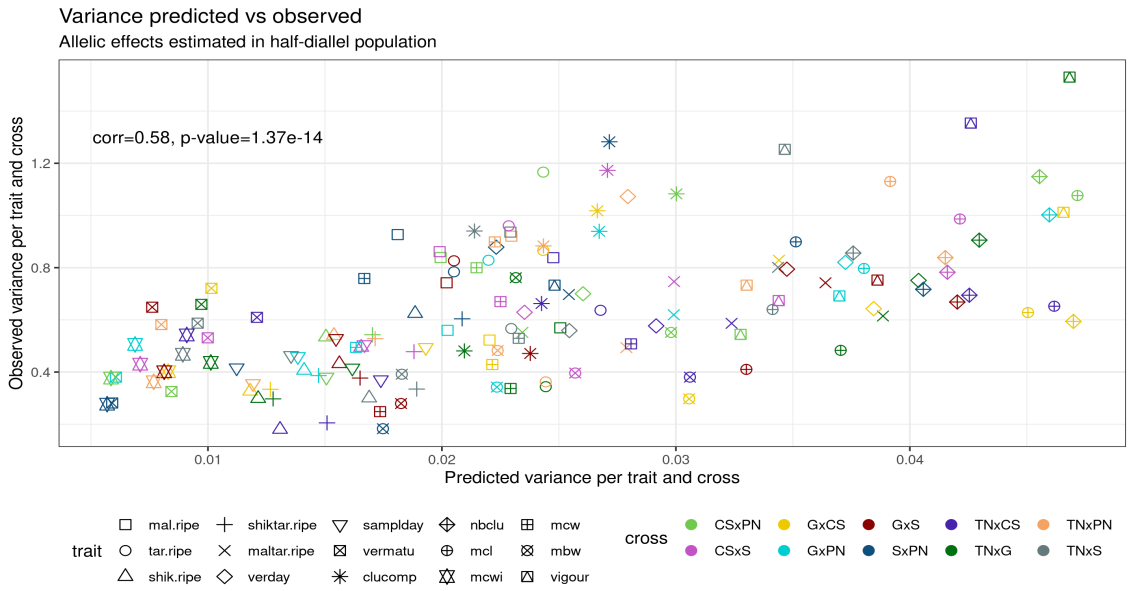


FIGURE 3.1 – Observed vs predicted cross variances for 15 traits in the half-diallel. Observed variances were calculated from the genotypic BLUPs. Predicted variances were computed using **predCrossVar** R-package, for each cross and trait.

3.4 Chapter general discussion

The selection of parents for future crosses should aim at getting offspring with high genotypic values by maximizing cross means and variances. Cross mean and variance predictions, associated with a selection index can be combined into a multi-trait usefulness criterion, as defined by Wolfe et al. (2021). However, crosses with superior mean are likely to also have lower variance (Neyhart & Smith, 2019; Zhong & Jannink, 2007). When using the usefulness criterion, the associated GP model might be adapted to account for directional dominance, i.e., inbreeding depression.

Chapter 4

Interest of phenomic prediction as an alternative to genomic prediction in grapevine

Charlotte Brault, Juliette Lazerges, Agnès Doligez, Miguel Thomas, Martin Ecarnot, Pierre Roumet, Yves Bertrand, Gilles Berger, Thierry Pons, Pierre François, Loïc Le Cunff, Patrice This, Vincent Segura.

UMR AGAP Institut, Université de Montpellier, Cirad, INRAE, Institut Agro-Montpellier SupAgro, F-34000 Montpellier, France

UMT Geno-Vigne®, IFV-INRAE-Institut Agro, Montpellier, France

Abstract

Phenomic prediction has been defined as an alternative to genomic prediction by using spectra instead of molecular markers. Reflectance at a given wavelength reflects the biochemical composition within a tissue, this composition being itself genetically determined. Thus, a relationship matrix built from spectra could potentially capture genetic relationship. This new methodology has been successfully applied in several cereal species but little is known so far about its interest in perennial species. Besides, phenomic prediction has only been tested for a restricted set of traits, mainly related to yield or phenology. This study aims at applying phenomic prediction for the first time in grapevine for 15 traits, using spectra collected on two tissues and over two consecutive years, in two populations, a half-diallel and a diversity panel. First, we characterized the genetic signal in spectra and under which condition phenomic prediction could be maximized, then phenomic predictive ability was compared to the genomic one. Finally, spectra and SNPs were combined into a single model, to test if it could enhance genomic predictive ability.

We found that the co-inertia between spectra and genomic data was stable across tissues or years, but variable across populations, with co-inertia around 0.3 for the diversity panel and 0.6 for the half-diallel population. Differences between populations were also observed for predictive ability of phenomic prediction, with an average of 0.27 for the diversity panel and 0.35 for the half-diallel. For both populations, there was a correlation across traits between predictive ability of genomic and phenomic prediction, with a slope around 1 and an intercept of –0.2, thus suggesting that phenomic prediction could be applied for any trait.

Overall, our results suggest the usefulness of NIRS as a low-cost alternative to genotyping for predicting complex traits in grapevine.

4.1 Introduction

Viticulture has to face two major threats in the next decades, diseases and climate change, which impact both yield and wine quality. Plant breeding could help mitigating these impacts by mobilizing grapevine genetic diversity (Morales-Castilla et al., 2020). However, grapevine breeding is currently slow because of the long juvenile period and cumbersomeness of field trials. Genomic prediction (GP), first proposed by Meuwissen et al. (2001) is a promising tool to speed up breeding programs and increase selection accuracy, by using genomic information to predict breeding values of candidates to selection. Even though genotyping costs have decreased drastically during the last decades, they can still be prohibitive when hundreds of selection candidates have to be genotyped. That is why Rincent et al. (2018) proposed to switch from genomic markers to near-infrared spectra (NIRS) measured on plant tissues, in a new concept called phenomic prediction (PP). The relationship matrix based on NIRS is indeed expected to share similarities with the genomic relationship matrix, because reflectance at a wavelength is determined by the molecular composition of the analyzed sample (Beer-Lambert law), which in turn is determined by genetic and environmental factors. As PP uses endophenotypes such as NIRS, it may better account for genotype-by-environment interactions (GxE) than GP, as well as for non-additive genetic effects. In addition, besides being cheaper, NIR measurements are high-throughput, which is required for screening the large populations typically evaluated in breeding programs. One step further, Robert et al. 2021 (in press) proposed a definition of genomic-like omics based (GLOB) prediction, which encompasses both phenomic and other omics-based prediction as in Schrag et al. (2018). GLOB is a particular configuration where NIRS (or other omics) used for model training and prediction come from different environments.

Rincent et al. (2018) found that PP could reach higher predictive ability (PA) than GP with wheat grain NIRS and equivalent PA with poplar wood NIRS for some traits. In wheat, when predicting across environments, PP was still more accurate than GP for most traits.

Other studies, such as Lane et al. (2020) in maize reported PA for PP, but in this study, GP was

not implemented for comparison. Krause et al. (2019) applied PP in wheat in a single environment with hyperspectral imaging from different phenological stages, they found higher PA with PP than with GP for most time-points studied. Indeed, this might be explained thanks to GxE interactions, because NIRS on training set (TS) and validation set (VS) were measured in a single environment. Several studies also reported an increase in PA when combining genomic and phenomic matrices in a single prediction model (Cuevas et al., 2019; Galán et al., 2020). Nevertheless, PP is still in its infancy, as it has been mostly applied to cereals with grain and leaves as tissues. Many issues remain, in particular which could be the best way to implement PP in breeding programs. In the case of perennial species, such as grapevine, year effect is known to strongly affect phenotype, and how behaves PP in this context remains to be studied. Also, in the case of woody perennial, wood matter offers another kind of material for collecting spectra which could be complementary to leaves. Rincent et al. (2018) found in wheat that combining NIRS collected on leaves and grains could enhance the PA, but the gain was not systematic. More work is thus required to devise a strategy for implementing PP in breeding programs.

In grapevine, GS has been already implemented and gave promising results on different populations (Brault et al., 2021a; Brault et al., 2021b; Flutre et al., 2020; Fodor et al., 2014). However, so far to our knowledge only one study has evaluated PP in woody perennials (in poplar, Rincent et al., 2018) and consequently none on grapevine.

The aim of this study was to understand under which configuration PP could be implemented in grapevine breeding programs. For that, we first provided a thorough characterization of the genetic signal in spectra. Specifically, we performed a co-inertia analysis (Dolédec & Chessel, 1994) to assess the pairwise relationship between genotyping and NIRS matrices. This methodology was already used in ecology and multi-omics studies but has never been applied in this context (Meng et al., 2014; Min et al., 2019).

Then, we compared multiple modalities for performing PP, such as using raw NIRS vs derived BLUPs over a single or two years and over a single or two tissues. Finally, two distinct questions, never addressed before in grapevine, were answered: how do phenomic PA performs compared to genomic PA? Can adding NIRS to genotypic data increase PA?

4.2 Material and Methods

4.2.1 Plant Material

Our plant material is composed of a diversity panel reflecting the whole genetic diversity of *Vitis vinifera* (Nicolas et al., 2016) and a half-diallel (Tello et al., 2019), better reflecting populations used in breeding programs.

The diversity panel is composed of 279 varieties, with an equal proportion of individuals from each of the three gene pools: Wine West (WW), Wine East (WE) and Table East (TE)

(Nicolas et al., 2016). This panel was overgrafted on Marselan in 2009, itself grafted on Fercal. Field location is in the Domaine du Chapitre experimental vineyard of Institut Agro | Montpellier SupAgro in Villeneuve-lès-Maguelone (South of France). The panel is replicated in five randomized complete blocks, each variety being represented by one plot of a single vine in each block.

The half-diallel is composed of 676 individuals from ten bi-parental populations (hereafter named crosses) in a half-diallel mating design between five parents: Syrah (S), Grenache (G), Cabernet-Sauvignon (CS), Terret Noir (TN) and Pinot Noir (PN) (Tello et al., 2019). All of them, except Grenache, belong to the WW gene pool (Brault et al., 2021b). Each cross comprises between 64 and 70 offspring. This population was planted in 2005 and grafted on Richter 110. Field location is the same experimental vineyard, a few kilometers away from the diversity panel field trial. The half-diallel is replicated in two randomized complete blocks, each offspring being represented by one plot of two consecutives vines in each block.

4.2.2 Phenotyping

We studied the same 15 traits in both trials, related to (i) berry composition at harvest, with malic acid (**mal.ripe**), tartaric acid (**tar.ripe**), shikimic acid (**shik.ripe**) concentrations, and shikimic / tartaric acid (**shiktar.ripe**) and malic / tartaric acid (**maltar.ripe**) ratios, (ii) morphological traits, with mean berry weight (**mbw**), mean cluster weight (**mcw**), mean cluster length (**mcl**), mean cluster width (**mcwi**) and cluster compactness (**clucomp**), (iii) phenology traits, with véraison date (onset of ripening, **verday**), harvest date (**samplday**) and the interval between véraison and harvest (**vermatu**), (iv) vigor (**vigour**). Details about phenotypic measurements, statistical processing and heritability can be found in Brault et al. (2021b). For prediction, we used the Best Linear Unbiased Predictors (BLUP) of genotypic values from Flutre et al. (2020) in the diversity panel and Brault et al. (2021b) in the half-diallel. Briefly, a mixed linear model was fitted for eliminating experimental confounding effects and in order to extract BLUPs of genotypic values. In the following, only BLUPs of genotypic values were used for the diversity panel, whereas the sum of genotypic and cross BLUPs were used for the half-diallel.

4.2.3 SNP genotyping

We used a set of 32,894 SNP markers common to both populations. Details about genotyping and marker processing are given in Tello et al. (2019) for the half-diallel and in Flutre et al. (2020) for the diversity panel. The selection of common SNPs was done in Brault et al. (2021b). 622 out of 676 individuals were successfully genotyped in the half-diallel, and 277 out of 279 individuals in the diversity panel.

4.2.4 Spectra measurements

Spectra were measured in both trials on dried wood and leaves collected during two consecutive years (2020 and 2021). For wood tissue, two shoots were cut per plot, on two vines in

the half-diallel and one in the diversity panel. These wood shoots were approximately 3 cm long. Wood was harvested on January 27^{th} in 2020 and January 14^{th} in 2021. For leaf tissue, four discs were sampled per plot, on two adult leaves per vine for two different vines in the half-diallel and on four leaves per vine in the diversity panel. Leaf disks had diameters of circa 1 cm and 0.5 cm in 2020 and 2021, respectively. Leaf tissue harvest occurred on July 1^{st} 2020 and June 16^{th} 2021. Two blocks were used in both trials, leading to a total of four wood shoots and eight leaf discs per genotype. After harvest, shoots and leaves were dried at $60°C$ until the weight stopped decreasing, and then stored in a cold chamber until measurements.

Spectra were measured with a reflectance probe plugged to an infra-red spectrometer (LabSpec 2500 Portable Vis/Nir spectrometer device; Analytical Spectral Devices, Inc.) and its associated software IndicoPro. A reference spectrum was taken twice a day, using Spectralon ®. For each wood shoot, two scans were taken, one on each end of the shoot. For each leaf disc, one scan was taken, on the adaxial surface. Thus, for each tissue, four scans were produced per plot (i.e., per genotype x block combination). Wavelengths ranged from 350 to 2,500 nm, with a 1 nm step. Each scan consisted in an average over 10 spectra, automatically computed during spectrum acquisition. In total around 1,800 and 5,400 scan were collected on the diversity panel and the half-diallel populations for each year and tissue, respectively.

4.2.5 Spectra processing

Spectra were processed separately within each trial. The first 50 wavelengths (visible range) were cut, because of instabilities. The average of the four spectra per plot were then carried out over the 2,101 remaining wavelengths. From these averaged raw spectra (**raw**), five pre-processing were then applied: smoothing (**smooth**) using Savitzky and Golay (1964) procedure, normalization or standard normal variate (**snv**) which consists in centering and scaling (Barnes et al., 1989), detrend (**dt**) for removing baseline (Barnes et al., 1989), and first and second derivative on normalized spectra (**der1**and **der2**, respectively), also for removing baseline and exacerbate some parts of the signal.

On each of these six spectra matrices (**raw**, **smooth**, **snv**, **dt**, **der1** and **der2**), we applied a mixed model over the reflectance at each wavelength, to compute variance components and derive NIRS genotypic BLUPs for each possible combination of three modalities at the tissue level times three modalities at the year level (Table 4.1).

The base mixed model was:
$reflectance_{ijk} = \mu + \underline{geno_i} + \underline{cross/subpop_i} + block_j + \underline{x} + \underline{y} + \epsilon_{ijk}$

With $reflectance$ the reflectance at a given wavelength, μ the intercept, $geno$ the random genotypic effect, $cross/subpop$ the random effect for cross (10 levels in the half-diallel) or subpopulation (3 levels in the diversity panel) effect, $block$ the fixed effect of block, x and y the random effects for plot coordinates, and $epsilon$ the residuals.

Factors could then be added to this base model, depending on the modality combination (Table 4.1).

	Wood	**Leaves**	**Wood + leaves**
2020	Base model	Base model	Base model $+\underline{geno:tissue} + \underline{cross:tissue}$
2021	Base model	Base model	Base model $+\underline{geno:tissue} + \underline{cross:tissue}$
2020 + 2021	Base model $+\underline{geno:year} + \underline{cross:year}$	Base model $+\underline{geno:year} + \underline{cross:year}$	Base model $+\underline{geno:tissue} + \underline{cross:tissue} + \underline{geno:year} + \underline{cross:year}$

TABLE 4.1 – Mixed model fitted, depending on the modality combination. *cross* effect is replaced by *subpop* for the diversity panel.

NIRS BLUPs used further were the sum of genotypic and cross or subpopulation BLUPs. For comparison, we also computed genotypic BLUPs from models without *cross* or *subpopulation* effects.

For comparison purpose and to evaluate the benefit of fitting a mixed model per wavelength to extract a genotypic BLUP, we also computed for each of the 6 spectra matrices (**raw**, **smooth**, **snv**, **dt**, **der1** and **der2**) the averaged spectra per genotype.

4.2.6 Variance components and co-inertia

Variance components from mixed models were extracted at each wavelength and compared between modality combinations and populations.

We also compared relationship matrices obtained independently from SNPs (that is, the genomic relationship matrix, GRM) and NIRS BLUPs (that could be called the phenomic relationship matrix), using co-inertia analysis (Dolédec & Chessel, 1994). Briefly, the co-inertia between two matrices X and Y (from SNP and wood NIRS for example) is computed as: $coinertia(X,Y) = trace(XQ_XX^TDYQ_YY^TD)$, with Q_X and Q_Y the weights associated with X and Y columns (SNP markers and reflectances), which were set to 1, and D the weights associated with X and Y rows (individuals), which were set to 1/n with n the number of individuals.

Then, a measure of correlation between X and Y can be computed as the RV coefficient: $RV = \frac{coinertia(X,Y)}{\sqrt{coinertia(X,X)}\sqrt{coinertia(Y,Y)}}$

We applied co-inertia analysis to SNPs, wood and leaf NIRS BLUPs, in order to estimate pairwise RV coefficients between these matrices.

4.2.7 Heritability assessment

Heritability values of phenotypic data were assessed for both populations in Flutre et al. (2020) for the diversity panel and in Brault et al. (2021b) for the half-diallel.

Heritability values were also assessed for reflectance data at each wavelength, after mixed model fitting. As for phenotypic data, heritability formula was:
$H^2 = \frac{\sigma^2_{geno}+\sigma^2_{cross}}{\sigma^2_{geno}+\sigma^2_{cross}+\frac{\sigma^2_{geno:year}+\sigma^2_{cross:year}}{n_{year}}+\frac{\sigma^2_x+\sigma^2_y+\sigma^2_\epsilon}{n_{rep}\times n_{year}}}$, with the variance components previously estimated in the mixed model, $n_{rep.year}$ the mean number of replicates per year, and n_{year} the number of year (one or two, depending on the modality).

4.2.8 Phenomic and genomic prediction models

Three methods were compared for the implementation of PP and GP, based on two models. Models were fitted separately for each population and trait.

RR-BLUP vs GBLUP/HBLUP model type

— In RR-BLUP, we fitted the following model: $y = X\beta + \epsilon$, with y the vector of genotypic BLUPs from phenotypic data, X the matrix for marker genotypes (additively coded as in Brault et al., 2021b) or wavelength data (from NIRS BLUPs for each of the nine above-mentioned modality combinations), β the marker or wavelength effects and ϵ the residual effects. This model was fitted using R/glmnet package version 4.1-2 (Friedman et al., 2010). In RR-BLUP, marker or wavelength effects are shrunk towards zero, according to a regularization parameter, chosen by an inner cross-validation (CV).

— In HBLUP (GBLUP model, using the NIRS relationship matrix H), we fitted the following model: $y = u + \epsilon$, with y the vector of genotypic BLUPs from phenotypic data, u the random effects for genomic or phenomic estimated breeding value, with $u \sim MN(0, \sigma^2_u K)$, K being the relationship matrix from markers or spectra, σ^2_u the genetic variance, and ϵ the random residual effect, $\epsilon \sim MN(0, \sigma^2_\epsilon I$, I being the identity matrix. $K = \frac{X_{sc}X^T_{sc}}{nb\,of\,X_{sc}\,columns}$, X_{sc} the previously described X matrix scaled on allelic frequencies or wavelength reflectances. This model was fitted using R/lme4GS package version 0.1 (Caamal-Pat et al., 2021).

Cross-validation

PP and GP models were assessed within each population and for each trait using CV. In the half-diallel, 10-fold CV was applied, while in the diversity panel, 5-fold CV was applied; CV was repeated 10 times. For each CV replicate, predicted values from all folds were combined and compared with observed genotypic BLUPs. We computed predictive ability (PA) as Pearson's correlation between the observed and predicted genotypic values. In the half-diallel, PA was calculated within each cross, as it was done in Brault et al. (2021b).

Multi-matrix model fitting

Using R/lme4GS allowed us to fit a single model involving several variance-covariance matrices, such as: $y = \sum_{j=1}^{q} u_j + \epsilon$, with $u_j \sim MN(0, \sigma_j^2 K_j)$, and K_j the relationship matrix from SNPs, wood NIRS or leaf NIRS previously described. We fitted this multi-matrix model using two or three variance-covariance matrices: SNPs + wood NIRS, SNPs + leaf NIRS, wood NIRS + leaf NIRS and SNPs + wood NIRS + leaf NIRS.

4.3 Results

4.3.1 Characterization of genetic signal in spectra

Variance components

Variance components for the nine modality combinations studied in each population are shown in Figure 4.1 for **der1** pre-processing. In both populations, genotypic variance was maximized in single-year and single-tissue analyses, with an essentially comparable magnitude between wood and leaves on one hand and 2020 and 2021 on the other hand. In multi-tissue analyses, genotypic variance drastically decreased and was mostly replaced by *geno:tissue* variance, while in multi-year analyses, genotypic variance was only partly replaced by *geno:year* variance. A strong x effect (row effect) was observed, while barely no y effect was present.

Comparing populations, the *cross* variance in the half-diallel was larger than the *subpop* variance in the diversity panel. The variance of interactions between *cross* or *subpop* and *year* or *tissue* remained low. The *geno:year* interaction was more important in the diversity panel than in the half-diallel.

Heritability

Genotypic variance results were consistent with heritability values calculated for each wavelength (distributions of heritability values for each pre-process are given in Figure S1). When comparing raw and pre-processed spectra, it was clear that the lowest heritability values generally corresponded to **raw** and **smooth** spectra. Heritability values for other pre-processes were close to each other, **der1** yielding the highest heritability values overall (Figure S1). Including both wood and leaf NIRS in the mixed model resulted in very low heritability values (Figure S2), hence we excluded this modality in the following analyses. The analysis wavelength by wavelength has showed that NIRS carry some genetic variance, with a moderate magnitude. To further characterize this genetic signal over the entire spectral range, we then carried out a co-inertia analysis between NIRS and SNP matrices.

Comparison of matrices from SNPs, NIRS and phenotypes, using co-inertia analysis

Co-inertia analysis was conducted on single-tissue modalities only. Figure 4.2 shows for each population the relative co-inertia between matrices of SNPs, wood and leaf NIRS BLUPs of

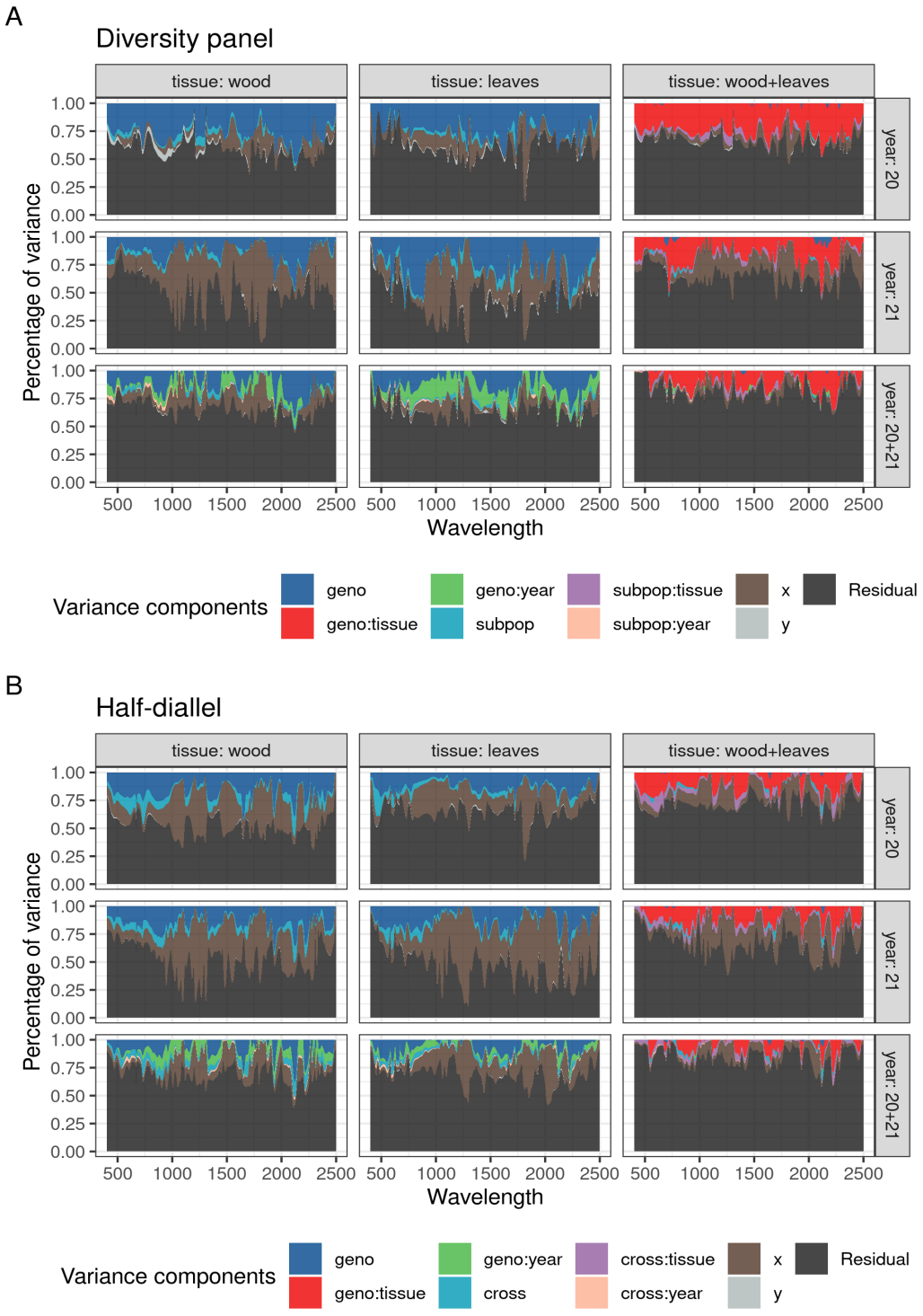


FIGURE 4.1 – Variance components from the mixed models applied to NIRS after **der1** pre-process. A: in the diversity panel population, B: in the half-diallel population.

genotype + cross or *subpopulation* effects for "2 years" modalities. For both populations, correlation with SNPs was similar between wood and leaf NIRS. However, this correlation was

nearly twice higher in the half-diallel than in the diversity panel. It is noteworthy that in both populations, the correlation between the SNP matrix and NIRS BLUPs matrices (obtained from wood or leaves) was higher than between the two NIRS BLUPs matrices obtained on wood and leaves.

We also carried out the co-inertia analysis with matrices derived from NIRS BLUPs of genotype effect for a model containing either a genotype effect only or both *genotype* and *cross* or *subpopulation* effects (Figure S2).

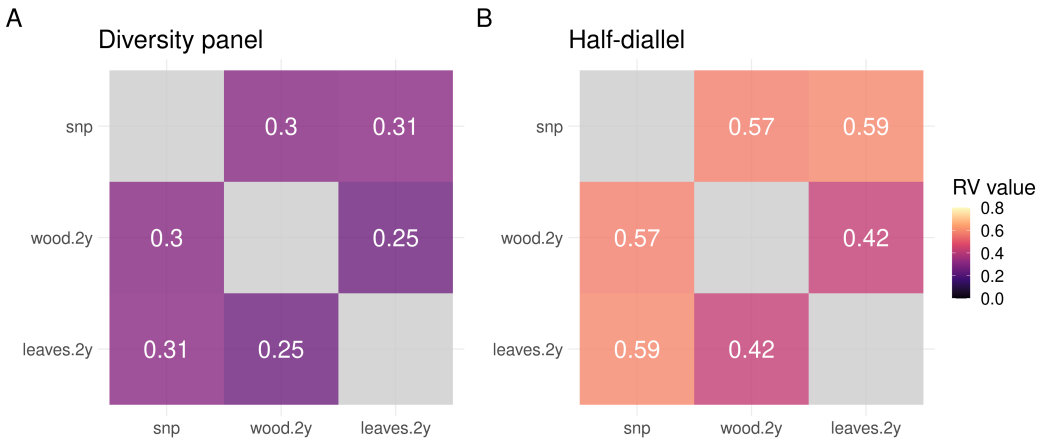


FIGURE 4.2 – Correlation between the genomic relationship matrix ("snp") and the relationship matrices derived from wood and leaf NIRS BLUPs of genotype + cross or subpopulation effects with both years included in the mixed model ("wood.2y", "leaves.2y", respectively). A: in the diversity panel, B: in the half-diallel.

Using such matrices strongly decreased correlation with the SNP matrix, as compared to using matrices derived from BLUPs of *genotype* + *cross* or *subpopulation* effects (Figure S2). Therefore, in subsequent prediction analyses, we used only the latter matrices including *cross* or *subpopulation* effect. Matrices from multi-year NIRS BLUPs generally displayed a slightly higher correlation with the SNP matrix than the single-year BLUPs, and this effect was more pronounced in the half-diallel (Figure S3).

4.3.2 Phenomic prediction using BLUPs vs *base* spectra and RR-BLUP vs HBLUP

In each population, using spectra BLUPs instead of *base* spectra almost always resulted in higher PA (Figure 4.3). However, differences were observed depending on the method and population. The method yielding the highest PA was HBLUP (implemented with lme4GS) in the half-diallel and RR-BLUP (implemented with glmnet) in the diversity panel. However, it is worth mentioning that differences between methods were found to be more pronounced in the half-diallel than in the diversity panel. The highest differences between *base* spectra and BLUPs were observed for the best method in each population.

Thus, we retained spectra BLUPs in all cases, lme4GS in the half-diallel and glmnet in the diversity panel.

We observed higher variance in PA in the half-diallel than in the diversity panel, because in the half-diallel, PA distribution was over 10 crosses in addition to the 6 years x tissues modality combinations retained (see above). Average PA for the best method was slightly higher in the half-diallel (0.31) than in the diversity panel (0.26).

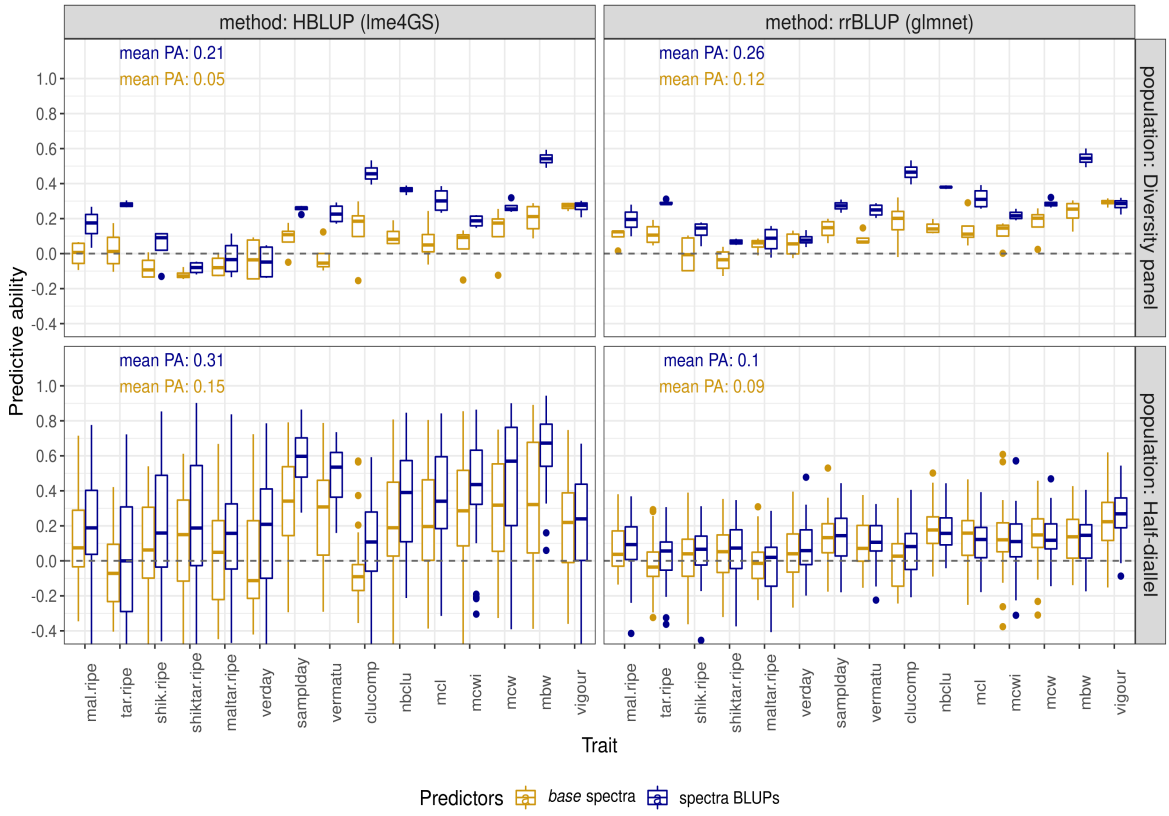


FIGURE 4.3 – Predictive ability of phenomic prediction for 15 traits, in two populations with two methods, using base spectra (in golden) or spectra BLUPs (in dark blue) after der1 pre-process. For each trait, PA distribution was over the 6 modalities retained for years and tissues (and also over the 10 crosses in the half-diallel).

We compared PA for all pre-processes, after selecting the best method for each population (Figure S4). We found that **der1** and **der2** pre-processes gave close results, with a slight superiority of **der1** overall. Therefore, we kept only this pre-process in subsequent analyses.

4.3.3 Phenomic prediction using NIRS collected over one or two years and tissues

We further compared PP models including a single vs both NIRS BLUP matrices obtained from wood and leaves. For each tissue configuration, we used the NIRS BLUPs derived from the above-described year modalities (2020, 2021 or both). For single tissue configurations, we used the best method selected above in each population, and one NIRS BLUP matrix was

fitted. For the wood+leaves configuration in both populations, two NIRS BLUP matrices (one for wood and one for leaves) were fitted using lme4GS package.

For both populations, the nine configurations tested resulted in close PA distributions (Figure 4.4). Yet, "two years" modalities and “two tissues” configurations overall gave the best average PA values. We thus retained only multi-year and multi-tissue PP results for subsequent comparison with GP.

Finally, PA was slightly higher in the half-diallel on average, with a larger variance originating from differences between crosses (see hereafter).

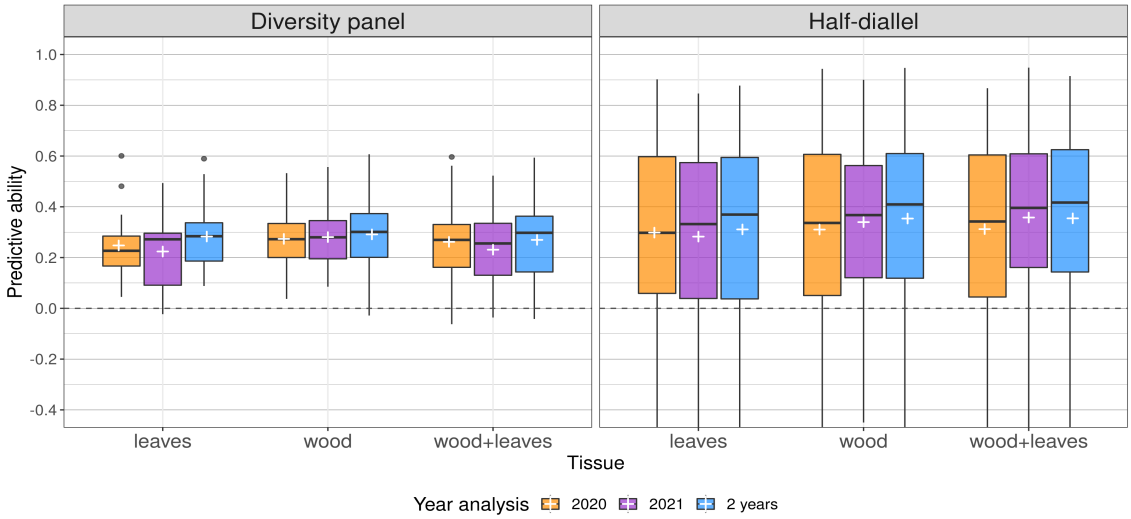


FIGURE 4.4 – Predictive ability of phenomic prediction with a single vs both years and tissues, over the 15 traits in both populations and the 10 crosses in the half-diallel. Prediction models were fitted with glmnet in the diversity panel (except for wood+leaves configuration) and with lme4GS in the half-diallel. In both populations, models were carried out after der1 pre-processing. The white cross indicates the average PA for each configuration.

If we now turn to details per trait, the results show that, within the diversity panel, each trait displayed nearly the same PA for the different tissue configurations, for PA values above 0.2 (Figure S5). In the half-diallel, there were larger differences between tissue configurations. However, this factor still had far less impact on PA than cross or trait.

Overall average PA of PP for "2 years" and "wood+leaves" configuration was 0.27 in the diversity panel and 0.35 in the half-diallel (Figure 4.4). PA values per trait ranged from -0.04 for **shiktar.ripe** to 0.59 for **mbw** in the diversity panel (Figure S5A), and from 0.09 for **mal.ripe** to 0.72 for **mbw** in the half-diallel (note that in the half diallel, PA values per trait are averaged over the 10 crosses). However, large differences in PA of PP were observed within a trait at the cross level in the half-diallel for “2 years” and “wood+leaves” configuration, such as for **tar.ripe**, from –0.49 for GxCS to 0.74 for TNxS (Figure 4.5B and Figure S5B). Comparatively, differences at the cross level were lower for GP (Figure S6). The best predicted cross with PP over all traits was GxS (average PA of 0.41) and the worst one was TNxG (0.29) (Figure S7).

For some crosses and traits, PA values could be above 0.8, the maximum PA for PP being 0.91 for **mbw** and TNxG (Figure S5B).

4.3.4 Comparison of PP with GP

Before comparing PP with GP, we applied GP on both populations with the two methods previously compared for PP (Figure S6). We found that lme4GS was overall the best method in both populations, hence we retained this method for the following comparison. Like this was the case for PP, differences between methods appeared to be more pronounced in the half-diallel than in the diversity panel.

The PA reached by PP was close to that of GP in both populations for a few traits and for some half-diallel crosses (**samplday**, **vermatu** and **mbw**) (Figure 4.5). PP even outperformed GP for some crosses and traits in the half-diallel, such as for CSxPN, GxCS, GxPN and **vigour**, or GxCS, SxPN, TNxPN and **clucomp**. Differences in PA between PP and GP were lower in the diversity panel than in the half-diallel.

In the diversity panel, PA of PP was significantly higher (non-overlapping error bar) than PA of GP for one trait (mcl) and non-significantly different for two other traits (**clucomp** and **vermatu**) (Figure 4.5A). In the half-diallel, PA of PP was significantly higher than PA of GP for 28 trait x cross combinations out of 150, while this difference was not significant in 17 other cases (Figure 4.5B). In all other cases, PA of PP was lower than PA of GP.

In Figure 4.6, we further compared mean PAs of PP and GP per trait in each population. In both populations, the slope of the regression model was close to 1 and the intercept to –0.2. This suggests that PA of PP and GP follow the same ranking, independently of the trait. However, this regression had a much lower R^2 in the half-diallel than in the diversity panel.

4.3.5 Enhancing genomic prediction using NIRS

Another possible way of using NIRS is to add it into the predictive model together with SNPs, in order to increase PA. We thus implemented multi-BLUP models with SNPs and NIRS BLUPs and compared them to GP models in each population.

Overall, for both populations and for all traits, differences in PA between GP and different combined GP+PP models were small (Figure S8). In the diversity panel, combining wood NIRS with SNPs led to the best PA (0.405), closely followed by leaves NIRS + SNPs (0.403), wood NIRS + leaves NIRS + SNPs (0.402) and SNPs alone (0.400). In the half-diallel, SNPs alone gave the highest PA (0.595), followed by wood NIRS + leaves NIRS + SNPs (0.587), leaves NIRS + SNPs (0.576), and wood NIRS + SNPs (0.569).

Nevertheless, adding NIRS to a predictive model could lead to minor (non-significant) improvements in PA for some traits, compared to classic GP. Combining GP + PP from wood NIRS slightly increased PA over the GP model for two traits in the diversity panel (**clucomp** and **mcl**) (Figure S8A). In the half-diallel, the difference in average PA with GP was much

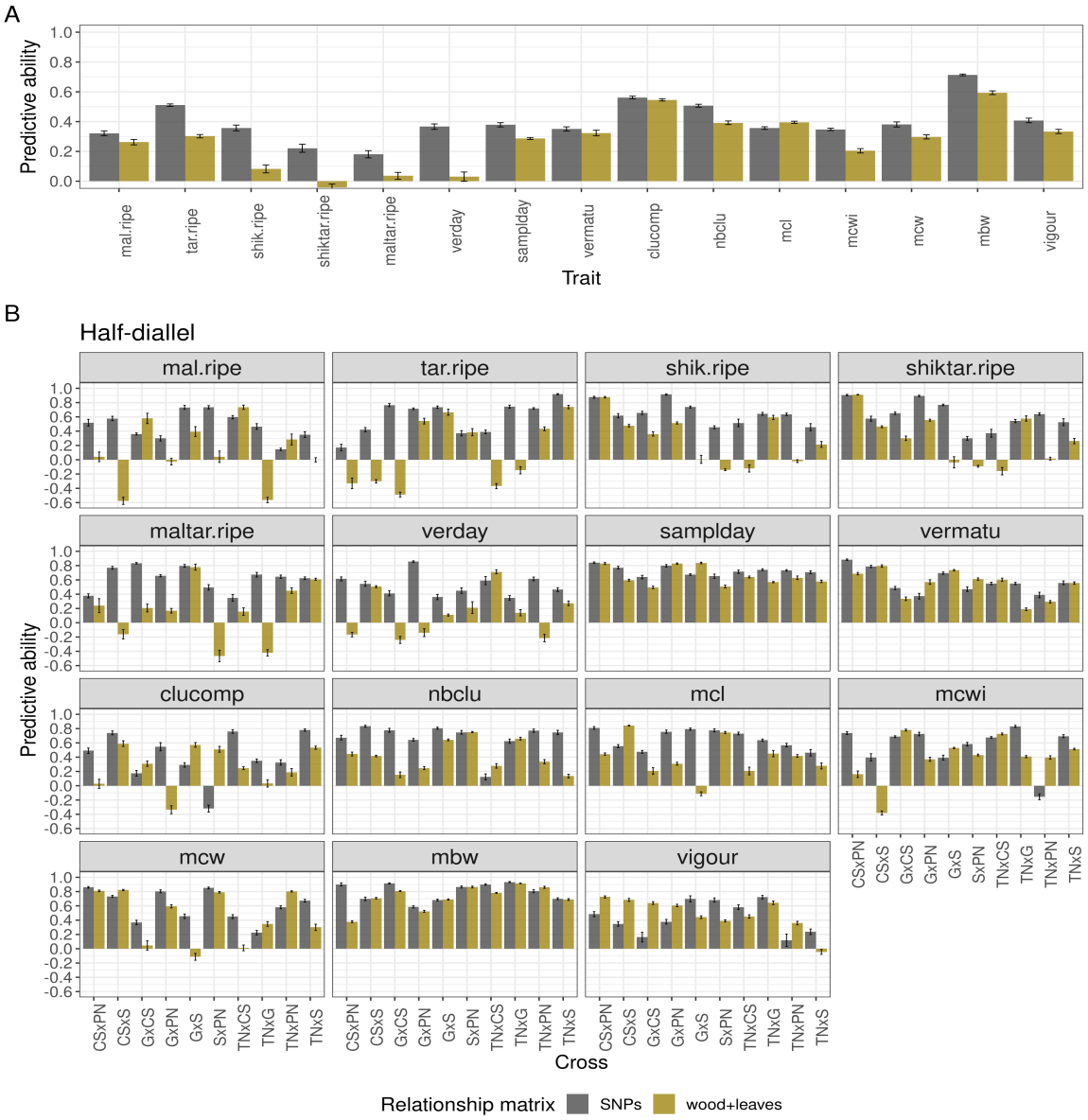


FIGURE 4.5 – Predictive ability in two settings for 15 traits with lme4GS: SNPs: GP; wood+leaves: PP with two variance-covariance matrices, for wood and leaf NIRS, for "2 years" NIRS BLUPs derived after der1 pre-process. Prediction models were fitted with lme4GS. Error bars correspond to the 95% confidence interval around the mean, based on CV repetitions.

more variable among traits, with an increase for vigour, clucomp, vermatu and samplday, and a decrease for **mal.ripe**, **tar.ripe**, **shik.ripe**, **shiktar.ripe**, **maltar.ripe**, **nbclu**, **mcl** and **mcwi** (Figure S8B).

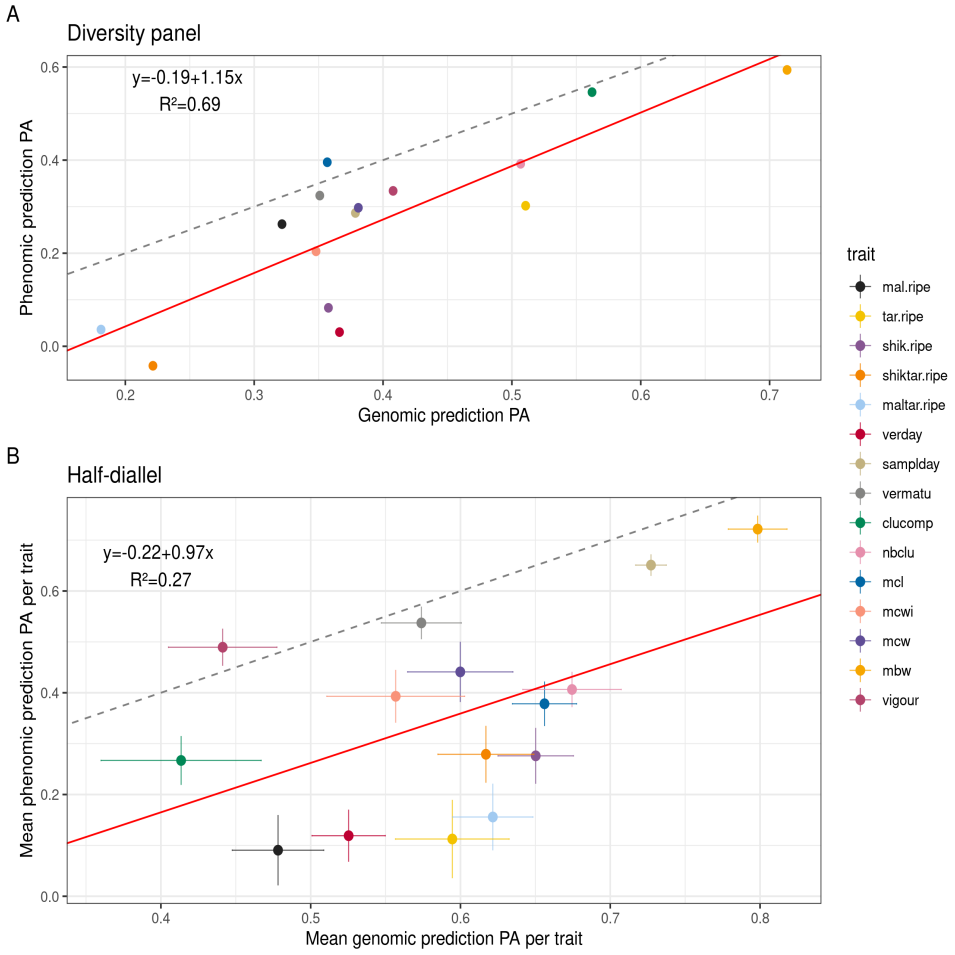


FIGURE 4.6 – Predictive ability of phenomic against genomic prediction. A: in the diversity panel, B: in the half-diallel. The regression formula corresponds to the following linear model: PA of $PP \sim PA$ of GP, with 15 observations (traits). In the half-diallel, PA was averaged across the ten crosses, hence standard error around each point is displayed. The red line is the regression line and the gray dashed line is the identity line.

4.4 Discussion

So far, PP has only been implemented in a reduced number of species and traits. This study provides the first use of PP in grapevine, within two complementary populations: a diversity panel and a half-diallel. Besides, we tested PP for 15 traits, belonging to four categories: berry composition, phenology, morphological traits and vigour. We first showed that NIRS variability was partly of genotypic origin. We then tested several parameters (mean vs BLUP, tissue, year, method), to optimize both PP and GP. Finally, we found that PP could yield PA values close to or even higher than GP ones.

4.4.1 NIRS variance components and co-inertia with SNPs

Genotype and derived interaction variables had a fairly moderate impact on total variance observed between spectra (Figure 4.1). The genotypic effect was best captured in single-tissue analyses. This was not surprising, because the genetic signal at a given wavelength relies on molecules specific to each tissue. Then, mixing both tissues into a single model led to no overall genetic effect and to strong *geno:tissue* interaction. This tendency was also observed, to a smaller extent, in the multi-year analyses. This also suggests that different tissues bring non redundant genetic information. This was confirmed by co-inertia analysis, which evidenced that NIRS matrices from wood and leaves were more correlated to the SNP matrix than to each other.

Interestingly, co-inertia analysis showed that multi-year NIRS BLUP matrices were slightly more correlated with the GRM than single-year ones, despite lower genotypic variance. This implies that the genotypic part of NIRS estimated by multi-year analysis could be more related to the genetic signal. Thus, genetic signal ignoring genotype-by-environment interactions could be better captured when several years are combined, this was also the case in Galán et al. (2021) for which multi-year spectra resulted in higher PA values.

Comparatively to genotype-related effects, among non-residual variance components, “x” effect displayed a large variance along wavelengths (Figure 4.1). This effect actually corresponds to a row effect and might be due to the experimental design. Indeed, leaf discs and wood shoots were both sampled and scanned row by row. However, we cannot determine whether this “x” effect comes from the tissue sampling, i.e., sampling time (over a day), soil heterogeneity; or from the NIRS measurement step, i.e., device calibration, differential storage time, air humidity. Our results underline the importance of accounting such experimental effects in order to improve the genetic signal capture and thus prediction. In further experiments, one could increase the number of spectra per plot and randomize NIRS measurements, in order to determine if the "x” effect observed here was due to measurement or sampling and to reduce it. Other studies that fitted a linear model for each wavelength did not introduce field coordinates as effects (e.g. (Galán et al., 2020; Krause et al., 2019; Lane et al., 2020). But the first and last studies were based on hyperspectral images taken with aircraft flights, that is with an experimental design less prone to plot location effect, and the second study fitted a linear model with only block and environmental effects.

Galán et al. (2020) found a mean heritability value of 0.73 for wavelength reflectances, which is substantially higher than the values we observed (Figure S1). However, we did not use the same heritability formula. Montesinos-López et al. (2017b) also reported overall higher heritability values ranging from 0.6 to 0.8 for most time points, with strong variations depending on the environment (water condition) and time-point.

We found higher heritability and genetic variance in the diversity panel than in the half-diallel. Yet, PA were generally higher in the half-diallel. In Rincent et al. (2018), genetic variance estimates per wavelength between wheat and poplar were consistent with PA in

these species, i.e., they evidenced higher PA values in wheat than in poplar. On the opposite, our results on co-inertia analysis were consistent with PA values: correlation between SNP and NIRS BLUPs matrices was higher in the half-diallel than in the diversity panel (Figure 4.2). This suggest that co-inertia analysis is more relevant to compare modalities for NIRS BLUP than variance components.

The higher correlation observed between SNP and NIRS BLUP in the diallel with respect to the diversity panel is likely to be explained by the higher genetic structure in the half-diallel, or because the half-diallel is in better health than the diversity panel, which is older and over-grafted. Actually, it was surprising that NIRS could capture genetic structure, i.e., in our case the *subpopulation* effect in the diversity panel and the *cross* effect in the half-diallel. Although variance components for *subpopulation* and *cross* remained moderate (Figure 4.1), adding the corresponding BLUP effects to genotypic effects led to a sharp increase in correlation between NIRS and SNP matrices (Figure S2). Further in-depth studies are required to better understand whether this observation could be specific to some subpopulations or families.

4.4.2 Optimizing PP

Among the parameters tested, some had substantial impact on PA, while others had only negligible impact. Namely, using NIRS via BLUP analysis instead of merely average spectra per genotype led to a strong increase in PA (Figure 4.3). This was probably associated with the strong x effect we observed in variance analysis. Such a difference had never been reported before, as studies obtained PP results either from *base* (such as Cuevas et al. (2019), Rincent et al. (2018)) or BLUE (such as Krause et al. (2019), Lane et al. (2020)) spectra, without comparing both modalities.

Surprisingly, the prediction method also had notable impact on PA: using RR-BLUP or HBLUP / GBLUP models gave different PA in the half-diallel, while differences in PA between methods were lower in the diversity panel (Figure 4.3). Yet, HBLUP/GBLUP and RR-BLUP models are expected to perform similarly when the regularization parameter in ridge regression is equal to $\sigma_e^2/sigma_g^2$ (Habier et al., 2007). In our analysis, this parameter value was chosen by cross-validation using cv.glmnet function. The higher relatedness between genotypes within the half-diallel than within the diversity panel (Brault et al. (2021b), Figure 1a) may boost HBLUP and GBLUP models compared to RR-BLUP in this population. In future investigations, one could use variable selection method such as LASSO to select the most relevant wavelengths for computing the relationship matrix from NIRS BLUP. Such variable selection was performed by Galán et al. (2020) and resulted in higher PA.

On the opposite, using single-year, single tissue, multi-year, or multi-tissue NIRS BLUPs and all pre-processes except **smooth** gave very similar results over all traits and crosses (Figure 4.4), with a slight superiority of multi-year modality overall. This was consistent with the results of co-inertia analysis (Figure S3). In Rincent et al. (2018), the multi-tissue analysis for wheat with leaf and grain combined gave similar PA as for single-tissue analysis. As the

combination of two tissues for PP was only done in one other study (Rincent et al., 2018), further work needs to be done to assess these conclusions.

For a given trait, both tissues tested gave similar PA for the diversity panel (Figure S5A). For the half-diallel, more differences were observed between tissues, and much larger differences were observed between crosses (Figure S5B). However, no cross was consistently well or poorly predicted for all traits, suggesting a strong cross x trait interaction. These large disparities among crosses were consistent with the GP results obtained in the same population by Brault et al. (2021b).

4.4.3 Comparison between PP and GP

PP is supposed to better account for GxE than GP. However, it was shown in Rincent et al. (2018) that PP could still reach good PA values when NIRS for TS were taken in an environment different from the one in which VS was phenotyped, i.e., when accounting for GxE was not possible. In this study, we could not assess whether PP accuracy partly relied on location-related GxE, because phenotypes and NIRS came from a single location. Nevertheless, phenotypes were measured in 2011-2012 and 2013-2017 in the diversity panel and half-diallel populations, respectively, whereas NIRS were measured in 2020-2021 in both populations. Vintage (year) effect is also part of GxE and it is likely that 2020 or 2021 could display some differences in terms of weather with phenotyping years. For training and validation model, we used genotypic BLUPs of both phenotypic data, thereby removing year and *geno:year* effects. We found that PA seemed not to be impacted by NIRS year for all traits studied, suggesting that vintage has a negligible effect on PA when genotypic BLUPs are used.

As a prospect, one could specifically extract genotype x year and genotype x location variance components from phenotypic and NIRS data and test if PP could be useful to predict this GxE part. Montesinos-López et al. (2017a) studied GxE using spectra and they compared models including or not GxE and wavelength-by-environment interaction. They reported that the inclusion of GxE provided no increase of PA while including wavelength-by-environment interaction was the best modality.

We found that PP could compete with GP for some traits in both populations, despite moderate genetic variance estimated from NIRS. However, the number of traits for which PP outperformed GP remained low. These results were close to those of Rincent et al. (2018) on poplar. In our case, one explanation could be that NIRS came from tissues sampled in 2020 and 2021, while phenotypes were measured in 2011-2012 and 2013-2017 in the diversity panel and half-diallel, respectively. Thus, we couldn't take into account for GxE from vintage effect. As a perspective, it would be interesting to compare PA when spectra are measured the same year as phenotyping or not. In such case, one could explicitly model vintage effects in spectra to further increase PA.

Nevertheless, even when PP does not outperform GP, it may still be interesting in breeding, because of its lower cost and increased throughput compared to genotyping. Moreover, when

a trait was well-predicted with GP, we found that it was also well-predicted with PP, with a global shift of –0.2 in PA (Figure 4.6). This suggests that PP PA truly relies on genetic variability and that PP could be applied indifferently for all traits. Even though this study is the first one implementing PP on so many traits (15), these conclusions remain to be confirmed on other species and traits. Based on data from Rincent et al. (2018) and on the relative GP and PP reliability that we observed, we are still expecting a positive genetic gain by switching from GP to PP.

We implemented that setting, in order to test whether combining NIRS and SNP could increase PA compared to GP, by taking other genetic effects into account. However, as we used NIRS BLUPs, we only maximized the genetic variance part of spectra, we thus intentionally excluded GxE. Therefore, the fact that adding NIRS to GP model did not result in any increase in PA is consistent with our spectra processing. Cuevas et al. (2019) and Galán et al. (2020) found slight to noticeable improvement in PA when NIRS was added to the model, compared to GP model with SNPs only; difference in PA was at most 0.01 in Cuevas et al. (2019) and up to 0.1 in Galán et al. (2020). Both studies are however so different than ours that it is difficult to explain these different behaviors.

As a conclusion, we provided the first implementation of PP in grapevine. The number of traits studied allowed us to put forward a correlation between PA of GP and PP, suggesting that PP relies on a genetic basis. Such a correlation was never reported before. We expect that the shift of PA between PP and GP of –0.2 would be reduced if year of phenotyping and spectra measurement are the same. Still, PP has shown its interest for breeding over a wide range of traits.

Chapter 5

General discussion

Over chapters 2, 3 and 4, I applied various GP and PP models to three complementary populations for a range of traits. The overall goal was to provide guidelines for improving selection of new grapevine varieties by using genomic or phenomic prediction. More specifically, I found that (i) multivariate GP methods only provided a slight increase in PA over univariate ones, (ii) decoupling PA into cross mean and within-cross predictions was a requirement to better understand which parameters affect across-population PA, (iii) using NIRS instead of SNP markers allowed estimating genetic relatedness and predicting phenotypes with slightly and consistently lower PA over the studied traits. Overall, I found that both GP and PP could be useful to speed up grapevine breeding programs.

5.1 Strengths and weaknesses of the three populations and phenotypic data studied

In my thesis, I analyzed data from three complementary grapevine populations: (i) a bi-parental population, from a reciprocal cross between Syrah and Grenache, hereafter referred to as SxG, (ii) a multi-parental population, from a half-diallel mating design between five parents, Syrah, Grenache, Cabernet-Sauvignon, Pinot Noir and Terret Noir, (iii) a diversity panel, composed of 279 varieties chosen to be representative of *V. vinifera* L. genetic diversity. I used SxG in chapter 2 and the two other populations in chapters 3 and 4. Here I discuss and compare the interest of each population and the associated available phenotypic data for evaluating the potential interest of genomic and phenomic prediction in grapevine breeding. These three populations were planted and phenotyped before my thesis and I tried to take advantage of the best of each population to answer my research questions.

5.1.1 SxG progeny

The advantage of a bi-parental population is its genetic homogeneity, with no genetic structure. Therefore, this population was adapted to the statistical design and comparison of GP methods in a simple context with controlled genetic parameters. Besides, segregating alleles at QTLs are necessarily coming from one of the two parents, which makes it easier to dissect genetic architecture. Indeed, I also used this population for QTL detection and candidate

gene exploration in chapter 2. It was therefore adapted for studying the genetic basis of many traits of interest in grapevine breeding. Moreover, with its large size (N = 188) and narrow genetic diversity (only full-sibs), this population was well-suited for assessing maximum potential PA by within-population cross-validation.

However, QTLs found and marker effects estimated in this population cannot easily be extrapolated or transferred to other breeding populations. Some desirable alleles might not be present in Syrah and Grenache, thus extending prediction to other parents and crosses was required.

5.1.2 Half-diallel

The half-diallel population (5 parents) corrects the main drawback of SxG, since it has a higher genetic diversity, with potentially more alleles at each locus, and therefore more possibilities for transferring QTLs results and prediction equations to other breeding populations. In particular, the half-diallel design allowed to test prediction between half-sibs, which can be interesting in some breeding schemes (see 5.5.1). Moreover, the differences between crosses were very interesting to study and model, especially given that the five parents displayed varying genetic relatedness with each other.

However, its large size (N = 622) might be prohibitive for setting up trials. On the other hand, each cross of the half-diallel had between 64 and 70 offspring, making it difficult to apply within-cross CV. Indeed, evaluating PA in these small-sized crosses may be questionable, because of sampling bias in small samples. It is noteworthy that one of the ten families is a cross between Syrah and Grenache, with different individuals from those of the SxG population. PA in these two populations could be compared to assess the extent of this sampling bias in PA estimation.

5.1.3 Diversity panel

This population was designed to bear high genetic diversity, each grapevine subpopulation was balanced, thus reflecting the small genetic structure existing within *V. vinifera* (Bacilieri et al., 2013). This population of 279 individuals is closer to what we expect from a universal TS, which should encompass all genetic diversity. Nevertheless, the relatedness of the current wine breeding programs, i.e., bi-parental crosses from WW subpopulation, with this panel is lower than with the half-diallel. And LD was much lower in the panel, leading to lower PA given the number of markers available. Still, this panel was useful to estimate the interest of new potential crosses.

A new population is currently being developed to correct some drawbacks of the diversity panel as a universal TS, by specifically including parents of current breeding programs. This new population (SelGenVit panel) consists mostly in wine varieties, with the latest INRA-ResDur (as presented in 1.1.3) varieties bearing several resistance genes, as well as other

varieties with new sources of resistance to Pierce's disease or black rot (Rex et al., 2014; Riaz et al., 2009), and varieties adapted to climate change.

5.1.4 Comparison between these three populations

Each of the three populations presented above have advantages and drawbacks for their use in genetics and breeding (Figure 5.1).

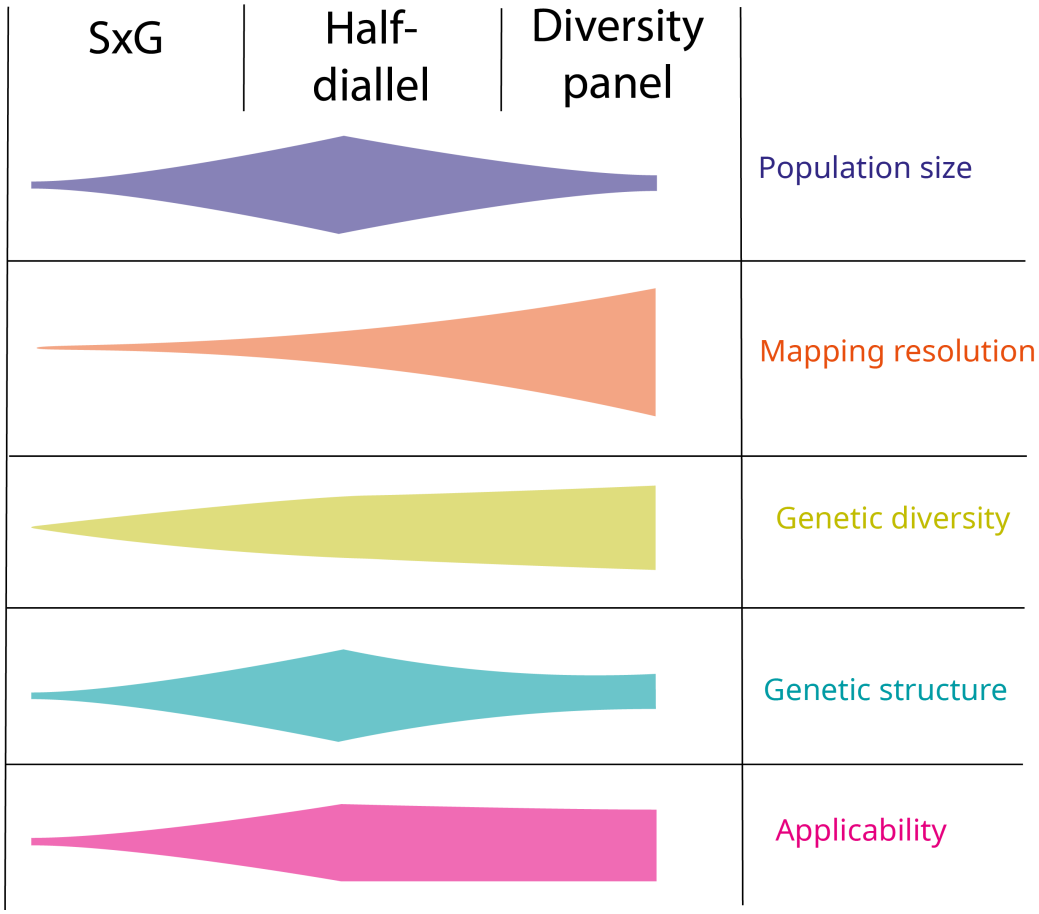


FIGURE 5.1 – Comparison of the three grapevine populations studied. Applicability refers to the usefulness of each population for the future grapevine breeding programs for wine.

Concerning traits measured in the different populations, the diversity panel had the highest number of traits (128), many of which were related to anthocyanins and tannins. The SxG population was phenotyped in a high-throughput phenotyping platform under semi-controlled conditions, with water stress applied to part of the plants. In SxG and diversity panel field trials, water stress was also applied to half of the blocks in some years. The half-diallel population had the fewer number of traits phenotyped (26). The 15 traits in common between the half-diallel and diversity panel populations have been studied. This number of traits allowed us to shed light on interesting correlations in chapters 3 and 4, even though

they remain to be confirmed with a larger number of traits. More traits are thus available for some populations and could be added in further analyses.

5.2 Comparison of results with other published studies

5.2.1 Specificities of grapevine for genetics and breeding

Quantitative genetics and genomic selection have been widely applied to cattle and cereal species. Methods have been specifically developed to work on these species, e.g., adapting to heterozygous genome in cattle, to homozygous genome in wheat, or to hybrid development in maize. Unlike most cultivated plant species, grapevine has a highly heterozygous genome with strong inbreeding depression, thus back-crossing and homozygous line development is not possible. Moreover, grapevine is a perennial species, with a generation time of three to four years, compared to a few months for some cereals. This implies that field trials require much more time, but also space and money to get the same amount of phenotypic data. However, once a population is planted in the field, phenotypic data can be collected over several years. Grapevine is a plant with vegetative propagation, clones of a genotype can thus be tested in different environments and crossing may involve varieties that have been existed for centuries, unlike cattle individuals.

These fundamental differences partly explain why genetic gain is lower in grapevine than in other species. However, a few authors had already implemented GS in grapevine, and found it promising to improve breeding (Flutre et al., 2020; Fodor et al., 2014; Migicovsky et al., 2017; Viana et al., 2016), which I confirmed with more traits, populations and configurations. Despite its specificities, grapevine shares similarities with other heterozygous perennial species such as fruit or forest trees.

5.2.2 Comparison of results with other species

Comparison between univariate and multivariate GP models

Previous studies reported better performance of multivariate GP models over univariate ones (Calus & Veerkamp, 2011; Dagnachew & Meuwissen, 2019; Guo et al., 2014; Jia & Jannink, 2012; Jiang et al., 2015). In chapter 2, I found only minor improvement in PA when comparing multivariate to univariate models on experimental data and no improvement on simulated data. However, all the previously cited studies had a p/n ratio lower than the one I had in SxG. These results suggest that TS size is a possible parameter to explain the relative performance of multivariate models over univariate ones (Brault et al., 2021a; Obozinski et al., 2011). As in grapevine, large TS sizes will probably never be common due to cumbersomeness of field trials, univariate methods were preferred in subsequent chapters.

The present study was the first large-scale study on the comparison of statistical methods for GP in grapevine, in particular multivariate vs univariate ones, given the number of traits and methods studied.

Across-population genomic prediction

In chapter 3, I applied across-population GP, by using the diversity panel as TS and half-diallel crosses as VS. Before comparing my results with other species, it is useful to outline that I used the "across-population" term because there was no direct relationship between TS and VS and because, of the three subpopulations of the diversity panel, only the WW one was more related to the half-diallel.

Indeed, across-population GP results vary depending on species and how "population" is defined. In cattle, "across-population" often refers to across-breed, these breeds being much more differentiated than grapevine subpopulations. In other plant species, such differentiated groups of varieties are called heterotic groups, as in maize. Crosses are made within each heterotic group to produce lines which will then be crossed across heterotic groups to give hybrids with superior phenotypes, taking advantage of heterosis. Thus, the PA values I obtained with across-population prediction are hardly comparable with these species. Still, comparison with other studies on heterozygous perennial species using a genetic resource panel as TS is possible, as in apple with Roth et al. (2020) and Cazenave et al. (2021).

The present study is the first so far focused on across-population GP on experimental data in grapevine for several different traits. Besides, analyzing separately cross means and within-cross genotypic values was not done in many studies before. Finally, the search for parameters affecting PA and TS optimization for cross mean prediction had never been done before in grapevine.

Phenomic prediction

To the best of my knowledge, PP results have only been published for wheat (Cuevas et al., 2019; Krause et al., 2019; Rincent et al., 2018), rye (Galán et al., 2020, 2021), maize (Lane et al., 2020) and poplar (Rincent et al., 2018). Overall PA values reported in chapter 4 for grapevine were slightly lower than those for these species, except for poplar (Rincent et al., 2018).

In chapter 4, I provided for the first time a thorough characterization of genetic signal in spectra with estimation of variance components and co-inertia analysis between SNPs and NIRS matrices, showing that NIRS capture a large part of genetic signal. I also evidenced that using BLUPs of spectra instead of *base* spectra gave notably better PA, which had never been reported before.

5.3 Components of the predictive model and prediction optimization

The predictive model used in chapters 2, 3 and 4 was: $y = X\beta + \epsilon$ (RR-BLUP version), with y the vector of BLUPs of genotypic values from phenotypic data, X the design matrix for marker data for GP or reflectance for PP, β the vector of marker or reflectance effects and

epsilon the residual vector. I will describe in what follows these components and how they were modified or optimized in my thesis.

5.3.1 Phenotype measurement

Phenotyping remains the strongest bottleneck in breeding (Crossa et al., 2021), especially in grapevine, due to cumbersomeness of field trials, long juvenile phase and highly transformed final product. However, precise phenotyping is crucial for genetic determinism studies and prediction accuracy. In grapevine, a given population is planted for several years. Hence, phenotypic data are often measured during several years on each vine. I used BLUPs of genotypic values for each trait, by fitting a model including both design and year effects. Thus, compared to other species, we often have phenotypic data on less individuals per genotype (for a given year), but this is compensated by the fact that phenotypic data are collected over several years.

In some cases, I used traits derived from a combination of phenotypes, e.g., malic / tartaric acids in chapters 3 and 4, or the normalized difference in chapter 2. In some cases, it resulted in an increase in broad-sense heritability, which could be explained because the resulting trait is actually closer to a physiological response.

We used NIRS instead of genotyping in PP, but a much more classical use of NIRS is as a high-throughput phenotype, i.e., as a proxy for a trait of interest difficult to measure (Crossa et al., 2021). Using hyperspectral imaging is interesting for predicting cereal yield (Rutkoski et al., 2016). In grapevine, traits of interest are mainly cluster and berry composition traits. Still, the use of portable spectrometer on cluster is relevant for high-throughput fruit phenotyping (Diago et al., 2015). Another way to improve phenotyping throughput at the wine level is to use new micro-vinification devices (Ducasse et al., 2019).

We found in chapters 2 and 3 that broad-sense heritability was a major driver of PA for within-population GP. Thus, increasing phenotyping precision is a substantial lever for getting higher PA, even though when TS and VS were less related, other parameters were also involved in determining PA. Increasing heritability could be done by raising the number of clone and block replicates, the number of years, or by refining trait measurement.

5.3.2 Input matrix structure

Genomic matrix

In chapter 2, I directly used a genomic design matrix based on SSRs or SNPs, without deriving a genomic relationship matrix. I did not find any improvement in PA when including dominance allelic effects into this matrix together with additive ones. However, modelling both additive and dominance effects in a single matrix prevents from using different variances for additive and dominance allelic effects, which could hamper PA. Other statistical models, multi-kernel or multi-matrix based, allow to handle several different variance-covariance matrices. I used such methods in chapter 4 with lme4GS (Caamal-Pat et al., 2021),

when I fitted NIRS matrices from two tissues or combined NIRS and SNP data in the same model. Dominance effect could thus be tested again using such models. Non-additive effects such as dominance and epistasis can be incorporated in GP models using non-linear methods. Indeed, I tested to estimate marker effects with Gradient Boosting but it did not give higher PA in chapter 2. However, this does not imply that non-additive effects are negligible (Ubbens et al., 2021). Other ways of incorporating non-additive effects in GP models include the reparametrization of the genomic relationship matrix (Vitezica et al., 2013; Vitezica et al., 2017).

In chapter 2, I also estimated specific additive effects for each of four imputed alleles at a given locus, but this did not result in higher PA compared to using a bi-allelic matrix coded additively. Using haplotypes is another approach for taking into account both allelic diversity and LD extent, which could be interesting when training GP models in diversity panels; however, it resulted in no or very small increase in PA in previous studies (Ballesta et al., 2019; Hess et al., 2017; Won et al., 2020).

Another manner of improving PA is to add QTLs as cofactors in the prediction model, so that allelic effects associated with these QTLs could have a greater variance (Fodor et al., 2014; Nsibi et al., 2020). This resulted in these two studies in an increase in PA up to 0.3 for traits under simple genetic architecture. However, this implies to detect QTLs within TS and that these QTLs also segregate in VS. In fact, I tested to add QTLs as cofactors in the prediction model in data simulated in chapter 1, but it did not increase PA compared to other marker selection methods as EN or LASSO (results not shown). Likewise, Campbell et al. (2021) proposed to compute trait-specific genomic relationship matrix (TGRM), by adding weights computed from loci effects on the phenotype, based on prior high-throughput phenotyping, and thus increased inter-population PA by about 0.05.

Using whole-sequence data, MacLeod et al. (2016) proposed to *a priori* split SNP markers into three categories depending on their potential impact on phenotype, based on prior biological information, such as candidate genes or missense variants for example. The resulting BayesRC method fits a GP model with varying genetic variance attributed to each marker category.

Phenomic matrix

In PP, the genotyping matrix is replaced by a phenomic matrix, with reflectance measured across wavelengths on a given tissue for all genotypes. As this matrix is numeric, I used the same statistical methods as with GP. In terms of matrix structure, the phenomic matrix displays high auto-correlation between close wavelengths and even at a longer range. To some extent, this is similar to the LD observed between loci. For the phenomic relationship matrix, I observed a lower structure compared to the GRM (data not shown), and this was confirmed by Galán et al. (2021). Besides, Galán et al. (2020) increased PA by selecting wavelengths with LASSO, this raises the question of how many wavelengths are necessary to capture all useful information in spectra.

One advantage of using NIRS is its lower dimension compared to SNPs; for example, in chapter 4 I used 2,101 wavelengths for PP, compared to 32,894 SNPs for GP. For hyperspectral imaging, one spectrum is available for each pixel and the number of pixels depends on the resolution. Studies that applied PP averaged spectra over pixels, in order to reduce the dimension to one spectrum per plot (Galán et al., 2020; Krause et al., 2019).

Use of other omics?

The association of genotype to phenotype remains challenging, because of GxE and due to complex genomic interactions. To address these issues, several studies proposed to use downstream omics such as transcriptomics or metabolomics (Guo et al., 2016; Schrag et al., 2018; Wade et al., 2021; Westhues et al., 2017; Xu et al., 2016). They included transcriptomic and metabolomic data alone or in combination with genomic data. For combining several types of omics, some studies used a multi-matrix model, as I did in chapter 4, while others concatenated predictor matrices. Wade et al. (2021) established that information redundancy between SNP markers and transcriptomics was a key factor explaining the usefulness of genomics and transcriptomics combination. Similarly in chapter 4, redundancy between SNP markers and NIRS matrix could explain why I did not find any superiority of adding NIRS into the GP model. Transcriptomic data for a few genes are available in SxG progeny and could be tested in the future (Huang et al., 2013; Huang et al., 2014).

With the drop of sequencing costs, it is conceivable that genome sequence could be directly used in GP model instead of array data for example. In grapevine, a set of 3,000 varieties is currently being sequenced at low resolution in the frame of an international research consortium. This raises the question of the effective usefulness of whole genome sequencing (short reads) vs genotyping in GP, despite the proven theoretical interest due to adding more causal polymorphisms for inter-population prediction (MacLeod et al., 2016). Using simulated data, Pérez-Enciso (2014), Pérez-Enciso et al. (2017) found no advantage of using sequence over array genotypic data, which might be due to the lower quality of markers from sequencing than from genotyping. Still, genome sequencing is useful for marker imputation for increasing SNP density (Pégard et al., 2019), which is expected to increase PA when using the diversity panel as TS.

Thanks to long reads sequencing methods, other areas of research become possible, as studies on the pangenome, which is the entire set of genes found for a species (Tettelin et al., 2005). The specific part of pangenome includes structural variants (SVs) such as copy number variation, transposable elements or presence-absence variation. Structural variants are suspected to play a role in phenotypic plasticity (see Wellenreuther et al. (2019) for a review), and in grapevine, SVs are associated with domestication traits, such as sex determination (Zhou et al., 2019). Whole genome sequencing is a promising prospect to account for SV(s), with adapted prediction models (Kosugi et al., 2019).

5.3.3 Statistical estimation of variable effects

I mostly used frequentist methods, because they were fast, easy to implement and to optimize. In chapter 2, I compared several univariate GP methods and some of their multivariate counterparts, showing no substantial advantage of the latter ones with our population size. I also re-analyzed simulated data from Jia and Jannink (2012) and found that multivariate BayesA provided higher PA than multivariate LASSO. In chapter 4, I implemented two GP methods and kept PA from the best one in order to get rid of the variation due to the method. Retrospectively, I should also have used the GBLUP model for the half-diallel population in chapter 3, because it resulted in the highest PA within this population in chapter 4. However, differences between RR-BLUP and GBLUP were small in the diversity panel, and in scenario 2, GP model was trained in this population. Therefore, we expect close results for this scenario, when switching from RR-BLUP to GBLUP.

I showed that GP method performance depends on (i) the genetic architecture (chapters 2 and 3), (ii) the prediction scenario (chapter 3), (iii) the input matrix structure (chapter 4). All these parameters are in interaction with each other. Therefore, there is no best GP method that could be used universally, which is consistent with Azodi et al. (2019).

These findings suggest that GP methods generally have small impact on PA and that optimizing input data and phenotyping would have greater impact on PA.

5.4 Accounting for GxE

In my thesis, I did not consider GxE interactions, because I had data only for a few years and only from one site for most populations. However, GxE is an issue in grapevine breeding because it is admitted in France, with the PDO system, that each variety should be grown in specific environments. An alternative to multi-site trials for studying GxE is to apply abiotic stress on one replicate of the population at a single location. This was done in chapter 2 with water stress applied both in a high-throughput phenotypic platform and in the field.

5.4.1 Impact of site and year on phenotypic variation

In grapevine, several studies evidenced the effects of GxE and GxY on phenotypic data (Gonçalves et al., 2016; Gonçalves et al., 2020; Migicovsky et al., 2021; Rasoli et al., 2015; Rustioni et al., 2019). In the statistical analysis of phenotypic data in these studies, GxE was most of the time significant and could affect yield from –8.2 to +33.5% for example.

In the sense of variance decomposition, environmental effects encompass location characteristics such as soil composition, precipitation and temperature across seasons. Thus, accounting for GxE in prediction models requires prior climatic recordings of environments for the training and prediction sets. However, local environmental characteristics may vary across years, at least for temperature and precipitations. These yearly variations cannot be predicted in advance. Then, the breeder needs to account for genetic response to macro-environment,

with average climate characteristics, while mitigating genetic response to year-to-year variations (Gonçalves et al., 2020).

From our data in a single site, I have derived multi-year BLUPs. Since only a few years are involved, it was not possible to estimate climatic components. However, for traits with significant genotype-by-year interactions (GxY), one could train a GP or PP model with phenotypes of a given year and predict phenotypes of another year, to test the robustness of prediction across years.

5.4.2 Traits related to drought tolerance

Drought is a highly complex stress, thus grapevine response to drought is also highly complex and involves several polygenic traits (Coupel-Ledru et al., 2016; Coupel-Ledru et al., 2014; Marguerit et al., 2012; Trenti et al., 2021), as I confirmed in Article I. Depending on the environmental scenario considered, i.e., extreme, severe or moderate drought, grapevine genetic response will vary (Tardieu, 2012; Tardieu et al., 2018). In chapter 2, I used phenotypic data measured in a phenotypic platform under semi-controlled conditions. However, results from this experiment are likely to diverge from experimental measurements of response to drought in the field, as environmental variables are different, as well as the managing system. Moreover, using a high-throughput platform severely limits the number of varieties that can be phenotyped at a time. It would thus be valuable to have a proxy of these traits that can be measured in the field. Proxys of physiological traits can be measured with spectra and is already routinely used for the biochemical characterization of plant tissues (Grzybowski et al., 2021; Osborne, 2006). For most of traits, I found a low broad-sense heritability for traits related to response to drought in part 2.4, thus suggesting low GxE for these traits, except for $\delta 13C$.

5.4.3 Statistical models for handling GxE

One way of studying GxE is to measure the same trait under different environments. In this case, we expect some genetic correlation between these phenotypes, that could be handled by using multivariate models. In chapter 2, I tested multivariate methods on several correlated traits related to water deficit. Concerning the computational burden and the limited number of traits that can be handled together, I decided to focus on univariate GP methods for the next chapters since they are more transferable to breeding.

I also could have tested trait-assisted prediction, a scenario in which a multivariate GP model is fitted with VS partly phenotyped for secondary traits in another environment (Fernandes et al., 2018; Jia & Jannink, 2012; Nsibi et al., 2020). This option could help reducing phenotyping costs, but genetic correlation must be high between target and secondary traits. In grapevine, such a configuration would be valuable with berry composition traits, by using NIRS wavelengths measured on clusters. Some multivariate methods such as in Runcie et al. (2021) are specifically adapted for handling both multi-environment trials (METs) with

incomplete design and secondary traits. However, one must be careful when comparing univariate and such multivariate GP models using secondary traits, because most of the time, secondary traits are measured on VS individuals. Thus, these individuals share the same non-genetic variance, which leads to an over- or under-estimation of PA, and may lead to a wrong choice of GP model, as demonstrated in Runcie and Cheng (2019). This pitfall could be avoided by correcting the estimation of accuracy, or by using clones of VS genotypes to measure secondary traits.

Another approach is to combine ecophysiological mechanistic modeling and QTL detection or GP to find alleles associated with a response parameter to environmental factors (Marguerit et al., 2012; Vivin et al., 2017). Leveraging such biological information that includes response to the environment from crop growth model or from phenological stages allows to increase PA when predicting across-environments (Messina et al., 2018; Millet et al., 2019). Burgueño et al. (2012) conceived a method to model similarities between environments through a covariance structure. Models such as AMMI (additive main effect and multiplicative interaction) or GGE (genotype main effect plus genotype-by-environment interaction) use phenotypes from METs to fit G and GxE based on principal component decomposition (Yan et al., 2007).

5.4.4 The use of PP for specifically handling GxE

As spectra reflect the biochemical composition of a tissue, it is expected that spectra variability partly reflects genotypic and GxE variances. But many questions remain unanswered concerning GxE in spectra. What is the extent of GxE captured by spectra? I had spectra from one location over two years. Thus, I could only assess the GxY variance, which was moderate in the mixed model. In Rincent et al. (2018), several environments were available and the proportion of GxE depended on the tissue analyzed, with large variability for leaves (wheat) and wood (poplar) and a more stable proportion of GxE variance for grain (wheat). Overall, the magnitude of GxE in this study was higher than the proportion of variance I found for GxY in this work. In chapter 4, I fitted a mixed model for each wavelength reflectance and further retained only BLUPs for genotype and cross or subpopulation effects, thus excluding GxY effect. If we had phenotypic data measured during the same years as spectra, we could test the possibility to use specific GxY BLUPs to predict phenotypes from the same year. Likewise, if we had different locations with phenotypes and spectra, we could test whether using genotype-by-location interaction in spectra results in higher PA for the environment considered.

In Montesinos-López et al. (2017a), authors reported that models integrating GxE resulted in similar PA than models without GxE. However, integrating wavelength-by-environment interaction increased PA.

5.5 Practical guidelines for present and future grapevine breeding programs

In this section, I will first provide guidelines for the Inter-Rhône breeding program which is in progress, then I will advise on how future grapevine breeding programs could be conducted, according to my results.

5.5.1 Present breeding programs: Inter-Rhône example

Program characteristics

The aim of this program is to conceive new cultivars for *Vallée du Rhône*, by crossing emblematic varieties from this region with polygenic resistant cultivars.

For that purpose, new genotypes were produced in 2016-2017-2018 by crossing Syrah and Grenache with 4 genotypes bearing two resistance genes to PM and DM. The 155 remaining genotypes after MAS for major genes were planted in 2018-2019-2020 at *domaine de Piolenc* (Orange, South of France). Phenotypes will be available over the next few years and it is planned to genotype these individuals with GBS.

Genotypic value prediction

If we retain the GP option (the alternative PP option being described in 5.5.2), this progeny should be genotyped using the same enzyme for GBS as in the three populations studied in the thesis. Then, we could use phenotypic data from the half-diallel population with crosses related to Syrah or Grenache, depending on the target population, to train the GP model, thus approaching what I coined scenario 1b. In chapter 3, this scenario gave an average PA of 0.39 across traits, with PA superior to 0.4 for acid traits at ripe stage except for malic acid. However, in scenario 1b, three crosses were used to train GP model for each parent, VS being the fourth cross with one common parent, while for Inter-Rhône progenies, we can directly use the four crosses with Syrah and the four crosses with Grenache. Thus, compared to scenario 1b, one more cross will be used to train the GP model. On the one hand, we expect that adding one more cross compared to chapter 3 will increase PA, but on the other hand, the other parents of Inter-Rhône crosses are inter-specific resistant hybrids that are less related to the TS compared to other half-diallel parents, a configuration that will likely decrease PA.

Predicting genotypic values can only be done for traits that have been measured in the TS. My results can thus be integrated into this breeding program in four ways. (i) If traits are already available in the half-diallel, such as yield, berry composition (sugar/acid balance) or phenology, they can be used directly for prediction. (ii) If some traits are missing in the half-diallel, they must be phenotyped in the following years in the adequate populations of the half-dialell. (iii) If these traits are available in the diversity panel or SxG populations, they could be used for prediction, but with expected loss in PA. (iv) If they are not, the phenotyping of missing traits and related prediction can be done after a first selection of individuals

for already available traits. In the latter case, enough genotypes must be retained, for keeping some genetic variability for yet unphenotyped traits.

After having predicted genotypic values for all candidates, the final choice between individuals remains to be done. The selection index theory proposed to put weights on traits depending on their relative importance (more details about selection index are given in part 5.5.3).

5.5.2 Future breeding programs

I established guidelines and advices for future grapevine breeding programs. A summary of breeding steps is in Figure 5.2.

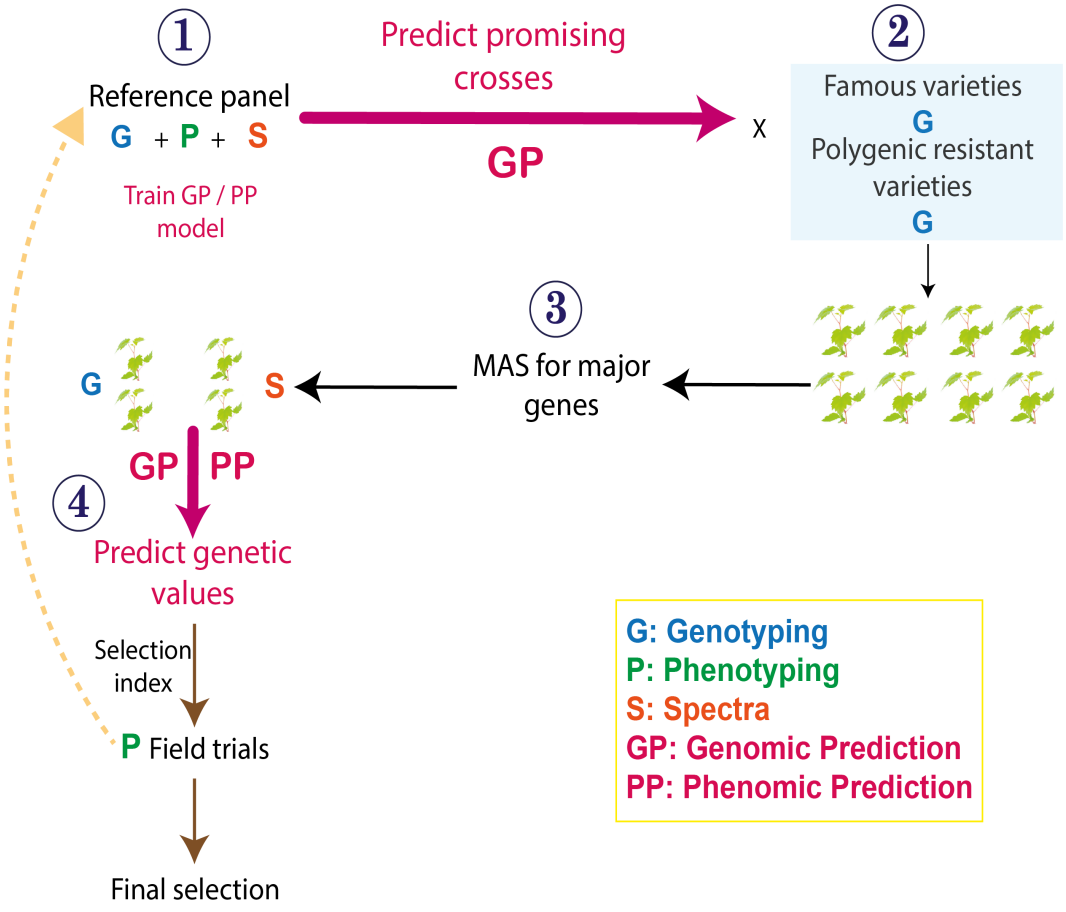


FIGURE 5.2 – Future grapevine breeding schemes involving GP and/or PP.

Prediction of future crosses with genotyping

In future breeding programs, crosses to be made are not yet determined. Since we are still seeking for resistant genotypes, it is likely that future crosses will include as parents genotypes resistant to PM, DM and possibly other diseases on the one hand and genotypes known for their wine quality on the other hand.

The choice of crosses to be done could be determined by GP, using RR method. I showed in chapter 3 that the average of parent's observed genotypic values was different from the observed average offspring value. Conversely, cross mean GP was very accurate, especially in scenario 1b. This means that cross mean prediction implementation is a crucial step for improving grapevine breeding scheme. To do so, a reference population and candidate parents have to be genotyped with the same SNPs and the reference population must be phenotyped for the target traits. Then, GP model is trained within the reference population and cross mean predicted based on the average parental genotype, defined for each cross as in chapter 3. When I used the diversity panel as reference population, PA was lower than in scenario 1b but still satisfying for breeding. In particular, I found that per cross PA of cross mean was highly correlated with the genetic distance between parents (on a PCA projection), crosses with the closest parents yielding the most accurate predictions of cross mean, and that TS optimization improved PA for the least accurately predicted crosses. If these results were to be confirmed with more crosses and more than 15 traits, this implies that predicting cross mean should be applied to crosses with fairly related parents, while for crosses between distant parents, TS should be optimized. Other crosses could be done for genetically distant parents to avoid the risk of missing a good cross due to low prediction accuracy.

Prediction of cross variance is also possible, despite a lower accuracy for scenario 1a (see chapter 3), which could help selecting crosses altogether with cross mean and deciding on the number of offspring to generate for each cross.

After the achievement of the most promising crosses, MAS should be applied for traits controlled by major genes, to retain only seedlings with polygenic resistance for example. Then, two options are possible for further selecting individuals within crosses, as described in Figure 5.2. The choice between PP and GP will depend on additional results which will confirm or refute the interest of PP.

Option 1: applying genomic prediction

Based on the results from chapter 2, I found that within-full-sibs GP gave the highest PA, but it is not the most desirable configuration for breeding, as it requires to wait for phenotyping part of the population to be predicted. Still, this configuration is currently applied in other grapevine breeding programs (called Martell and Edgarr) for Cognac and Rosé wines, respectively. Scenario 1b from chapter 3 resulted in only a small decrease in PA compared to scenario 1a but it requires phenotyping several half-sib families related to the cross to be predicted. Therefore, if the selected crosses involve parents of the half-diallel and the target traits have been phenotyped in the half-diallel, one could use half-sib families from this population to train the GP model, as described for Inter-Rhône program. Otherwise, I advise to use the SelGenVit panel, which has been specifically conceived for GP in future breeding programs (see 5.1.3). This panel is being planted in three locations in 2021 and 2022 and first phenotypic data will thus be available starting from 2024. The advantage of using a single

reference population is that phenotyping effort is concentrated on one population. Hence, it may be worth to measure costly traits in this population.

In chapter 3, I found that RR method more often resulted in higher PA than LASSO for across-population prediction. Therefore, I advise to use RR rather than LASSO for GP. In chapter 4, I found that GBLUP method was better than RR-BLUP method for GP within the half-diallel and hypothesized that the difference was due to structure within the genomic matrix. Within the diversity panel, both methods yielded similar PA. Within the SelGenVit panel, we expect that GBLUP will also be better than RR-BLUP, due to potentially more structure than in the diversity panel. I would thus recommend to use GBLUP rather than RR-BLUP. However, the relative performance of these methods is susceptible to change from within- to across-population GP, but this will be easy to test with the data available on the diversity panel.

In a nutshell, GP model using GBLUP could be trained with phenotypic and genotypic data from the reference panel (half-diallel or SelGenVit panel) to predict genotypic values of offspring within crosses.

Option 2: applying phenomic prediction

In chapter 4, I found that PP provided PA values close to those given by GP, with an average decrease of –0.2. Nevertheless, PP could still result in more genetic gain than GP because NIRS is much cheaper and high-throughput than genotyping (see Rincent et al. (2018) for an estimation of the potential comparative gain). Indeed, switching from GP to PP would allow to save money, that can be reinjected into more individuals to be phenotyped and/or more phenotyping precision. These variables are actually present in the breeder's equation and increasing their value would provide larger genetic gain.

However, there is an overall lack of knowledge for applying PP in grapevine breeding programs. In particular, I have not tested across-population PP. If we apply within-population PP, the population to be predicted must be partly phenotyped in order to train the PP model and predict the genotypic values of selection candidates. Such a configuration would require to wait for phenotypic data, which will drastically reduce the genetic gain of PP option. Therefore, I advise to test across-population PP before implementing this option into grapevine breeding programs. This will be quick since data are available. Actually, Galán et al. (2021) provided encouraging results in rye by using more or less related genotypes between TS and VS, and found that across-population decrease in PA was lower for PP than for GP.

Conception of breeding cycles

After having performed the final selection of candidates to be released, these individuals should be integrated into the reference panel, and another breeding cycle should be performed the same way. If GP option is selected, released candidates must be genotyped, and

phenotypic data are already available, coming from field trials. These released candidates should also be added to the pool of potential parents. This breeding cycle is expected to last nine years, which implies that every two cycles the reference panel should be replanted, incorporating by the way the released candidates.

One may also use pre-breeding steps to introduce new alleles for disease resistance. Sources of resistance against Pierce's disease or black-rot are already characterized and should be introgressed into breeding material (Rex et al., 2014; Riaz et al., 2009). Using microvine (Chaïb et al., 2010) would be highly beneficial for this purpose, as its generation time is reduced from 3 to 1 years.

Switching to GP or PP instead of pedigree and phenotyping selection will induce a quicker loss of genetic diversity and increase in homozygosity (Colleau et al., 2017; da Silva et al., 2021). Furthermore, switching from within-family selection alone to combined across- and within-family selection will reduce genetic diversity (da Silva et al., 2021). As we know that grapevine suffers from strong inbreeding depression and that a decrease in genetic variance will decrease genetic gain, it is important to tackle this issue. One possibility is to include a constraint for maintaining genetic diversity when choosing crosses to be done (Allier et al., 2019; Fernández et al., 2021).

5.5.3 Defining grapevine ideotype and associated selection index

In my thesis, I developed tools for accelerating the selection of new grapevine varieties, based on already available phenotypic data. For the full implementation of my work, grapevine ideotype should be clearly identified by wine-growers and wine-growing corporations so that selected genotypes meet their expectations. Such an ideotype was conceived by Viviane Bécart at Inter-Rhône. This ideotype comprises nine traits under four categories. Traits were late harvest date, wilting, defoliation, index of growth stop, yield and yield stability, high acidity in must, anthocyanin content and wine appreciation. This work should be continued in partnership with ecophysiologists and enologists for grapevine response to drought and for wine quality, respectively. Once traits for ideotype have been defined, phenotyping reference panel populations must be a priority.

The ideotype is not an addition of traits, we must consider genetic correlation between these traits. For example, within the half-diallel population, we found negative correlations between tartaric acid and harvest date, or between harvest date and vigor (data not shown). This was also confirmed with another study on a bi-parental population, that total acidity was negatively correlated with date of ripening (Bayo-Canha et al., 2012). Restricted selection indices have been developed to handle such cases and allow to select for a target trait, while holding other traits at their initial value (Kempthorne & Nordskog, 1959). Other indices are useful to impose more complex constraints on ideal trait value (Itoh & Yamada, 1988; Lin, 2005).

Beyond selection index, selection intensity must also be determined as a function of accuracy for reaching a desired genetic gain from the breeder's equation.

5.6 Further research prospects

In this section, I will propose some research avenues worth exploring to further optimize genotypic value prediction in grapevine, in addition to more specific prospects already included throughout general discussion.

5.6.1 Incorporating non-additive genetic effects

Concerning dominance effects, most studies focused on species commercialized as F1 hybrids, for which dominance is expected due to hybrid vigor. In heterozygous species such as grapevine, dominance is inevitable but its extent remains to be studied, depending on the trait. In the half-diallel, I observed transgressive segregation patterns for many traits, thus suggesting non-additive genetic effects.

Incorporating additive and dominance effects into the prediction model does not always lead to an increase in PA, as shown in Chapter 2 and discussed in 5.3.2. For example, Yamamoto et al. (2021) in strawberry reported transgressive segregation but additive and dominant GP models did not increase PA. Similarly, Resende et al. (2017) in Eucalyptus found a superiority of the additive plus dominance model only for one trait out of the three studied, despite a documented dominance effect. It is likely that the superiority of such models would also depend on statistical modelling, e.g., using a dominance genomic relationship matrix or a design matrix for coding additive and dominance effects as in 2.2.2 .

In the long term, using GP models that only include additive effects will increase allelic frequency of alleles with positive additive substitution effects, thus leading to a decrease in heterozygosity (Werner et al., 2020). Hence, even if GP models including only additive or all genetic effects provide identical PAs, it would be beneficial in the long term to include non-additive effects into the GP model.

Testing models including non-additive effects using an adapted genomic relationship matrix (Vitezica et al., 2013; Vitezica et al., 2017), could be quite straightforward as a prospect of my work, as data are already available.

5.6.2 Designing trials for studying GxE

The diversity and SelgenVit panels are currently planted in different locations. Once phenotypic data will be available, the magnitude of GxE could be measured and specific models integrating GxE could be tested.

Other research groups have created, phenotyped and genotyped grapevine populations for QTL detection (Vezzulli et al., 2019). As the availability of data after publication becomes

more and more the norm, one could use these data and integrate them into GP models. However, most of the time, the number of genotypes in common over different locations is limited, making the use of this data statistically challenging. On the other hand, several grapevine varieties (not from created populations) are planted and studied in many places around the world and could be used to broadly estimate GxE on a restricted set of genotypes, extending published results such as those of Rustioni et al. (2019).

Across-environment prediction could be classified into three scenarios. We could have phenotypic data for a population in a given environment and predict phenotypic data for the same genotypes in a new environment. Another possibility is to predict unknown genotypes in unknown environments. Finally, an intermediate scenario would be to have an incomplete design with genotypes measured in one or two different environments and predict missing combinations. In the first scenario, ecophysiological and crop growth models will be more adapted than genetic models, because a combination of location and year can be considered as an environment. The second case is the most challenging one; to handle both new environments and new genotypes, one needs to combine GP with ecophysiological modelling, as has been done in Millet et al. (2019). Finally, the last scenario is the most classical one in breeding, and multivariate modelling could allow to handle such cases (Runcie et al., 2021).

5.6.3 Optimizing the use of spectra

Using spectra to replace genotyping for GP

Whalen et al. (2020) proposed a new original way of using spectra in breeding. As reflectance at a given wavelength is genetically determined, they hypothesized that reflectance at each wavelength could be modelled as a trait with a complex genetic architecture. Thus, each variation in allelic inheritance could be tagged by a QTL linked to reflectance variation. Therefore, these high-throughput phenotypes (HTP) could be used to impute segregation states of non-genotyped varieties in a progeny (parents need to be genotyped). Using simulated data, Whalen et al. (2020) showed that imputation accuracy was dependent on HTP heritability and genome size. Then, GP was implemented based on these imputed genotypic data. So far, this method is a proof-of-concept and has only been adapted to inbred species. Further work is needed before this could be applied to grapevine breeding.

Better understanding how PP works

We have seen in the previous paragraph that Whalen et al. (2020) hypothesized that reflectance at a given wavelength could be modelled as a trait under additive genetic architecture, with QTLs distributed along the entire genome. Is this approximation far from reality? Are there some genomic regions non "covered" by spectra? On the other hand, are there new regions such as structural variants leading to phenotype variants that are captured by spectra and not by traditional genotypic markers? Reflectance at a given wavelength is expected to be linked to several loci. Thus, reflectance variation summarizes allelic variation across several loci, which makes it nearly impossible to decipher the genetic determinism underlying a

trait from spectra. In chapter 4, I found correlation between PA of PP and GP, meaning that a trait well predicted under GP will also be well predicted under PP. However, this correlation deserves to be confirmed with more traits, which could be done using available data not included in this study (see 5.1.4). In particular, is PP working for simple genetic architecture with major QTLs? I expect that PP would be more adapted to infinitesimal genetic architecture. It is also anticipated that PP could better capture non-additive genetic effects, as it relies on expressed genome.

In order to shed light on some of these issues, we could use simulation by using existing spectra and simulating phenotypes from existing genotypes under various genetic architectures. Thus, we could check whether PP at least partly relies on the genetic relatedness captured by spectra, rather than on correlation between the trait and spectra or GxE.

Consequences of applying PP to breeding programs

Concerning the implementation of PP in breeding programs, some adaptations are needed for PP compared to using GP. First, in chapter 4 I used mature leaves and wood to measure spectra, but the advantage of GP is to save time. Thus, one needs to test PP with spectra measured on young leaves on seedlings, as is done for genotyping in GP, or on wood after only one year of growth.

In chapter 3, I used parental average genotype at each locus to predict cross mean, based on allelic effects previously estimated with GP. It is likely that using the same procedure with spectra will result in lower accuracy, because spectra are only partly related to SNP markers. But the extent of the decrease needs to be characterized. However, predicting cross mean using GP is not the costliest part, as only parents and reference panel need to be genotyped.

Unlike GP, PP does not primary relies on additive genetic effects. Then, it could more easily take into account non-additive effects compared to GP. This has an impact on genetic variance of selected genotypes and could be beneficial to genetic gain in the long term, as discussed above (see 5.6.1).

As pointed out in the previous section, across-population PP should be tested, to know how to practically implement PP in breeding programs.

Spectra modalities

So far, I only implemented PP using spectra from NIRS measured on dried samples. Many other sensors are used for measuring spectra, such as Red-green-blue (RGB) imaging, hyperspectral imaging which could be integrated in an unoccupied aerial vehicle (UAV), or laser-imaging detection and ranging which takes 3-dimensional images (Persa et al., 2021). All these spectra have their own characteristics in terms of portability, range of wavelength and dimensionality. So far, to the best of my knowledge, studies that implemented PP only used one reflectance spectrum per plant. When several pixels were available from hyperspectral imaging, spectra from all pixels were averaged at the plant level. Using multi-dimensional

data remains statistically challenging but handling all these pixels in PP might increase PA as more information is brought into the model. To do so, one could use multi-kernel models, or deep-learning methods such as convolutional neural network (CNN), a method specifically adapted for processing images (Jiang & Li, 2020).

Use of spectra to estimate phenomic relatedness

The aim of some grapevine breeding programs, such as the Inter-Rhône one, is to select varieties phenotypically close to emblematic varieties with polygenic disease resistance. To address that, we could use spectra that are reflecting endophenotypes, i.e., intermediate between genotypes and phenotypes. Within a bi-parental population, all full-sibs get the same proportion of each parental genome, yet some individuals may share more phenotypic similarities with one parent. This similarity could be better captured by spectra, by estimating a "phenomic relatedness". This method is currently being tested on populations from local French breeding programs.

5.7 Conclusion

I have investigated the interest of genomic and phenomic prediction on grapevine breeding, by testing several parameters affecting PA, such as training population, genetic architecture of traits or statistical method. In a nutshell, it emerges that GP and PP could be applied in breeding programs, although PP requires more studies to be fully optimized. In order to enhance PA, one could incorporate more biological or environmental information in the predictive model, such as endophenotypes or environmental covariates.

Chapter 6

Résumé en français

6.1 Introduction

6.1.1 A propos de la vigne

La vigne est cultivée sur 7,4 millions d'hectares dans le monde. La France arrive en cinquième position en termes de surface cultivée avec 789 000 hectares représentant 3% de la surface agricole cultivée (OIV, 2017). Néanmoins, la viticulture est un secteur économiquement majeur en France qui représente 31,2% des exportations agricoles (FRANCEAGRIMER, 2019).

La vigne cultivée correspond à l'espèce *V. vinifera* du genre *Vitis* qui contient environ 80 espèces inter-fertiles. La vigne a subi une première domestication il y a environ 9 000 ans dans le Caucase, puis une domestication secondaire, avec une spécialisation selon l'usage du raisin (consommation fraîche, raisins secs ou transformation en vin). Cette domestication secondaire est à l'origine de la (faible) structuration génétique en trois groupes observée aujourd'hui chez la vigne cultivée : cuve est, cuve ouest et table est (BACILIERI et al., 2013 ; PÉROS et al., 2011).

La vigne cultivée est une espèce diploïde avec une grande diversité génétique et un haut niveau d'hétérozygotie (LAUCOU et al., 2011). Elle souffre de dépression de consanguinité, ce qui rend difficile l'obtention de descendants issus d'autofécondation d'une variété.

Avec la possibilité de multiplication végétative, certaines variétés existent depuis des centaines d'années, c'est le cas du Pinot Noir, du Chardonnay ou encore du Gamay qui existent depuis le Moyen-Age, voire avant (BOWERS et al., 1999 ; LACOMBE et al., 2013).

A la fin du 19ème siècle, le mildiou, l'oïdium et le phylloxéra, ont été introduits en Europe. Quasiment toutes les variétés de vigne cultivées sont sensibles à ces ravageurs. Des traitements à base de soufre et de cuivre ont été développés pour lutter contre le mildiou et l'oïdium mais aucun traitement efficace n'existe contre le phylloxéra. Pour lutter contre le phylloxéra, deux solutions sont possibles : d'une part l'hybridation avec des sources de résistance génétique issues d'autres espèces du genre *Vitis* et d'autre part l'utilisation de variétés de porte-greffe résistantes, issues d'autres espèces de *Vitis*. La première option a conduit à la création des hybrides producteurs directs et leur abandon quasi complet pendant la seconde

moitié du XXème siècle, alors que la seconde option est toujours en vigueur aujourd'hui dans une très grande majorité de vignobles de par le monde.

Aujourd'hui, la vaste majorité des variétés cultivées sont toujours sensibles au mildiou et à l'oïdium et nécessitent donc de nombreux traitements phytosanitaires. Pour réduire l'utilisation de produits phytosanitaires, la sélection de variétés résistantes est une voie prometteuse. Dans ce contexte, le programme d'amélioration génétique français INRA-ResDur vise à combiner plusieurs sources de résistance au mildiou et à l'oïdium pour plus de durabilité (SCHNEIDER et al., 2019a). Néanmoins, les étapes de sélection pour les autres caractères impliquent toujours le phénotypage des descendants issus de ces croisements, ce qui prend six à sept ans pour la vigne. La durée totale du cycle de sélection est actuellement de 15 ans.

Un idéotype correspond à une variété idéale, c'est-à-dire, à la variété qui remplit les objectifs de sélection pour tous les caractères.

Comme nous l'avons vu, la résistance au mildiou et à l'oïdium est un critère majeur chez la vigne pour diminuer l'utilisation de produits phytosanitaires. Toutefois, la vigne est également sensible à beaucoup d'autres maladies comme les maladies du bois ou encore les insectes qui attaquent la vigne et transmettent des virus.

L'autre menace qui pèse sur la vigne est le changement climatique. Même si la vigne est notamment cultivée dans des régions au climat relativement chaud comme le Sud du pourtour méditerranéen, les projections climatiques prévoient une augmentation des sécheresses, l'augmentation globale des températures et une plus grande probabilité des événements climatiques extrêmes (IPCC et al., 2021). Les conséquences du changement climatique sur la vigne sont une augmentation du stress hydrique, l'avancée des stades phénologiques, et plus particulièrement de la date de récolte, et le changement de la composition des baies avec notamment une augmentation du degré d'alcool et une baisse de l'acidité. Toutefois, les cépages cultivés actuellement ne représentent qu'une faible partie de la diversité génétique existante chez *V. vinifera* et la sélection génétique devrait pouvoir permettre d'atténuer les effets du changement climatique (MORALES-CASTILLA et al., 2020 ; MYLES, 2013).

6.1.2 Prédiction de la valeur génétique

Avant la prédiction génomique, la variation phénotypique était caractérisée au niveau moléculaire à l'aide de QTL (*Quantitative Trait Loci*), c'est-à-dire des zones du génome statistiquement associées à la variation d'un caractère. Une fois les QTL identifiés, le caractère d'intérêt pouvait être prédit à l'aide de marqueurs moléculaires associés aux QTLs via la sélection assistée par marqueurs (SAM). Cette méthode est par exemple utilisée actuellement pour suivre les gènes de résistance au mildiou et à l'oïdium chez la vigne. Néanmoins, cette méthode devient rapidement inapplicable lorsque le nombre de QTL à suivre augmente. De plus, la majorité des caractères quantitatifs sont contrôlés par un grand nombre de QTL avec chacun de petits effets, ce qui rend l'application de cette méthodologie impossible (certains QTL ne sont pas détectés et il y a trop de QTL à suivre).

La sélection génomique (SG), permise par l'accès au génotypage haut-débit, a été proposée par MEUWISSEN et al. (2001) pour répondre à ces limites. Le principe est d'utiliser une population d'entraînement phénotypée et génotypée pour estimer les effets statistiques associés aux marqueurs moléculaires et d'utiliser ces estimations pour prédire les valeurs génotypiques de la population de sélection sur la base du génotypage uniquement.

Plusieurs méthodes statistiques existent pour l'estimation des effets des marqueurs, les plus populaires sont le RR-BLUP ou GBLUP, les méthodes à inférence bayésienne comme le BayesB, Bayes Cpi, BayesR, ou encore les méthodes non-linéaires comme le random forest, ou le RKHS (de los CAMPOS et al., 2013). Les premières méthodes permettent uniquement d'estimer des effets alléliques additifs de substitution, donc une reformulation du modèle est nécessaire pour estimer la dominance ou l'épistasie. Les méthodes non-linéaires sont plus à même de capturer des effets non-linéaires (et donc non-additifs) mais ils demandent plus de temps de calcul et des jeux de données plus importants. Selon l'architecture génétique des caractères, le classement entre les méthodes peut varier mais les performances globales des différentes méthodes sont proches (AZODI et al., 2019).

Les principales variables qui affectent la précision de prédiction génomique sont : l'apparentement entre les populations d'entraînement et de validation, le type de marqueurs moléculaires associé à la densité de génotypage, et le phénotypage associé à la structure de la population. Ces paramètres interagissent entre eux et avec la méthode utilisée pour la SG.

Globalement, la SG permet d'accélérer le cycle de sélection, avec une prédiction des valeurs génétiques à un stade précoce. Ce point est particulièrement important chez la vigne où le temps de génération est très long et les essais phénotypiques coûteux. De plus, la SG permet d'augmenter la précision de sélection et de gérer la diversité génétique sur le long terme (CONSORTIUM et al., 2021).

Malgré des avancées considérables, le prix du génotypage peut constituer un frein à l'application de la SG dans les programmes d'amélioration, où des milliers d'individus doivent être génotypés. C'est pourquoi, RINCENT et al. (2018) ont proposé la sélection phénomique (SP), comme alternative à la SG. Le principe repose sur l'hypothèse que les spectres mesurés sur les tissus des plantes dépendent de leur composition en différentes molécules, qui est déterminée génétiquement. Les spectres reflètent donc le génome des individus, en interaction avec l'environnement. Cette méthode a été testée avec succès sur le blé et le peuplier et quelques autres espèces mais reste nouvelle et de nombreuses questions demeurent.

6.1.3 Problématique de la thèse

Cette thèse résulte d'un partenariat entre l'interprofession viticole Inter-Rhône, l'IFV (Institut Français de la Vigne et du Vin) et INRAE. Elle s'est déroulée au sein de l'UMT GénoVigne et de l'équipe DAAV de l'UMR AGAP Institut. L'objectif était de répondre aux problématiques d'Inter-Rhône en termes de création variétale pour accélérer la sélection de nouvelles variétés de vigne résistantes au mildiou et à l'oïdium.

Si la SG a démontré son utilité dans les programmes d'amélioration des animaux et de diverses espèces végétales (CONSORTIUM et al., 2021), son application chez la vigne reste limitée à quelques études préliminaires restreintes. L'objectif de ma thèse était donc d'implémenter la SG et la SP dans des configurations variées et sur un grand nombre de caractères pour étudier leurs intérêts respectifs en vue d'optimiser la sélection de nouvelles variétés de vigne. Pour cela, j'ai utilisé trois populations de vigne déjà génotypées et phénotypées.

Dans un premier temps, j'ai comparé plusieurs méthodes de prédiction génomique dans une population bi-parentale, en testant notamment si des méthodes multivariées (multi-caractères) pouvaient avoir un intérêt (Chapitre 2) . Dans un second temps, j'ai appliqué la SG dans un contexte plus proche de celui des programmes de sélection, c'est-à-dire en inter-population, avec une population d'entraînement plus ou moins apparentée avec la population de validation (Chapitre 3). Enfin, j'ai testé l'implémentation de la SP, en comparant sa précision à celle de la SG (Chapitre 4). Enfin, une discussion générale a permis de discuter certains points de mes travaux, de proposer quelques pistes d'application en sélection, et des perspectives de recherche.

6.2 Prédiction génomique univariée et multivariée appliquées à une population bi-parentale

La modélisation statistique multivariée permet d'analyser conjointement plusieurs caractères. Jusqu'à présent, il a été montré que cette modélisation permettait d'obtenir une meilleure précision de prédiction des phénotypes en prenant en compte les corrélations génétiques et non génétiques entre les caractères e.g., JIA et JANNINK (2012). Par ailleurs, certaines méthodes de prédiction génomique font de la sélection de variables (effets alléliques) ; on peut donc s'en servir pour l'étude du déterminisme génétique des caractères (détection de QTL). J'ai comparé plusieurs méthodes de SG (ridge regression, elastic net, LASSO, gradient boosting, et régression pénalisée structurée) avec des méthodes classiques de détection de QTL par cartographie d'intervalle (SIM pour *Simple Interval Mapping* et MIM pour *Multiple Interval Mapping*).

Dans ce contexte, j'ai utilisé une population bi-parentale de 188 individus, issue d'un croisement réciproque entre les variétés Syrah et Grenache. Cette population a été phénotypée en conditions semi-contrôlées (plateforme PhenoArch) pour 7 caractères liés à la tolérance au déficit hydrique (la transpiration diurne et nocturne spécifique, le potentiel hydrique foliaire, la différence de potentiel hydrique entre sol et feuille, la conductance hydraulique, la différence de production de biomasse et l'efficience d'utilisation de l'eau) (COUPEL-LEDRU et al., 2016 ; COUPEL-LEDRU et al., 2014). Cette même population a été plantée au champ, un stress hydrique a été appliqué sur un des deux blocs, et 7 caractères agronomiques ou liés à la tolerance à la sécheresse ont été étudiés (le poids de bois de taille, le poids de baie, la date de véraison, la fertilité, le rendement, le nombre de grappes et le $\delta C13$ qui reflète le stress hydrique subit par la plante au cours de la saison).

Pour tester le potentiel intérêt de la modélisation multivariée, et des méthodes de SG pour la détection de QTL, j'ai construit une nouvelle carte génétique dense avec des marqueurs SNP (Single Nucleotide Polymorphism), et j'ai simulé deux caractères avec une corrélation génétique, en faisant varier l'architecture génétique et l'héritabilité. Par ailleurs, j'ai utilisé les données expérimentales obtenues en champ et en plateforme de phénotypage pour valider les résultats de prédiction et détecter de nouveaux QTL.

Sur les données simulées, les résultats de prédiction génomique ont montré que certaines méthodes étaient particulièrement sensibles à l'architecture génétique avec une baisse de précision lorsque l'architecture génétique était complexe (50 QTL), plus particulièrement pour SIM et MIM. D'une manière générale, une meilleure précision de prédiction était associée à une plus grande héritabilité. Pour les méthodes multivariées, les précisions de prédiction n'étaient pas supérieures aux meilleures méthodes univariées. Pour la détection de QTL, les méthodes de SG faisant de la sélection de variable surpassaient SIM et MIM en termes de précision de sélection lorsque 50 QTL étaient simulés. Toutefois, l'utilisation de méthodes de SG pour la détection de QTL requiert que la valeur du paramètre de pénalisation soit adaptée à la sélection de variable et non à la prédiction. Ceci a été rendu possible grâce à des extensions de ces méthodes comme la *Stability Selection* ou le *marginal False Discovery Rate*.

Sur les données expérimentales obtenues en plateforme de phénotypage, les méthodes de SG avaient les meilleurs résultats pour la prédiction, avec une très légère supériorité globale des méthodes multivariées. Les précisions de prédiction obtenues variaient de –0.1 à 0.68 selon les caractères et les méthodes, avec une moyenne de 0.38 pour la meilleure méthode (elastic net multivarié). Ces valeurs sont encourageantes pour l'application de la SG sur des caractères liés à la tolérance à la sécheresse. (COUPEL-LEDRU et al., 2016 ; COUPEL-LEDRU et al., 2014) avaient déjà appliqué une détection de QTL sur ces données. En comparaison, j'ai utilisé de nouvelles méthodes pour la détection de QTL, qui ont prouvé leur utilité sur des données simulées ; de plus, j'ai utilisé une carte génétique plus dense. Ces nouveautés m'ont permis de réduire les intervalles de confiance des QTL et de trouver de nouveaux QTL pour ces caractères.

Sur les données expérimentales obtenues au champ, j'ai appliqué les meilleures méthodes précédemment définies pour la SG. Le classement entre les méthodes était similaire à celui observé sur les données de la plateforme de phénotypage. Pour chaque caractère, j'ai calculé la différence normalisée entre modalités (présence vs absence de stress hydrique) pour chaque génotype, considérant cette différence comme un nouveau caractère reflétant la réponse au stress. La précision de prédiction de la réponse au stress hydrique était faible pour tous les caractères et pour toutes les méthodes testées.

6.3 Prédiction génomique inter-population

La SG est souvent appliquée au sein d'une population génétiquement homogène à la fois pour entraîner et valider le modèle. Cependant, cette configuration correspond rarement à

celle d'un programme d'amélioration, où la population de sélection correspond à des croisements qui ne sont pas encore réalisés et qui n'ont pas servi à entraîner le modèle. Dans cette partie, j'ai donc testé la SG dans des configurations plus réalistes en création variétale. Pour cela, j'ai testé l'utilisation d'un panel de diversité, composé de 279 variétés représentant l'ensemble de la diversité génétique de *Vitis vinifera*, comme potentielle population d'entraînement universelle. La population de validation était composée de 622 individus répartis en 10 populations bi-parentales équilibrées issues des croisements entre 5 variétés, chacune de ces 5 variétés étant impliquée dans 4 croisements (plan de croisement appelé demi-diallèle). Ces deux populations ont été phénotypées différentes années et génotypées avec des SNP, dont 32 894 sont en commun entre ces deux populations. Les quinze caractères communs étudiés sont classés en quatre catégories : caractères morphologiques pour le poids, la largeur, la longueur, la compacité de la grappe, le nombre de grappes et le poids des baies ; caractères de composition des baies pour le malate, le tartrate et le shikimate au stade mûr et des ratios associés ; caractères de phénologie avec la date de véraison (début de la maturation), la date de récolte et l'intervalle entre les deux, et le caractère de vigueur.

J'ai décomposé mon étude en deux sous-objectifs. Dans un premier temps, j'ai prédit la moyenne des descendants d'un croisement, ce qui correspondant à la première étape de sélection des parents pour réaliser les meilleurs croisements ; puis j'ai prédit les valeurs des individus au sein de chaque croisement, ce qui correspond à la seconde étape de sélection une fois les croisements réalisés. Pour chaque étape, j'ai mesuré la précision de prédiction avec deux méthodes de SG : ridge regression (RR) et LASSO. La performance relative de ces deux méthodes dépendant de l'architecture génétique du caractère, j'ai gardé la valeur de la prédiction de la meilleure méthode afin de s'abstraire de cet effet. J'ai conçu trois scénarios de prédiction : dans le scénario 1a, le modèle est entraîné et validé au sein du demi-diallèle par validation croisée ; dans le scénario 1b, trois populations issues de croisement avec un parent en commun sont utilisées pour entraîner le modèle, la prédiction se faisant dans le quatrième croisement du demi-diallèle impliquant ce même parent. Enfin, dans le scénario 2, le panel de diversité est utilisé pour entraîner le modèle et les croisements demi-diallèle servent à sa validation. Le scénario 1a correspond à la précision de prédiction maximale, le scénario 2 correspond à ce qui pourrait être implémenté dans un programme de sélection et le scénario 1b est un intermédiaire entre ces deux scénarios. Enfin, j'ai testé différents paramètres pouvant expliquer la précision de prédiction obtenue à chaque étape.

L'analyse des données phénotypiques montre que les populations du demi-diallèle affichent des moyennes et des variances différentes selon les caractères. Au niveau génétique, l'apparentement entre le demi-diallèle et le panel de diversité est partiel, avec 1/3 des individus du panel (de la sous-population cuve ouest) étant plus apparentés avec le demi-diallèle que les autres (des sous-populations cuve est et table est).

Pour la prédiction de la moyenne des descendants des croisements, la distribution des précisions de prédiction étaient diférentes selon que l'on considérait les moyennes par croisement tous caractères confondus ou les moyennes par caractère tous croisements confondus. Par

croisement, les précisions de prédiction étaient autour de 0.6 pour les scénarios 1a et 1b et beaucoup plus variables (de –0.3 à 0.72) pour le scénario 2, avec une forte corrélation négative avec la distance génétique entre les parents des croisements.

Pour la prédiction des performances des individus au sein de chaque croisement, le scenario 1a permettait d'obtenir la meilleure précision de prédiction, cette dernière diminuant avec le scénario 1b et le scénario 2, ce qui était attendu. Les précisions de prédiction demeuraient toutefois très variables selon le croisement et le caractère considérés. Pour les scénarios 1a et 1b, le paramètre le plus déterminant pour la précision de prédiction était l'héritabilité, alors que pour le scénario 2, les paramètres les plus déterminants étaient l'héritabilité, la différentiation entre les croisements pour chaque caractère, et le modèle de prédiction (RR ou LASSO).

Pour améliorer la précision de prédiction et étant donné le faible apparentement global entre le panel de diversité et les croisements du demi-diallèle, nous avons échantillonné dans le panel les individus les plus apparentés avec chaque croisement pour constituer une nouvelle population d'entraînement optimisée. Cette optimisation a permis d'améliorer la prédiction de la moyenne des croisements, lorsque qu'ils étaient mal prédits. Pour la prédiction de la valeur des individus au sein d'un croisement, l'optimisation de la population d'entraînement n'a permis qu'une légère amélioration de la précision de prédiction en comparaison avec un échantillonnage aléatoire.

Par ailleurs, nous avons également implémenté la prédiction de la variance des populations issues des croisements du demi-diallèle, dans le scénario 1a. En effet, les croisements les plus prometteurs pour un caractère donné sont ceux produisant des descendants avec à la fois une moyenne et une variance élevées, ce qui permet d'augmenter la probabilité de générer les meilleurs individus. La prédiction de la variance s'est avérée moins précise que celle de la moyenne, avec une corrélation de 0.58 entre la variance prédite et la variance observée.

En conclusion, il s'agit de la première étude sur la prédiction génomique en inter-population chez la vigne utilisant des données expérimentales et impliquant plusieurs croisements. Les précisions de prédiction obtenues pour la moyenne des croisements et les valeurs des individus dans chaque croisement sont prometteuses pour une implémentation pratique de la SG chez la vigne.

6.4 Intérêt de la prédiction phénomique comme alternative à la prédiction génomique chez la vigne

Dans ce chapitre, j'ai testé l'utilisation de spectres au lieu des marqueurs moléculaires pour la prédiction des phénotypes. Cette méthodologie a été originellement proposée par RINCENT et al. (2018) et je l'ai appliquée pour la première fois chez la vigne. L'intérêt de l'utilisation de spectres est qu'ils sont moins chers et plus rapides à acquérir que les données de génotypage. L'hypothèse sous-jacente est que la réflectance à une longueur d'onde donnée dépend

de la composition biochimique dans le tissu considéré et que cette composition est déterminée génétiquement. Cette hypothèse a été testée avec succès sur le blé et le peuplier dans la première étude de RINCENT et al. (2018). Ces résultats ont été confirmés pour plusieurs autres espèces céréalières mais la prédiction phénomique est une nouvelle méthodologie et beaucoup d'interrogations demeurent. Pour mieux comprendre l'intérêt de la prédiction phénomique comparativement à la prédiction génomique, nous avons mesuré en 2020 et 2021 des spectres en proche infra-rouge sur des échantillons séchés de feuilles et de bois, dans le demi-diallèle et le panel de diversité précédemment étudiés. J'ai ainsi pu tester plusieurs modalités d'utilisation des spectres. D'abord, j'ai ajusté un modèle linéaire mixte pour la réflectance à chaque longueur d'onde, afin d'en extraire la composante génétique. Ce modèle était modifié selon que les deux années et/ou les deux tissus étaient inclus ou non. A partir des valeurs génétiques extraites à chaque longueur d'onde, j'ai construit une matrice d'apparentement phénomique, et je l'ai comparée à la matrice d'apparentement génomique. Les résultats montrent que l'année ou le tissu ne modifie presque pas la corrélation (ou co-inertie) entre ces matrices. L'analyse des composantes de la variance par longueur d'onde montre que l'ajustement d'un modèle unique pour les deux tissus conduit à une variance génétique très faible. J'ai donc écarté ce modèle pour la prédiction phénomique. La prédiction phénomique a été implémentée au sein du demi-diallèle et du panel de diversité pour les 15 caractères précédemment étudiés dans le Chapitre 3. Deux méthodes de prédiction ont été comparées : RR-BLUP et GBLUP. La première se base sur une matrice de marqueurs SNP pour la SG et une matrice de réflectance pour la SP. La seconde se base sur la matrice d'apparentement génomique ou phénomique précédemment calculée. J'ai comparé l'utilisation de la moyenne des spectres bruts et l'utilisation des BLUP de réflectance issus des modèles linéaires mixtes décrits ci-dessus. L'utilisation des BLUP de réflectance a conduit à une augmentation drastique de la précision de prédiction. Cette conclusion n'avait jamais été mise en évidence auparavant. En revanche, de manière générale, l'année, le tissu et le pré-traitement des spectres avaient un effet négligeable sur la précision de prédiction. Toutefois, utiliser des spectres issus de deux années et combiner les deux tissus dans un modèle multi-matriciel donnait globalement la meilleure précision de prédiction. J'ai comparé les meilleures précisions de prédiction ainsi obtenues pour la prédiction phénomique, avec celles de la prédiction génomique. Pour quelques caractères dans quelques croisements, la précision de prédiction phénomique était supérieure à celle de la génomique mais en moyenne elle était de –0.2. De façon intéressante, j'ai observé dans les deux populations une corrélation significative entre les deux précisions de prédiction, avec une pente de régression proche de 1. Cela suggère que le rang des précisions de prédiction est conservé entre la prédiction génomique et phénomique, quel que soit le caractère.

Enfin, j'ai combiné dans un même modèle de prédiction les matrices d'apparentement génomique et phénomique, en testant les différents tissus (bois, feuille ou bois + feuille). Globalement, l'addition des matrices phénomiques conduisait à un léger gain de précision de prédiction dans le panel de diversité et à une légère diminution dans le demi-diallèle, comparé à la prédiction génomique seule. Ceci peut s'expliquer par le fait que j'ai cherché à maximiser la

variance génotypique des spectres et à exclure les autres composantes de la variance. De plus, comme les phénotypes ont été mesurés plusieurs années avant les spectres, la précision de prédiction ne peut reposer que sur la composante génétique seule, et non sur le lien entre la nature du spectre et le caractère à prédire, comme l'utilisation de spectres pris sur bois pour prédire la vigueur.

6.5 Discussion générale

6.5.1 Données et populations étudiées

Cette thèse avait pour objectif de proposer des optimisations de la sélection de nouvelles variétés de vigne en faisant appel à la prédiction génomique ou phénomique. Bien que la prédiction génomique ne soit pas une méthodologie nouvelle (la publication de référence date de 2001), les premières applications chez la vigne datent de 2014 et il y a eu assez peu de travaux publiés. Pourtant, la SG gagnerait d'autant plus à être appliquée que le temps de génération est long et que la taille des essais phénotypiques chez la vigne devient vite un problème. Pour mesurer l'intérêt de l'implémentation de la SG dans les programmes d'amélioration vigne, j'ai d'abord comparé plusieurs méthodes de SG, dans une population bi-parentale homogène génétiquement. Les avantages de cette population sont : 1) que l'apparentement théorique est identique et maximal pour tous les individus, 2) que cette descendance a été phénotypée pour de nombreux caractères relatifs à la réponse au stress hydrique, à la fois en conditions semi-contrôlées et au champ. Ensuite, je me suis placée dans un cadre plus proche de celui rencontré dans les programmes d'amélioration, avec une population d'entraînement génétiquement plus éloignée de la population de validation. Pour cela, j'ai utilisé un panel de diversité comme population d'entraînement et des populations bi-parentales apparentées en demi-diallèle comme populations de validation. Ces populations ont été phénotypées pour 15 caractères en commun. L'avantage de ces populations est que les résultats obtenus sont plus facilement extrapolables à d'autres croisements puisque la diversité génétique explorée est plus grande, même si les précisions de prédiction obtenues sont moins élevées. Enfin, dans une optique de baisse de coût et d'augmentation du débit, j'ai testé la sélection phénomique comme alternative à la sélection génomique. Pour simplifier l'analyse, j'ai implémenté la SP seulement en intra-population, en utilisant les populations et les caractères étudiés dans la partie précédente.

6.5.2 Décomposition du modèle de prédiction

Nous pouvons discuter les résultats obtenus en analysant pour chacune des composantes du modèle de prédiction, ses caractéristiques et modifications possibles pour augmenter la précision de prédiction. Le modèle de prédiction global est : $y = X\beta + \epsilon$, avec y le vecteur des phénotypes ajustés pour un caractère donné, X la matrice génotypique des SNP ou la matrice phénomique des réflectances pour chaque longueur d'onde, pour chaque individu, β les effets estimés des SNP ou des réflectances, et ϵ le vecteur de la résiduelle.

Les données phénotypiques sont une composante majeure de l'équation de prédiction. Notamment, nous avons démontré que l'héritabilité du caractère était un paramètre déterminant de la précision de prédiction en intra-population. L'héritabilité peut être améliorée en augmentant le nombre de répétitions, ou en décomposant le caractère en ses composantes physiologiques pour diminuer la complexité du caractère. Néanmoins, le coût de phénotypage constitue encore un goulot d'étranglement dans de nombreux programmes d'amélioration.

La matrice d'entrée du modèle peut prendre des formes différentes. En SG, il s'agit, pour chaque individu et SNP, du nombre d'allèles en commun avec la séquence de référence, ainsi les valeurs possibles sont 0, 1 et 2. En SP, il s'agit de la réflectance pour un individu à une longueur d'onde donnée, soit une valeur numérique comprise entre 0 et 1. D'un point de vue génétique, les populations issues des croisements du demi-diallèle et les sous-populations du panel sont plus différenciées dans la matrice d'apparentement génomique que phénomique.

La méthode d'estimation des effets des variables (estimation du vecteur β) est susceptible d'impacter la précision de prédiction, en interaction avec les autres facteurs mentionnés. A titre d'exemple, j'ai comparé plusieurs méthodes univariées et multivariées dans le chapitre 2, et montré dans le Chapitre 4 la supériorité de GBLUP sur RR-BLUP dans le cas d'une population plus structurée.

6.5.3 Recommandations pour améliorer les programmes d'amélioration vigne

Dans le cadre de la sélection de variétés résistantes à Inter-Rhône, les cépages Syrah et Grenache ont été croisés avec 4 génotypes résistants au mildiou et à l'oïdium. Parmi les descendants, 155 individus ont une résistance polygénique à ces deux maladies et sont génotypés par la technologie GBS (SNP). Pour ce programme spécifiquement, je conseille de prédire précocement les valeurs génétiques des individus avec la SG, en utilisant comme populations d'entraînement les croisements du demi-diallèle apparentés à Syrah ou à Grenache. Ainsi, on peut s'attendre à ce que les précisions de prédiction soient proches de celles obtenues dans le scénario 1b du chapitre 3.

Pour les futurs programmes d'amélioration vigne, un nouveau panel créé dans le cadre du projet ANR SelGenVit pourra être utilisé. Il est composé de 132 variétés et inclut des individus avec des gènes de résistance au mildiou, à l'oïdium mais aussi à la maladie de Pierce ou au black-rot, provenant d'autres espèces de *Vitis*.

Prédiction de la moyenne des croisements Dans le chapitre 3, nous avons vu que la moyenne observée des descendants ne correspondait que rarement à la moyenne observée des parents. Néanmoins, la moyenne des descendants d'un croisement peut être facilement prédite par SG si les parents sont génotypés. La précision de prédiction est variable selon la proximité génétique des parents mais de bonnes précisions peuvent être atteintes si celle-ci est grande. Cette étape requiert qu'un panel de référence soit génotypé et phénotypé pour estimer les effets associés aux marqueurs. La prédiction peut se faire sur tous les croisements possibles entre les variétés de ce panel. En pratique, les programmes auront probablement comme objectif

d'utiliser au moins un parent résistant aux maladies. Dans ce cas, lorsque les croisements sont réalisés, une sélection assistée par marqueurs fait un premier tri parmi les descendants pour les caractères avec un gène majeur (Figure 5.2).

Option 1 : prédiction génomique Une fois que les individus résistants des populations ont été sélectionnés, une option est d'appliquer la SG. Pour cela, les descendants résistants doivent être génotypés. Si le panel de référence est utilisé, les précisions de prédiction attendues sont de l'ordre de celle obtenues dans le scénario 2 du chapitre 3.

Option 2 : prédiction phénomique Pour cette seconde option, les descendants ne sont plus génotypés mais des spectres seront mesurés. Ces mesures devront également être faites dans le panel de référence. Jusqu'ici chez la vigne, la SP n'a pas été testée en inter-population. Il faudra d'abord s'assurer que la diminution de précision de prédiction attendue en inter-population ne soit pas trop forte pour la SP par rapport à celle de la SG.

Gestion du cycle de sélection Les individus prédits au sein de chaque croisement par la SP ou la SG feront l'objet d'un index de sélection combinant les objectifs de sélection pour les différents caractères cibles, pour ne conserver que les individus les plus prometteurs. Ainsi, des essais en champ seront ensuite réalisés pour un nombre réduit d'individus en vue de l'inscription au catalogue. Ces individus pourront être incorporés au modèle de prédiction dans le panel de référence pour les prochains cycles de sélection. Au fur et à mesure des cycles, la diversité génétique va diminuer, ce qui va impacter le progrès génétique sur le long terme. Il faut être vigilant lors de l'étape de sélection des croisements, pour conserver des génotypes peu apparentés, malgré la plus faible précision de prédiction associée. Des contraintes peuvent être ajoutées au modèle de prédiction pour gérer la consanguinité.

6.5.4 Perspectives

La prise en compte de l'interaction GxE est capitale pour les futurs programmes d'amélioration. Pour cela, il est nécessaire d'avoir une population phénotypée plusieurs années dans plusieurs environnements. Certains modèles statistiques sont spécifiquement adaptés pour prendre en compte ces effets, en estimant des covariances entre environnements, en étudiant des normes de réaction à une variable environnementale ou en utilisant un modèle multivarié, où chaque caractère correspond à un phénotype dans un environnement donné.

Dans ma thèse, j'ai utilisé des spectres mesurés dans le proche infra-rouge sur des échantillons séchés en laboratoire mais il existe d'autres dispositifs pour mesurer des spectres. Dans la même gamme de longueur d'onde, il existe des spectromètres portables qui peuvent être utilisé au champ. Sinon, des spectromètres peuvent être embarqués sur des drones pour mesurer des images hyper-spectrales. Toutefois, les spectres issus de ces modalités sont probablement plus bruités que ceux que nous avons utilisés, il faut donc trouver un compromis entre le débit et la qualité des données. Plusieurs interrogations demeurent quant à l'impact de l'utilisation de la SP dans les programmes d'amélioration. La SP est-elle adaptée à tous

les types d'architecture génétique ? Est-ce qu'il y a des zones du génome qui ne sont pas couvertes par les spectres ? Est-ce que la SP permet de mieux prendre en compte les effets génétiques non-additifs ? L'utilisation de simulations permettrait de répondre à certaines de ces questions.

Bibliography

Abdollahi-Arpanahi, R., Gianola, D., & Peñagaricano, F. (2020). Deep Learning versus Parametric and Ensemble Methods for Genomic Prediction of Complex Phenotypes. *Genetics Selection Evolution*, *52*(1), 12. https://doi.org/10.1186/s12711-020-00531-z

Akdemir, D., & Isidro-Sánchez, J. (2019). Design of training populations for selective phenotyping in genomic prediction. *Scientific Reports*, *9*(1), 1446. https://doi.org/10.1038/s41598-018-38081-6

Allier, A., Lehermeier, C., Charcosset, A., Moreau, L., & Teyssèdre, S. (2019). Improving Short- and Long-Term Genetic Gain by Accounting for Within-Family Variance in Optimal Cross-Selection. *Frontiers in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.01006

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Ausseil, A.-G. E., Law, R. M., Parker, A. K., Teixeira, E. I., & Sood, A. (2021). Projected Wine Grape Cultivar Shifts Due to Climate Change in New Zealand. *Frontiers in Plant Science*, *0*. https://doi.org/10.3389/fpls.2021.618039

Azodi, C. B., Bolger, E., McCarren, A., Roantree, M., de los Campos, G., & Shiu, S.-H. (2019). Benchmarking Parametric and Machine Learning Models for Genomic Prediction of Complex Traits. *G3: Genes, Genomes, Genetics*, *9*(11)pmid 31533955, 3691–3702. https://doi.org/10.1534/g3.119.400498

Bacilieri, R., Lacombe, T., Le Cunff, L., Di Vecchi-Staraz, M., Laucou, V., Genna, B., Péros, J.-P., This, P., & Boursiquot, J.-M. (2013). Genetic structure in cultivated grapevines is linked to geography and human selection. *BMC Plant Biology*, *13*(1), 25. https://doi.org/10.1186/1471-2229-13-25

Ballesta, P., Maldonado, C., Pérez-Rodríguez, P., & Mora, F. (2019). SNP and Haplotype-Based Genomic Selection of Quantitative Traits in Eucalyptus globulus. *Plants*, *8*(9), 331. https://doi.org/10.3390/plants8090331

Barnaud, A., Laucou, V., This, P., Lacombe, T., & Doligez, A. (2010). Linkage disequilibrium in wild French grapevine, Vitis vinifera L. subsp . silvestris. *Heredity*, *104*(5), 431–437. https://doi.org/10.1038/hdy.2009.143

Barnaud, A., Lacombe, T., & Doligez, A. (2006). Linkage disequilibrium in cultivated grapevine, Vitis vinifera L. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, *112*(4)pmid 16402190, 708–716. https://doi.org/10.1007/s00122-005-0174-1

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*, *43*(5), 772–777. Retrieved August 25, 2021, from https://www.osapublishing.org/as/abstract.cfm?uri=as-43-5-772

Bayo-Canha, A., Fernández-Fernández, J. I., Martínez-Cutillas, A., & Ruiz-García, L. (2012). Phenotypic segregation and relationships of agronomic traits in Monastrell × Syrah wine grape progeny. *Euphytica*, *186*(2), 393–407. https://doi.org/10.1007/s10681-012-0622-3

Beavis, W. D., Smith, O. S., Grant, D., & Fincher, R. (1994). Identification of quantitative trait loci using a small sample of topcrossed and F4 progeny from maize. *Crop science (USA)*. Retrieved August 27, 2021, from https://scholar.google.com/scholar_lookup?title=Identification+of+quantitative+trait+loci+using+a+small+sample+of+topcrossed+and+F4+progeny+from+maize&author=Beavis%2C+W.D.+%28Pioneer+Hi-Bred+International%2C+Johnston%2C+IA.%29&publication_year=1994

Ben-Sadoun, S., Rincent, R., Auzanneau, J., Oury, F. X., Rolland, B., Heumez, E., Ravel, C., Charmet, G., & Bouchet, S. (2020). Economical optimization of a breeding scheme by selective phenotyping of the calibration set in a multi-trait context: application to bread making quality. *Theoretical and Applied Genetics*. https://doi.org/10.1007/s00122-020-03590-4

Bhattarai, G., Fennell, A., Londo, J. P., Coleman, C., & Kovacs, L. G. (2020). A Novel Grape Downy Mildew Resistance Locus from Vitis rupestris. *American Journal of Enology and Viticulture*. https://doi.org/10.5344/ajev.2020.20030

Bigard, A., Berhe, D. T., Maoddi, E., Sire, Y., Boursiquot, J.-M., Ojeda, H., Péros, J.-P., Doligez, A., Romieu, C., & Torregrosa, L. (2018). Vitis vinifera L. Fruit Diversity to Breed Varieties Anticipating Climate Changes. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.00455

Bouchez, A., Hospital, F., Causse, M., Gallais, A., & Charcosset, A. (2002). Marker-Assisted Introgression of Favorable Alleles at Quantitative Trait Loci Between Maize Elite Lines. *Genetics*, *162*(4), 1945–1959. https://doi.org/10.1093/genetics/162.4.1945

Bouquet, A. (1986). Introduction Dans l'espèce Vitis Vinifera L. d'un Caractère de Résistance à l'o (Uncinula Necator Schw. Burr.) Issu de l'espèce Muscadinia Rotundifolia (Michx.) Small. *Vignevini*, *12*, 141–146.

Bouquet, A., Pauquet, J., Adam-Blondon, A.-F., Torregrosa, L., Merdinoglu, D., & Wiedemann-Merdinoglu, S. (2000). Towards the obtention of grapevine varieties resistant to powdery and downy mildews by conventional breeding and biotechnology. *Progrès Agricole et Viticole*, *117*(18), 383. Retrieved July 28, 2021, from https://hal.inrae.fr/hal-02698762

Bouquet, A., Truel, P., & Wagner, R. (1981). Recurrent selection in grapevine breeding. *Agronomie*, 65–73. https://hal.archives-ouvertes.fr/hal-00884224

Bowers, J., Boursiquot, J.-M., This, P., Chu, K., Johansson, H., & Meredith, C. (1999). Historical Genetics: the Parentage of Chardonnay, Gamay, and Other Wine Grapes of Northeastern France. *Science*, *285*(5433)pmid 10477519, 1562–1565. https://doi.org/10.1126/science.285.5433.1562

Bowers, J., & Meredith, C. P. (1997). The parentage of a classic wine grape, Cabernet Sauvignon. *Nature Genetics*, *16*(1), 84–87. https://doi.org/10.1038/ng0597-84
Bandiera_abtest: a Cg_type: Nature Research Journals Primary_atype: Research

Brandariz, S. P., & Bernardo, R. (2019). Small ad hoc versus large general training populations for genomewide selection in maize biparental crosses. *Theoretical and Applied Genetics*, *132*(2), 347–353. https://doi.org/10.1007/s00122-018-3222-3

Brault, C., Doligez, A., Cunff, L., Coupel-Ledru, A., Simonneau, T., Chiquet, J., This, P., & Flutre, T. (2021a). Harnessing Multivariate, Penalized Regression Methods for Genomic Prediction and QTL Detection of Drought-Related Traits in Grapevine. *G3 Genes | Genomes | Genetics*, *11*(9). https://doi.org/10.1093/g3journal/jkab248

Brault, C., Segura, V., This, P., Cunff, L. L., Flutre, T., François, P., Pons, T., Péros, J.-P., & Doligez, A. (2021b). Across-population genomic prediction in grapevine opens up promising prospects for breeding. *bioRxiv*, 2021.07.29.454290. https://doi.org/10.1101/2021.07.29.454290

Breiman, L. (2001). Random Forests. *Machine Learning*, *1*(45), 5–32. https://doi.org/10.1023/A:1010933404324

Brendel, O., Pot, D., Plomion, C., Rozenberg, P., & Guehl, J.-M. (2002). Genetic parameters and QTL analysis of 13C and ring width in maritime pine. *Plant, Cell & Environment*, *25*(8), 945–953. https://doi.org/10.1046/j.1365-3040.2002.00872.x
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-3040.2002.00872.x

Browning, B. L., & Browning, S. R. (2016). Genotype Imputation with Millions of Reference Samples. *American Journal of Human Genetics*, *98*(1)pmid 26748515, 116–126. https://doi.org/10.1016/j.ajhg.2015.11.020

Burgueño, J., de los Campos, G., Weigel, K., & Crossa, J. (2012). Genomic Prediction of Breeding Values when Modeling Genotype × Environment Interaction using Pedigree and Dense Molecular Markers. *Crop Science*, *52*(2), 707. https://doi.org/10.2135/cropsci2011.06.0299

Caamal-Pat, D., Pérez-Rodríguez, P., Crossa, J., Velasco-Cruz, C., Pérez-Elizalde, S., & Vázquez-Peña, M. (2021). lme4GS: an R-Package for Genomic Selection. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/fgene.2021.680569

Calus, M. P., & Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution*, *43*(1). https://doi.org/10.1186/1297-9686-43-26

Campbell, M. T., Hu, H., Yeats, T. H., Brzozowski, L. J., Caffe-Treml, M., Gutiérrez, L., Smith, K. P., Sorrells, M. E., Gore, M. A., & Jannink, J.-L. (2021). Improving Genomic Prediction for Seed Quality Traits in Oat (Avena sativa L.) Using Trait-Specific Relationship Matrices. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/fgene.2021.643733

Canaguier, A., Grimplet, J., Di Gaspero, G., Scalabrin, S., Duchêne, E., Choisne, N., Mohellibi, N., Guichard, C., Rombauts, S., Le Clainche, I., Bérard, A., Chauveau, A., Bounon, R., Rustenholz, C., Morgante, M., Le Paslier, M.-C., Brunel, D., & Adam-Blondon, A.-F. (2017). A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). *Genomics Data*, *14*, 56–62. https://doi.org/10.1016/j.gdata.2017.09.002

Cantu, D., & Walker, M. A. (Eds.). (2019). *The Grape Genome*. Cham, Springer International Publishing. https://doi.org/10.1007/978-3-030-18601-2

Cazenave, X., Petit, B., Laurens, F., Durel, C.-E., & Muranty, H. (2021). Combining genetic resources and elite material populations to improve the accuracy of genomic prediction in apple, 2021.08.27.457920. https://doi.org/10.1101/2021.08.27.457920

Ceccarelli, S. (1989). Wide adaptation: how wide? *Euphytica*, *40*(3), 197–205. https://doi.org/10.1007/BF00024512

Chaïb, J., Torregrosa, L., Mackenzie, D., Corena, P., Bouquet, A., & Thomas, M. R. (2010). The grape microvine – a model system for rapid forward and reverse genetics of grapevines. *The Plant Journal*, *62*(6), 1083–1092. https://doi.org/10.1111/j.1365-313X.2010.04219.x
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-313X.2010.04219.x

Chang, C. J., Garnier, M., Zreik, L., Rossetti, V., & Bové, J. M. (1993). Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and its identification as a strain ofXylella fastidiosa. *Current Microbiology*, *27*(3)pmid 23835746, 137–142. https://doi.org/10.1007/BF01576010

Charmet, G., Storlie, E., Oury, F. X., Laurent, V., Beghin, D., Chevarin, L., Lapierre, A., Perretant, M. R., Rolland, B., Heumez, E., Duchalais, L., Goudemand, E., Bordes, J., & Robert, O. (2014). Genome-wide prediction of three important traits in bread wheat. *Molecular Breeding*, *34*(4), 1843–1852. https://doi.org/10.1007/s11032-014-0143-y

Cho, S., Kim, H., Oh, S., Kim, K., & Park, T. (2009). Elastic-net regularization approaches for genome-wide association studies of rheumatoid arthritis. *BMC Proceedings*, *3*, S25. https://doi.org/10.1186/1753-6561-3-s7-s25

Claverie, M., Notaro, M., Fontaine, F., & Wéry, J. (2020). Current knowledge on Grapevine Trunk Diseases with complex etiology: a systemic approach. *Phytopathologia Mediterranea*, *59*(1), 29. https://doi.org/10.14601/Phyto-11150

Coleman, C., Copetti, D., Cipriani, G., Hoffmann, S., Kozma, P., Kovács, L., Morgante, M., Testolin, R., & Di Gaspero, G. (2009). The powdery mildew resistance gene REN1 co-segregates with an NBS-LRR gene cluster in two Central Asian grapevines. *BMC genetics*, *10*pmid 20042081, 89. https://doi.org/10.1186/1471-2156-10-89

Colleau, J.-J., Palhière, I., Rodríguez-Ramilo, S. T., & Legarra, A. (2017). A Fast Indirect Method to Compute Functions of Genomic Relationships Concerning Genotyped and Ungenotyped Individuals, for Diversity Management. *Genetics Selection Evolution*, *49*(1), 87. https://doi.org/10.1186/s12711-017-0363-9

Consortium, R., Fugeray-Scarbel, A., Bastien, C., Dupont-Nivet, M., & Lemarié, S. (2021). Why and How to Switch to Genomic Selection: lessons From Plant and Animal Breeding Experience. *Frontiers in Genetics*, *0*. https://doi.org/10.3389/fgene.2021.629737

Coupel-Ledru, A. (2015, November 3). *Déterminisme Physiologique et Génétique de l'utilisation de l'eau Chez La Vigne* (thesis). Montpellier, SupAgro. Retrieved February 20, 2019, from http://www.theses.fr/2015NSAM0020

Coupel-Ledru, A., Lebon, E., Christophe, A., Gallo, A., Gago, P., Pantin, F., Doligez, A., & Simonneau, T. (2016). Reduced nighttime transpiration is a relevant breeding target for high water-use efficiency in grapevine. *Proceedings of the National Academy of Sciences*, *113*(32), 8963–8968. https://doi.org/10.1073/pnas.1600826113

Coupel-Ledru, A., Lebon, É., Christophe, A., Doligez, A., Cabrera-Bosquet, L., Péchier, P., Hamard, P., This, P., & Simonneau, T. (2014). Genetic variation in a grapevine progeny (Vitis vinifera L. cvs Grenache×Syrah) reveals inconsistencies between maintenance of daytime leaf water potential and response of transpiration rate under drought. *Journal of Experimental Botany*, *65*(21), 6205–6218. https://doi.org/10.1093/jxb/eru228

Crossa, J., Fritsche-Neto, R., Montesinos-Lopez, O. A., Costa-Neto, G., Dreisigacker, S., Montesinos-Lopez, A., & Bentley, A. R. (2021). The Modern Plant Breeding Triangle: Optimizing the Use of Genomics, Phenomics, and Enviromics Data. *Frontiers in Plant Science*, *12*pmid 33936136, 651480. https://doi.org/10.3389/fpls.2021.651480

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de Los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017). Genomic Selection in Plant Breeding: methods, Models, and Perspectives. *Trends in Plant Science*, *22*(11)pmid 28965742, 961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Cuevas, J., Montesinos-López, O., Juliana, P., Guzmán, C., Pérez-Rodríguez, P., González-Bucio, J., Burgueño, J., Montesinos-López, A., & Crossa, J. (2019). Deep Kernel for Genomic and Near Infrared Predictions in Multi-environment Breeding Trials. *G3 Genes | Genomes | Genetics*, *9*(9), 2913–2924. https://doi.org/10.1534/g3.119.400493

Dagnachew, B., & Meuwissen, T. (2019). Accuracy of Within-Family Multi-Trait Genomic Selection Models in a Sib-Based Aquaculture Breeding Scheme. *Aquaculture*, *505*, 27–33. https://doi.org/10.1016/j.aquaculture.2019.02.036

da Silva, É. D. B., Xavier, A., & Faria, M. V. (2021). Impact of Genomic Prediction Model, Selection Intensity, and Breeding Strategy on the Long-Term Genetic Gain and Genetic Erosion in Soybean Breeding. *Frontiers in Genetics*, *12*pmid 34539725, 637133. https://doi.org/10.3389/fgene.2021.637133

de los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., & Calus, M. P. L. (2013). Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. *Genetics*, *193*(2), 327–345. https://doi.org/10.1534/genetics.112.143313

de Bem Oliveira, I., Amadeu, R. R., Ferrão, L. F. V., & Muñoz, P. R. (2020). Optimizing whole-genomic prediction for autotetraploid blueberry breeding. *Heredity*, *125*(6), 437–448. https://doi.org/10.1038/s41437-020-00357-x

de Cortázar-Atauri, I. G., Duchêne, E., Destrac-Irvine, A., Barbeau, G., de Rességuier, L., Lacombe, T., Parker, A. K., Saurin, N., & van Leeuwen, C. (2017). Grapevine phenology in France: from past observations to future evolutions in the context of climate change. *OENO One*, *51*(2), 115–126. https://doi.org/10.20870/oeno-one.2017.51.2.1622

de Roos, A. P. W., Hayes, B. J., & Goddard, M. (2009). Reliability of Genomic Predictions Across Multiple Populations. *Genetics*, *183*(4), 1545–1553. https://doi.org/10.1534/genetics.109.104935

Desta, Z. A., & Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends in Plant Science*, *19*(9), 592–601. https://doi.org/10.1016/j.tplants.2014.05.006

Diago, M. P., Tardaguila, J., Aleixos, N., Millan, B., Prats-Montalban, J. M., Cubero, S., & Blasco, J. (2015). Assessment of cluster yield components by image analysis. *Journal of the Science of Food and Agriculture*, *95*(6)pmid 25041796, 1274–1282. https://doi.org/10.1002/jsfa.6819

Dolédec, S., & Chessel, D. (1994). Co-Inertia Analysis: An Alternative Method for Studying Species: Environment Relationships. *Co-inertia analysis: an alternative method for studying species: environment relationships*, *31*(3), 277–293.

Doligez, A., Adam-Blondon, A. F., Cipriani, G., Di Gaspero, G., Laucou, V., Merdinoglu, D., Meredith, C. P., Riaz, S., Roux, C., & This, P. (2006). An integrated SSR map of grapevine based on five mapping populations. *Theoretical and Applied Genetics*, *113*(3), 369–382. https://doi.org/10.1007/s00122-006-0295-1

Doligez, A., Bertrand, Y., Farnos, M., Grolier, M., Romieu, C., Esnault, F., Dias, S., Berger, G., François, P., Pons, T., Ortigosa, P., Roux, C., Houel, C., Laucou, V., Bacilieri, R., Péros, J.-P., & This, P. (2013). New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (Vitis vinifera L.) *BMC Plant Biology*, *13*(1), 217. https://doi.org/10.1186/1471-2229-13-217

Ducasse, M.-A., Roy, A., Caboulet, D., Saurin, N., Aguera, & Bes, M. (2019, June). *Validation of Microscale Red Wine Fermentations (<1kg) on Oenological Parameters, Color and Aroma Compounds: An Easy Red Wine Fermentation Method to Perform Many Experiments on Vines and Wines* (Symposium).

Duchêne, E. (2016). How can grapevine genetics contribute to the adaptation to climate change? *OENO One*, *50*(3). https://doi.org/10.20870/oeno-one.2016.50.3.98

Duchêne, E., Butterlin, G., Dumas, V., & Merdinoglu, D. (2012). Towards the adaptation of grapevine varieties to climate change: QTLs and candidate genes for developmental stages. *Theoretical and Applied Genetics*, *124*(4), 623–635. https://doi.org/10.1007/s00122-011-1734-1

Duchêne, E., Huard, F., Dumas, V., Schneider, C., & Merdinoglu, D. (2010). The challenge of adapting grapevine varieties to climate change. *Climate Research*, *41*(3), 193–204. https://doi.org/10.3354/cr00850

Eibach, R., Zyprian, E., Welter, L., & Töpfer, R. (2007). The use of molecular markers for pyramiding resistance genes in grapevine breeding, 6.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLOS ONE*, *6*(5), e19379. https://doi.org/10.1371/journal.pone.0019379

Erbe, M., Hayes, B. J., Matukumalli, L. K., Goswami, S., Bowman, P. J., Reich, C. M., Mason, B. A., & Goddard, M. (2012). Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *Journal of Dairy Science*, *95*(7), 4114–4129. https://doi.org/10.3168/jds.2011-5019

Falconer, D. S., & Mackay, T. F. C. (2009). *Introduction to quantitative genetics* (4. ed., [16. print.]). Harlow, Pearson, Prentice Hall
OCLC: 552422435.

Farquhar, G., & Richards, R. (1984). Isotopic Composition of Plant Carbon Correlates with Water-Use Efficiency of Wheat Genotypes. *Australian Journal of Plant Physiology*, *11*(6), 539–552. https://doi.org/10.1071/PP9840539
WOS:A1984AAV6800010

Feher, K., Lisec, J., Römisch-Margl, L., Selbig, J., Gierl, A., Piepho, H.-P., Nikoloski, Z., & Willmitzer, L. (2014). Deducing Hybrid Performance from Parental Metabolic Profiles of Young Primary Roots of Maize by Using a Multivariate Diallel Approach. *PLOS ONE*, *9*(1), e85435. https://doi.org/10.1371/journal.pone.0085435

Fernandes, S. B., Dias, K. O. G., Ferreira, D. F., & Brown, P. J. (2018). Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. *Theoretical and Applied Genetics*, *131*(3), 747–755. https://doi.org/10.1007/s00122-017-3033-y

Fernandez, O., Urrutia, M., Bernillon, S., Giauffret, C., Tardieu, F., Le Gouis, J., Langlade, N., Charcosset, A., Moing, A., & Gibon, Y. (2016). Fortune telling: metabolic markers of plant performance. *Metabolomics*, *12*(10), 158. https://doi.org/10.1007/s11306-016-1099-1

Fernández, J., Villanueva, B., & Toro, M. A. (2021). Optimum Mating Designs for Exploiting Dominance in Genomic Selection Schemes for Aquaculture Species. *Genetics Selection Evolution*, *53*(1), 14. https://doi.org/10.1186/s12711-021-00610-9

Flutre, T., Cunff, L. L., Fodor, A., Launay, A., Romieu, C., Berger, G., Bertrand, Y., Beccavin, I., Bouckenooghe, V., Roques, M., Pinasseau, L., Verbaere, A., Sommerer, N., Cheynier, V., Bacilieri, R., Boursiquot, J. M., Lacombe, T., Laucou, V., This, P., … Doligez, A. (2020). Genome-wide association and prediction studies using a grapevine diversity

panel give insights into the genetic architecture of several traits of interest. *bioRxiv*, 2020.09.10.290890. https://doi.org/10.1101/2020.09.10.290890

Fodor, A., Segura, V., Denis, M., Neuenschwander, S., Fournier-Level, A., Chatelet, P., Homa, F. A. A., Lacombe, T., This, P., & Le Cunff, L. (2014). Genome-Wide Prediction Methods in Highly Diverse and Heterozygous Species: proof-of-Concept through Simulation in Grapevine (R. Khanin, Ed.). *PLoS ONE*, *9*(11), e110436. https://doi.org/10.1371/journal.pone.0110436

Fournier-Level, A., Le Cunff, L., Gomez, C., Doligez, A., Ageorges, A., Roux, C., Bertrand, Y., Souquet, J.-M., Cheynier, V., & This, P. (2009). Quantitative Genetic Bases of Anthocyanin Variation in Grape (Vitis vinifera L. ssp. sativa) Berry: a Quantitative Trait Locus to Quantitative Trait Nucleotide Integrated Study. *Genetics*, *183*(3), 1127–1139. https://doi.org/10.1534/genetics.109.103929

Fournier-Level, A., Hugueney, P., Verriès, C., This, P., & Ageorges, A. (2011). Genetic mechanisms underlying the methylation level of anthocyanins in grape (Vitis vinifera L.) *BMC Plant Biology*, *11*(1), 179. https://doi.org/10.1186/1471-2229-11-179

FranceAgriMer. (2019). *La Filière Vin | FranceAgriMer - Établissement National Des Produits de l'agriculture et de La Mer*. La filière Vin | FranceAgriMer - établissement national des produits de l'agriculture et de la mer. Retrieved August 30, 2021, from https://www.franceagrimer.fr/filieres-Vin-et-cidre/Vin/La-filiere-Vin

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, *33*(1). https://doi.org/10.18637/jss.v033.i01

Fuchs, M., Golino, D. A., Martelli, G. P., & Meng, B. (Eds.). (2017). *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (1st ed. 2017). Cham, Springer International Publishing : Imprint: Springer. https://doi.org/10.1007/978-3-319-57706-7

Galán, R. J., Bernal-Vasquez, A.-M., Jebsen, C., Piepho, H.-P., Thorwarth, P., Steffan, P., Gordillo, A., & Miedaner, T. (2020). Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye. *Theoretical and Applied Genetics*. https://doi.org/10.1007/s00122-020-03651-8

Galán, R. J., Bernal-Vasquez, A.-M., Jebsen, C., Piepho, H.-P., Thorwarth, P., Steffan, P., Gordillo, A., & Miedaner, T. (2021). Early prediction of biomass in hybrid rye based on hyperspectral data surpasses genomic predictability in less-related breeding material. *Theoretical and Applied Genetics*, *134*(5), 1409–1422. https://doi.org/10.1007/s00122-021-03779-1

Galet, P. (2000). *Dictionnaire encyclopédique des cépages*. Paris, Hachette.

Gambetta, G. A., Herrera, J. C., Dayer, S., Feng, Q., Hochberg, U., & Castellarin, S. D. (2020). The Physiology of Drought Stress in Grapevine: Towards an Integrative Definition of Drought Tolerance. *Journal of Experimental Botany*, *71*(16), 4658–4676. https://doi.org/10.1093/jxb/eraa245

Gianola, D., & Fernando, R. L. (1986). Bayesian Methods in Animal Breeding Theory. *Journal of Animal Science*, *63*(1), 217–244. https://doi.org/10.2527/jas1986.631217x

Gianola, D., Fernando, R. L., & Stella, A. (2006). Genomic-Assisted Prediction of Genetic Value With Semiparametric Procedures. *Genetics*, *173*(3)pmid 16648593, 1761–1776. https://doi.org/10.1534/genetics.105.049510

Gianola, D., & van Kaam, J. B. C. H. M. (2008). Reproducing Kernel Hilbert Spaces Regression Methods for Genomic Assisted Prediction of Quantitative Traits. *Genetics*, *178*(4)pmid 18430950, 2289–2303. https://doi.org/10.1534/genetics.107.084285

Gonçalves, E., Carrasquinho, I., Almeida, R., Pedroso, V., & Martins, A. (2016). Genetic correlations in grapevine and their effects on selection. *Australian Journal of Grape and Wine Research*, *22*(1), 52–63. https://doi.org/10.1111/ajgw.12164
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajgw.12164

Gonçalves, E., Carrasquinho, I., & Martins, A. (2020). Measure to evaluate the sensitivity to genotype-by-environment interaction in grapevine clones. *Australian Journal of Grape and Wine Research*, *26*(3), 259–270. https://doi.org/10.1111/ajgw.12432
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajgw.12432

González-Diéguez, D., Legarra, A., Charcosset, A., Moreau, L., Lehermeier, C., Teyssèdre, S., & Vitezica, Z. G. (2021). Genomic Prediction of Hybrid Crops Allows Disentangling Dominance and Epistasis. *Genetics*. https://doi.org/10.1093/genetics/iyab026

González-Recio, O., Rosa, G. J. M., & Gianola, D. (2014). Machine learning methods and predictive ability metrics for genome-wide prediction of complex traits. *Livestock Science*, *166*, 217–231. https://doi.org/10.1016/j.livsci.2014.05.036

Granato, I., Cuevas, J., Luna-Vázquez, F., Crossa, J., Montesinos-López, O., Burgueño, J., & Fritsche-Neto, R. (2018). BGGE: a New Package for Genomic-Enabled Prediction Incorporating Genotype × Environment Interaction Models. *G3 Genes | Genomes | Genetics*, *8*(9), 3039–3047. https://doi.org/10.1534/g3.118.200435

Grattapaglia, D., Silva-Junior, O. B., Resende, R. T., Cappa, E. P., Müller, B. S. F., Tan, B., Isik, F., Ratcliffe, B., & El-Kassaby, Y. A. (2018). Quantitative Genetics and Genomics Converge to Accelerate Forest Tree Breeding. *Frontiers in Plant Science*, *9*, 1693. https://doi.org/10.3389/fpls.2018.01693

Grosman, J., & Doublet, B. (2012). Maladies Du Bois de La Vigne. Synthèse Des Dispositifs d'observation Au Vignoble, de l'observatoire 2003-2008 Au Réseau d'épidémio-Surveillance Actuel. *Phytoma – La Défense des Végétaux*, (651), 31–35. https://www.maladie-du-bois-vigne.fr/content/download/3817/37535/version/1/file/grosman-doublet+2012.pdf

Grzeskowiak, L., Costantini, L., Lorenzi, S., & Grando, M. S. (2013). Candidate loci for phenology and fruitfulness contributing to the phenotypic variability observed in grapevine. *Theoretical and Applied Genetics*, *126*(11), 2763–2776. https://doi.org/10.1007/s00122-013-2170-1

Grzybowski, M., Wijewardane, N. K., Atefi, A., Ge, Y., & Schnable, J. C. (2021). Hyperspectral reflectance-based phenotyping for quantitative genetics in crops: progress and challenges. *Plant Communications*, *2*(4), 100209. https://doi.org/10.1016/j.xplc.2021.100209

Guimier, S., Delmotte, F., Miclot, A., Fabre, F., Mazet, I., Couture, C., Schneider, C., & Delière, L. (2019). OSCAR, a national observatory to support the durable deployment of disease-resistant grapevine cultivars. *Acta Horticulturae*, (1248), 21–34. https://doi.org/10.17660/ActaHortic.2019.1248.4

Guo, G., Zhao, F., Wang, Y., Zhang, Y., Du, L., & Su, G. (2014). Comparison of single-trait and multiple-trait genomic prediction models. *BMC Genetics*, *15*(1), 30. https://doi.org/10.1186/1471-2156-15-30

Guo, Z., Magwire, M. M., Basten, C. J., Xu, Z., & Wang, D. (2016). Evaluation of the utility of gene expression and metabolic information for genomic prediction in maize. *Theoretical and Applied Genetics*, *129*(12), 2413–2427. https://doi.org/10.1007/s00122-016-2780-5

Habier, D., Fernando, R. L., & Dekkers, J. C. M. (2007). The Impact of Genetic Relationship Information on Genome-Assisted Breeding Values. *Genetics*, *177*(4), 2389–2397. https://doi.org/10.1534/genetics.107.081190

Habier, D., Fernando, R. L., & Garrick, D. J. (2013). Genomic BLUP Decoded: A Look into the Black Box of Genomic Prediction. *Genetics*, *194*(3), 597–607. https://doi.org/10.1534/genetics.113.152207

Habier, D., Fernando, R. L., Kizilkaya, K., & Garrick, D. J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC bioinformatics*, *12*pmid 21605355, 186. https://doi.org/10.1186/1471-2105-12-186

Habier, D., Tetens, J., Seefried, F.-R., Lichtner, P., & Thaller, G. (2010). The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genetics Selection Evolution*, *42*(1), 5. https://doi.org/10.1186/1297-9686-42-5

Hartung, J. S., Beretta, J., Brlansky, R. H., Spisso, J., & Lee, R. F. (1994). Citrus variegated chlorosis bacterium: axenic culture, pathogenicity, and serological relationships with other strains of Xylella fastidiosa. *Phytopathology*, *84*(6), 591–597. Retrieved August 4, 2021, from https://www.cabdirect.org/cabdirect/abstract/19942307663

Hayes, B., & Goddard, M. (2001). The distribution of the effects of genes affecting quantitative traits in livestock. *Genetics Selection Evolution*, *33*(3), 209. https://doi.org/10.1186/1297-9686-33-3-209

Heffner, E. L., Sorrells, M. E., & Jannink, J.-L. (2009). Genomic Selection for Crop Improvement. *Crop Science*, *49*(1), 1. https://doi.org/10.2135/cropsci2008.08.0512

Henderson, C. R. (1985). Equivalent Linear Models to Reduce Computations. *Journal of Dairy Science*, *68*(9), 2267–2277. https://doi.org/10.3168/jds.S0022-0302(85)81099-7

Hess, M., Druet, T., Hess, A., & Garrick, D. (2017). Fixed-Length Haplotypes Can Improve Genomic Prediction Accuracy in an Admixed Dairy Cattle Population. *Genetics Selection Evolution*, *49*(1), 54. https://doi.org/10.1186/s12711-017-0329-y

Hickey, J. M., Chiurugwi, T., Mackay, I., & Powell, W. (2017). Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature Genetics*, *49*(9), 1297–1303. https://doi.org/10.1038/ng.3920

Bandiera_abtest: a Cg_type: Nature Research Journals Primary_atype: Reviews Subject_term: Genetics;Genomics;Plant breeding Subject_term_id: genetics;genomics;plant-breeding

Hochberg, U., Rockwell, F. E., Holbrook, N. M., & Cochard, H. (2018). Iso/Anisohydry: a Plant–Environment Interaction Rather Than a Simple Hydraulic Trait. *Trends in Plant Science*, *23*(2), 112–120. https://doi.org/10.1016/j.tplants.2017.11.002

Holroyd, S. E. (2013). The Use of near Infrared Spectroscopy on Milk and Milk Products. *Journal of Near Infrared Spectroscopy*, *21*(5), 311–322. https://doi.org/10.1255/jnirs.1055

Huang, W., & Mackay, T. F. C. (2016). The Genetic Architecture of Quantitative Traits Cannot Be Inferred from Variance Component Analysis (X. Zhu, Ed.). *PLOS Genetics*, *12*(11), e1006421. https://doi.org/10.1371/journal.pgen.1006421

Huang, Y.-F., Bertrand, Y., Guiraud, J.-L., Vialet, S., Launay, A., Cheynier, V., Terrier, N., & This, P. (2013). Expression QTL mapping in grapevine—Revisiting the genetic determinism of grape skin colour. *Plant Science*, *207*, 18–24. https://doi.org/10.1016/j.plantsci.2013.02.011

Huang, Y.-F., Doligez, A., Fournier-Level, A., Le Cunff, L., Bertrand, Y., Canaguier, A., Morel, C., Miralles, V., Veran, F., Souquet, J.-M., Cheynier, V., Terrier, N., & This, P. (2012). Dissecting genetic architecture of grape proanthocyanidin composition through quantitative trait locus mapping. *BMC Plant Biology*, *12*(1), 30. https://doi.org/10.1186/1471-2229-12-30

Huang, Y.-F., Vialet, S., Guiraud, J.-L., Torregrosa, L., Bertrand, Y., Cheynier, V., This, P., & Terrier, N. (2014). A negative MYB regulator of proanthocyanidin accumulation, identified through expression quantitative locus mapping in the grape berry. *New Phytologist*, *201*(3), 795–809. https://doi.org/10.1111/nph.12557

IPCC, Shukla, P. R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H.-O., Roberts, D. C., Zhai, P., Slade, R., Connors, S., van Diemen, R., Ferrat, M., Haughey, E., Luz, S., Neogi, S., Pathak, M., Petzold, J., Portugal Pereira, J., Vyas, P., … Malley, J. (2021). IPCC, 2019: Summary for Policymakers. In: Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems.

Itoh, Y., & Yamada, Y. (1988). Selection indices for desired relative genetic gains with inequality constraints. *Theoretical and Applied Genetics*, *75*(5), 731–735. https://doi.org/10.1007/BF00265596

Jacquin, L., Cao, T.-V., & Ahmadi, N. (2016). A Unified and Comprehensible View of Parametric and Kernel Methods for Genomic Prediction with Application to Rice. *Frontiers in Genetics*, *0*. https://doi.org/10.3389/fgene.2016.00145

Jaillon, O., Aury, J.-M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyère, C., … The French–Italian Public Consortium for Grapevine Genome Characterization. (2007). The grapevine genome

sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, *449*(7161), 463–467. https://doi.org/10.1038/nature06148

Jannink, J.-L., Lorenz, A. J., & Iwata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Briefings in Functional Genomics*, *9*(2), 166–177. https://doi.org/10.1093/bfgp/elq001

Jia, Y., & Jannink, J.-L. (2012). Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics*, *192*(4)pmid 23086217, 1513–1522. https://doi.org/10.1534/genetics.112.144246

Jiang, J., Zhang, Q., Ma, L., Li, J., Wang, Z., & Liu, J.-F. (2015). Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. *Heredity*, *115*(1), 29–36. https://doi.org/10.1038/hdy.2015.9

Jiang, Y., & Li, C. (2020). Convolutional Neural Networks for Image-Based High-Throughput Plant Phenotyping: a Review. *Plant Phenomics*, *2020*. https://doi.org/10.34133/2020/4152816

Kadam, D. C., & Lorenz, A. J. (2019). Evaluation of Nonparametric Models for Genomic Prediction of Early-Stage Single Crosses in Maize. *Crop Science*, *59*(4), 1411–1423. https://doi.org/10.2135/cropsci2017.11.0668
_eprint: https://acsess.onlinelibrary.wiley.com/doi/pdf/10.2135/cropsci2017.11.0668

Kalia, R. K., Rai, M. K., Kalia, S., Singh, R., & Dhawan, A. K. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica*, *177*(3), 309–334. https://doi.org/10.1007/s10681-010-0286-9

Kempthorne, O., & Nordskog, A. W. (1959). Restricted Selection Indices. *Biometrics*, *15*(1)jstor 2527598, 10–19. https://doi.org/10.2307/2527598

Kosugi, S., Momozawa, Y., Liu, X., Terao, C., Kubo, M., & Kamatani, Y. (2019). Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. *Genome Biology*, *20*(1), 117. https://doi.org/10.1186/s13059-019-1720-5

Krause, M. R., González-Pérez, L., Crossa, J., Pérez-Rodríguez, P., Montesinos-López, O., Singh, R. P., Dreisigacker, S., Poland, J., Rutkoski, J., Sorrells, M., Gore, M. A., & Mondal, S. (2019). Hyperspectral Reflectance-Derived Relationship Matrices for Genomic Prediction of Grain Yield in Wheat. *G3 Genes | Genomes | Genetics*, g3.200856.2018. https://doi.org/10.1534/g3.118.200856

Krishnappa, G., Savadi, S., Tyagi, B. S., Singh, S. K., Mamrutha, H. M., Kumar, S., Mishra, C. N., Khan, H., Gangadhara, K., Uday, G., Singh, G., & Singh, G. P. (2021). Integrated genomic selection for rapid improvement of crops. *Genomics*, *113*(3)pmid 33610797, 1070–1086. https://doi.org/10.1016/j.ygeno.2021.02.007

Krivanek, A. F., Riaz, S., & Walker, M. A. (2006). Identification and molecular mapping of PdR1, a primary resistance gene to Pierce's disease in Vitis. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, *112*(6)pmid 16435126, 1125–1131. https://doi.org/10.1007/s00122-006-0214-5

Lacombe, T., Boursiquot, J.-M., Laucou, V., Di Vecchi-Staraz, M., Péros, J.-P., & This, P. (2013). Large-scale parentage analysis in an extended set of grapevine cultivars (Vitis vinifera

L.) *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, *126*(2)pmid 23015217, 401–414. https://doi.org/10.1007/s00122-012-1988-2

Lane, H. M., Murray, S. C., Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Rooney, D. K., Barrero-Farfan, I. D., Fuente, G. N. D. L., & Morgan, C. L. S. (2020). Phenomic selection and prediction of maize grain yield from near-infrared reflectance spectroscopy of kernels. *The Plant Phenome Journal*, *3*(1), e20002. https://doi.org/10.1002/ppj2.20002

_eprint: https://acsess.onlinelibrary.wiley.com/doi/pdf/10.1002/ppj2.20002

Laucou, V., Lacombe, T., Dechesne, F., Siret, R., Bruno, J.-P., Dessup, M., Dessup, T., Ortigosa, P., Parra, P., Roux, C., Santoni, S., Varès, D., Péros, J.-P., Boursiquot, J.-M., & This, P. (2011). High throughput analysis of grape genetic diversity as a tool for germplasm collection management. *Theoretical and Applied Genetics*, *122*(6), 1233–1245. https://doi.org/10.1007/s00122-010-1527-y

Laucou, V., Launay, A., Bacilieri, R., Lacombe, T., Adam-Blondon, A.-F., Bérard, A., Chauveau, A., de Andrés, M. T., Hausmann, L., Ibáñez, J., Paslier, M.-C. L., Maghradze, D., Martinez-Zapater, J. M., Maul, E., Ponnaiah, M., Töpfer, R., Péros, J.-P., & Boursiquot, J.-M. (2018). Extended diversity analysis of cultivated grapevine Vitis vinifera with 10K genome-wide SNPs. *PLOS ONE*, *13*(2), e0192540. https://doi.org/10.1371/journal.pone.0192540

Lauri, P.-É., Barigah, T. S., Lopez, G., Martinez, S., Losciale, P., Zibordi, M., Manfrini, L., Corelli-Grappadelli, L., Costes, E., & Regnard, J.-L. (2016). Genetic variability and phenotypic plasticity of apple morphological responses to soil water restriction in relation with leaf functions and stem xylem conductivity. *Trees*, *30*(5), 1893–1908. https://doi.org/10.1007/s00468-016-1408-3

Lebedev, V. G., Lebedeva, T. N., Chernodubov, A. I., & Shestibratov, K. A. (2020). Genomic Selection for Forest Tree Improvement: methods, Achievements and Perspectives. *Forests*, *11*(11), 1190. https://doi.org/10.3390/f11111190

Liang, Z., Owens, C. L., Zhong, G.-Y., & Cheng, L. (2011). Polyphenolic profiles detected in the ripe berries of Vitis vinifera germplasm. *Food Chemistry*, *129*(3), 940–950. https://doi.org/10.1016/j.foodchem.2011.05.050

Lin, C. Y. (2005). A Simultaneous Procedure for Deriving Selection Indexes with Multiple Restrictions1. *Journal of Animal Science*, *83*(3), 531–536. https://doi.org/10.2527/2005.833531x

Lorenz, A. J., & Smith, K. P. (2015–1, December). Adding Genetically Distant Individuals to Training Populations Reduces Genomic Prediction Accuracy in Barley. *Crop Science*, *55*(6), 2657–2667. https://doi.org/10.2135/cropsci2014.12.0827

Lucini, L., Miras-Moreno, B., Busconi, M., Marocco, A., Gatti, M., & Poni, S. (2020). Molecular basis of rootstock-related tolerance to water deficit in Vitis vinifera L. cv. Sangiovese: a physiological and metabolomic combined approach. *Plant Science*, 110600. https://doi.org/10.1016/j.plantsci.2020.110600

Lynch, M., & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, Mass, Sinauer.

MacLeod, I. M., Bowman, P. J., Vander Jagt, C. J., Haile-Mariam, M., Kemper, K. E., Chamberlain, A. J., Schrooten, C., Hayes, B., & Goddard, M. (2016). Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics*, *17*(1), 144. https://doi.org/10.1186/s12864-016-2443-6

Marchal, A., Legarra, A., Tisné, S., Carasco-Lacombe, C., Manez, A., Suryana, E., Omoré, A., Nouy, B., Durand-Gasselin, T., Sánchez, L., Bouvet, J.-M., & Cros, D. (2016). Multivariate genomic model improves analysis of oil palm (Elaeis guineensis Jacq.) progeny tests. *Molecular Breeding*, *36*(1). https://doi.org/10.1007/s11032-015-0423-1

Marguerit, E., Brendel, O., Lebon, E., Leeuwen, C. V., & Ollat, N. (2012). Rootstock control of scion transpiration and its acclimation to water deficit are controlled by different genes. *New Phytologist*, *194*(2), 416–429. https://doi.org/10.1111/j.1469-8137.2012.04059.x

Meng, B., Martelli, G. P., Golino, D. A., & Fuchs, M. (2017, July 5). *Grapevine Viruses: molecular Biology, Diagnostics and Management*. Springer.

Meng, C., Kuster, B., Culhane, A. C., & Gholami, A. M. (2014). A Multivariate Approach to the Integration of Multi-Omics Datasets. *BMC Bioinformatics*, *15*(1), 162. https://doi.org/10.1186/1471-2105-15-162

Menzel, A., Yuan, Y., Matiu, M., Sparks, T., Scheifinger, H., Gehrig, R., & Estrella, N. (2020). Climate change fingerprints in recent European plant phenology. *Global Change Biology*, *26*(4), 2599–2612. https://doi.org/10.1111/gcb.15000
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.15000

Merdinoglu, D., Schneider, C., Prado, E., Wiedemann-Merdinoglu, S., & Mestre, P. (2018). Breeding for durable resistance to downy and powdery mildew in grapevine. *OENO One*, *52*(3), 203–209. https://doi.org/10.20870/oeno-one.2018.52.3.2116

Merdinoglu, D., Wiedeman-Merdinoglu, S., Coste, P., Dumas, V., Haetty, S., Butterlin, G., & Greif, C. (2003). Genetic Analysis of Downy Mildew Resistance Derived from Muscadinia rotundifolia. *Acta Horticulturae*. Retrieved August 4, 2021, from https://pubag.nal.usda.gov/catalog/302359

Messina, C. D., Technow, F., Tang, T., Totir, R., Gho, C., & Cooper, M. (2018). Leveraging biological insight and environmental variation to improve phenotypic prediction: integrating crop growth models (CGM) with whole genome prediction (WGP). *European Journal of Agronomy*, *100*, 151–162. https://doi.org/10.1016/j.eja.2018.01.007

Meuwissen, T., Hayes, B., & Goddard, M. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, 11.

Meuwissen, T., Hayes, B., & Goddard, M. (2016). Genomic Selection: A Paradigm Shift in Animal Breeding. *Animal Frontiers*, *6*(1), 6–14. https://doi.org/10.2527/af.2016-0002

Michel, S., Löschenberger, F., Ametz, C., Pachler, B., Sparry, E., & Bürstmayr, H. (2019). Simultaneous selection for grain yield and protein content in genomics-assisted wheat

breeding. *Theoretical and Applied Genetics*, *132*(6), 1745–1760. https://doi.org/10.1007/s00122-019-03312-5

Migicovsky, Z., Cousins, P., Jordan, L. M., Myles, S., Striegler, R. K., Verdegaal, P., & Chitwood, D. H. (2021). Grapevine rootstocks affect growth-related scion phenotypes. *Plant Direct*, *5*(5), e00324. https://doi.org/10.1002/pld3.324
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pld3.324

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., & Myles, S. (2017). Patterns of genomic and phenomic diversity in wine and table grapes. *Horticulture Research*, *4*, 17035. https://doi.org/10.1038/hortres.2017.35

Millet, E. J., Kruijer, W., Coupel-Ledru, A., Alvarez Prado, S., Cabrera-Bosquet, L., Lacube, S., Charcosset, A., Welcker, C., van Eeuwijk, F., & Tardieu, F. (2019). Genomic prediction of maize yield across European environmental conditions. *Nature Genetics*, *51*(6), 952–956. https://doi.org/10.1038/s41588-019-0414-y

Min, E. J., Safo, S. E., & Long, Q. (2019). Penalized co-inertia analysis with applications to -omics data (O. Stegle, Ed.). *Bioinformatics*, *35*(6), 1018–1025. https://doi.org/10.1093/bioinformatics/bty726

Mira de Orduña, R. (2010). Climate change associated effects on grape and wine quality and production. *Food Research International*, *43*(7), 1844–1855. https://doi.org/10.1016/j.foodres.2010.05.001

Mondello, V., Songy, A., Battiston, E., Pinto, C., Coppin, C., Trotel-Aziz, P., Clément, C., Mugnai, L., & Fontaine, F. (2018). Grapevine Trunk Diseases: a Review of Fifteen Years of Trials for Their Control with Chemicals and Biocontrol Agents. *Plant Disease*, *102*(7)pmid 30673583, 1189–1217. https://doi.org/10.1094/PDIS-08-17-1181-FE

Montesinos-López, A., Montesinos-López, O. A., Cuevas, J., Mata-López, W. A., Burgueño, J., Mondal, S., Huerta, J., Singh, R., Autrique, E., González-Pérez, L., & Crossa, J. (2017a). Genomic Bayesian functional regression models with interactions for predicting wheat grain yield using hyper-spectral image data. *Plant Methods*, *13*(1), 62. https://doi.org/10.1186/s13007-017-0212-4

Montesinos-López, O. A., Montesinos-López, A., Crossa, J., de los Campos, G., Alvarado, G., Suchismita, M., Rutkoski, J., González-Pérez, L., & Burgueño, J. (2017b). Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods*, *13*(1), 4. https://doi.org/10.1186/s13007-016-0154-2

Morales-Castilla, I., García de Cortázar-Atauri, I., Cook, B. I., Lacombe, T., Parker, A. K., van Leeuwen, C., Nicholas, K. A., & Wolkovich, E. M. (2020). Diversity buffers winegrowing regions from climate change losses. *Proceedings of the National Academy of Sciences*, *117*(6), 2864–2869. https://doi.org/10.1073/pnas.1906731117

Morota, G., Koyama, M., M Rosa, G. J., Weigel, K. A., & Gianola, D. (2013). Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. *Genetics Selection Evolution*, *45*(1), 17. https://doi.org/10.1186/1297-9686-45-17

Mudge, K., Janick, J., Scofield, S., & Goldschmidt, E. E. (2009, October 7). A History of Grafting. In J. Janick (Ed.), *Horticultural Reviews* (pp. 437–493). Hoboken, NJ, USA, John Wiley & Sons, Inc. https://doi.org/10.1002/9780470593776.ch9

Myles, S. (2013). Improving fruit and wine: what does genomics have to offer? *Trends in Genetics*, *29*(4), 190–196. https://doi.org/10.1016/j.tig.2013.01.006

Myles, S., Boyko, A. R., Owens, C. L., Brown, P. J., Grassi, F., Aradhya, M. K., Prins, B., Reynolds, A., Chia, J.-M., Ware, D., Bustamante, C. D., & Buckler, E. S. (2011). Genetic structure and domestication history of the grape. *Proceedings of the National Academy of Sciences*, *108*(9)pmid 21245334, 3530–3535. https://doi.org/10.1073/pnas.1009363108

Naor, A., Schneider, D., Ben-Gal, A., Zipori, I., Dag, A., Kerem, Z., Birger, R., Peres, M., & Gal, Y. (2013). The effects of crop load and irrigation rate in the oil accumulation stage on oil yield and water relations of 'Koroneiki' olives. *Irrigation Science*, *31*(4), 781–791. https://doi.org/10.1007/s00271-012-0363-z

Naulleau, A., Gary, C., Prévot, L., & Hossard, L. (2021). Evaluating Strategies for Adaptation to Climate Change in Grapevine Production–A Systematic Review. *Frontiers in Plant Science*, *11*, 607859. https://doi.org/10.3389/fpls.2020.607859

Nejati-Javaremi, A., Smith, C., & Gibson, J. P. (1997). Effect of Total Allelic Relationship on Accuracy of Evaluation and Response to Selection. *Journal of Animal Science*, *75*(7), 1738–1745. https://doi.org/10.2527/1997.7571738x

Neyhart, J. L., & Smith, K. P. (2019). Validating Genomewide Predictions of Genetic Variance in a Contemporary Breeding Program. *Crop Science*, *59*(3), 1062–1072. https://doi.org/10.2135/cropsci2018.11.0716
_eprint: https://acsess.onlinelibrary.wiley.com/doi/pdf/10.2135/cropsci2018.11.0716

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biology and Technology*, *46*(2), 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024

Nicolas, S. D., Péros, J.-P., Lacombe, T., Launay, A., Le Paslier, M.-C., Bérard, A., Mangin, B., Valière, S., Martins, F., Le Cunff, L., Laucou, V., Bacilieri, R., Dereeper, A., Chatelet, P., This, P., & Doligez, A. (2016). Genetic diversity, linkage disequilibrium and power of a large grapevine (Vitis vinifera L) diversity panel newly designed for association studies. *BMC Plant Biology*, *16*(1), 74. https://doi.org/10.1186/s12870-016-0754-z

Norman, A., Taylor, J., Edwards, J., & Kuchel, H. (2018). Optimising Genomic Selection in Wheat: effect of Marker Density, Population Size and Population Structure on Prediction Accuracy. *G3: Genes, Genomes, Genetics*, *8*(9), 2889–2899. https://doi.org/10.1534/g3.118.200311

Nsibi, M., Gouble, B., Bureau, S., Flutre, T., Sauvage, C., Audergon, J.-M., & Regnard, J.-L. (2020). Adoption and Optimization of Genomic Selection To Sustain Breeding for Apricot Fruit Quality. *G3 Genes | Genomes | Genetics*, *10*(12), 4513–4529. https://doi.org/10.1534/g3.120.401452

Obozinski, G., Wainwright, M. J., & Jordan, M. I. (2011). Support union recovery in high-dimensional multivariate regression. *The Annals of Statistics*, *39*(1)arxiv 0808.0711, 1–47. https://doi.org/10.1214/09-AOS776

OIV. (2017). Distribution of the World's Grapevine Varieties. *Focus OIV 2017*.

Ojeda, H., Sanchis, F. M., Corbacho, L., Bouquet, A., & Carbonneau, A. (2010). Reduction of costs and emissions in vineyards by the use of low-input technologies adapted to productions of good quality wines: genotypes resistant to fungal diseases and minimal pruning-zero pruning system. *Progrès Agricole et Viticole*, *127*(21/22), 431–440. Retrieved June 22, 2021, from https://www.cabdirect.org/cabdirect/abstract/20113022440

Ollat, N., Bordenave, L., Tandonnet, J. P., Bourisquot, J. M., & Marguerit, E. (2016). Grapevine rootstocks: origins and perspectives. *Acta Horticulturae*, (1136), 11–22. Retrieved June 9, 2021, from https://pubag.nal.usda.gov/catalog/5643028

Ollat, N., Quénol, H., Barbeau, G., van Leeuwen, C., Darriet, P., Garcia de Cortazar Atauri, I., Bois, B., Ojeda, H., Duchêne, E., Lebon, E., Vivin, P., Torregrosa, L., Sablayrolles, J.-M., Teil, G., Lagacherie, P., Giraud-Héraud, E., Aigrain, P., & Touzard, J.-M. (2018). Adaptation to climate change of the French wine industry: a systemic approach – Main outcomes of the project LACCAVE (Sociedad Aragonesa de Gestión Agroambiental, S.L.U.(SARGA), Ed.). *E3S Web of Conferences*, *50*, 01020. https://doi.org/10.1051/e3sconf/20185001020

Olmo, D., Nieto, A., Adrover, F., Urbano, A., Beidas, O., Juan, A., Marco-Noales, E., López, M. M., Navarro, I., Monterde, A., Montes-Borrego, M., Navas-Cortés, J. A., & Landa, B. B. (2017). First Detection of Xylella Fastidiosa Infecting Cherry (Prunus Avium) and Polygala Myrtifolia Plants, in Mallorca Island, Spain. *Plant Disease*, *101*(10), 1820–1820. https://doi.org/10.1094/PDIS-04-17-0590-PDN

Osborne, B. (2006). Applications of near Infrared Spectroscopy in Quality Screening of Early-Generation Material in Cereal Breeding Programmes. *Journal of Near Infrared Spectroscopy*, *14*(2), 93–101. https://doi.org/10.1255/jnirs.595

Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, *103*(482), 681–686. https://doi.org/10.1198/016214508000000337

Parker, A. K., De Cortázar-Atauri, I., Van Leeuwen, C., & Chuine, I. (2011). General phenological model to characterise the timing of flowering and veraison of Vitis vinifera L.: grapevine flowering and veraison model. *Australian Journal of Grape and Wine Research*, *17*(2), 206–216. https://doi.org/10.1111/j.1755-0238.2011.00140.x

Parker, A. K., de Cortázar-Atauri, I. G., Chuine, I., Barbeau, G., Bois, B., Boursiquot, J.-M., Cahurel, J.-Y., Claverie, M., Dufourcq, T., Gény, L., Guimberteau, G., Hofmann, R. W., Jacquet, O., Lacombe, T., Monamy, C., Ojeda, H., Panigai, L., Payan, J.-C., Lovelle, B. R., … van Leeuwen, C. (2013). Classification of varieties for their timing of flowering and veraison using a modelling approach: a case study for the grapevine species Vitis vinifera L. *Agricultural and Forest Meteorology*, *180*, 249–264. https://doi.org/10.1016/j.agrformet.2013.06.005

Parker, A. K., García de Cortázar-Atauri, I., Gény, L., Spring, J.-L., Destrac, A., Schultz, H., Stoll, M., Molitor, D., Lacombe, T., Graca, A., Monamy, C., Storchi, P., Trought, M. C., Hofmann, R., & van Leeuwen, C. (2019). The temperature-based Grapevine Sugar Ripeness (GSR) model for adapting a wide range of Vitis vinifera L. cultivars in a changing climate. *21st GiESCO International Meeting: 'A Multidisciplinary Vision towards Sustainable Viticulture"*. Retrieved June 22, 2021, from https://hdl.handle.net/10182/12731
Accepted: 2020-09-21T21:21:26Z

Parker, A. K., García de Cortázar-Atauri, I., Gény, L., Spring, J.-L., Destrac, A., Schultz, H., Molitor, D., Lacombe, T., Graça, A., Monamy, C., Stoll, M., Storchi, P., Trought, M. C. T., Hofmann, R. W., & van Leeuwen, C. (2020). Temperature-based grapevine sugar ripeness modelling for a wide range of Vitis vinifera L. cultivars. *Agricultural and Forest Meteorology*, *285–286*, 107902. https://doi.org/10.1016/j.agrformet.2020.107902

Pauquet, J., Bouquet, A., This, P., & Adam-Blondon, A.-F. (2001). Establishment of a local map of AFLP markers around the powdery mildew resistance gene Run1 in grapevine and assessment of their usefulness for marker assisted selection: *Theoretical and Applied Genetics*, *103*(8), 1201–1210. https://doi.org/10.1007/s001220100664

Pech-Rouge, U. (2017). Variétés Résistantes: Patience et Longueur de Temps (295th ed.) [newspaper]. *La Vigne*. Retrieved August 4, 2021, from https://www1.montpellier.inra.fr/pechrouge/images/1703-la-vigne-varietes-resistantes-patience-et-longueur-de-temps.pdf

Pégard, M., Rogier, O., Bérard, A., Faivre-Rampant, P., Paslier, M.-C. L., Bastien, C., Jorge, V., & Sánchez, L. (2019). Sequence Imputation from Low Density Single Nucleotide Polymorphism Panel in a Black Poplar Breeding Population. *BMC Genomics*, *20*(1), 302. https://doi.org/10.1186/s12864-019-5660-y

Peressotti, E., Wiedemann-Merdinoglu, S., Delmotte, F., Bellin, D., Di Gaspero, G., Testolin, R., Merdinoglu, D., & Mestre, P. (2010). Breakdown of resistance to grapevine downy mildew upon limited deployment of a resistant variety. *BMC plant biology*, *10*pmid 20633270, 147. https://doi.org/10.1186/1471-2229-10-147

Pérez-Enciso, M. (2014). Genomic relationships computed from either next-generation sequence or array SNP data. *Journal of Animal Breeding and Genetics*, *131*(2), 85–96. https://doi.org/10.1111/jbg.12074
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jbg.12074

Pérez-Enciso, M., Forneris, N., de los Campos, G., & Legarra, A. (2017). Evaluating Sequence-Based Genomic Prediction with an Efficient New Simulator. *Genetics*, *205*(2), 939–953. https://doi.org/10.1534/genetics.116.194878

Péros, J.-P., Berger, G., Portemont, A., Boursiquot, J.-M., & Lacombe, T. (2011). Genetic variation and biogeography of the disjunct Vitis subg. Vitis (Vitaceae). *Journal of Biogeography*, *38*(3), 471–486. https://doi.org/10.1111/j.1365-2699.2010.02410.x
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2699.2010.02410.x

Persa, R., Ribeiro, P. C. d. O., & Jarquin, D. (2021). The use of high-throughput phenotyping in genomic selection context. *Crop Breeding and Applied Biotechnology*, *21*. https://doi.org/10.1590/1984-70332021v21Sa19

Piepho, H.-P., & Möhring, J. (2007). Computing Heritability and Selection Response From Unbalanced Plant Breeding Trials. *Genetics*, *177*(3), 1881–1888. https://doi.org/10.1534/genetics.107.074229

Pons, A., Allamy, L., Schüttler, A., Rauhut, D., Thibon, C., & Darriet, P. (2017). What is the expected impact of climate change on wine aroma compounds and their precursors in grape? *OENO One*, *51*(2), 141–146. https://doi.org/10.20870/oeno-one.2017.51.2.1868

Pou, A., Medrano, H., Tomàs, M., Martorell, S., Ribas-Carbó, M., & Flexas, J. (2012). Anisohydric behaviour in grapevines results in better performance under moderate water stress and recovery than isohydric behaviour. *Plant and Soil*, *359*(1), 335–349. https://doi.org/10.1007/s11104-012-1206-7

Rasoli, V., Farshadfar, E., & Ahmadi, J. (2015). Evaluation of Genotype × Environment Interaction of Grapevine Genotypes (Vitis vinifera L.) By Non Parametric Method. *Journal of Agricultural Science and Technology*, *17*(5), 1279–1289. Retrieved September 21, 2021, from http://jast.modares.ac.ir/article-23-1637-en.html

Rastas, P. (2017). Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics*, *33*(23), 3726–3732. https://doi.org/10.1093/bioinformatics/btx494

Resende, R. T., Resende, M. D. V., Silva, F. F., Azevedo, C. F., Takahashi, E. K., Silva-Junior, O. B., & Grattapaglia, D. (2017). Assessing the expected response to genomic selection of individuals and families in Eucalyptus breeding with an additive-dominant model. *Heredity*, *119*(4), 245–255. https://doi.org/10.1038/hdy.2017.37

Rex, F., Fechter, I., Hausmann, L., & Töpfer, R. (2014). QTL mapping of black rot (Guignardia bidwellii) resistance in the grapevine rootstock 'Börner' (V. riparia Gm183 × V. cinerea Arnold). *Theoretical and Applied Genetics*, *127*(7), 1667–1677. https://doi.org/10.1007/s00122-014-2329-4

Reynolds, A. (2015). *Grapevine Breeding Programs for the Wine Industry*. Elsevier Science. Retrieved June 11, 2021, from http://www.totalboox.com/book/id-5271892355917935666
OCLC: 969077965

Riaz, S., Boursiquot, J.-M., Dangl, G. S., Lacombe, T., Laucou, V., Tenscher, A. C., & Walker, M. A. (2013). Identification of mildew resistance in wild and cultivated Central Asian grape germplasm. *BMC Plant Biology*, *13*(1), 149. https://doi.org/10.1186/1471-2229-13-149

Riaz, S., Tenscher, A. C., Graziani, R., Krivanek, A. F., Ramming, D. W., & Walker, M. A. (2009). Using Marker-Assisted Selection to Breed Pierce's Disease-Resistant Grapes. *American Journal of Enology and Viticulture*, *60*(2), 199–207. Retrieved July 29, 2021, from https://www.ajevonline.org/content/60/2/199

Riedelsheimer, C., Technow, F., & Melchinger, A. E. (2012). Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. *BMC Genomics*, *13*(1), 452. https://doi.org/10.1186/1471-2164-13-452

Rincent, R., Laloë, D., Nicolas, S., Altmann, T., Brunel, D., Revilla, P., Rodríguez, V. M., Moreno-Gonzalez, J., Melchinger, A., Bauer, E., Schoen, C.-C., Meyer, N., Giauffret, C., Bauland, C., Jamin, P., Laborde, J., Monod, H., Flament, P., Charcosset, A., & Moreau, L. (2012). Maximizing the Reliability of Genomic Selection by Optimizing the Calibration Set of Reference Individuals: comparison of Methods in Two Diverse Groups of Maize Inbreds (Zea mays L.) *Genetics*, *192*(2)pmid 22865733, 715–728. https://doi.org/10.1534/genetics.112.141473

Rincent, R., Charpentier, J.-P., Faivre-Rampant, P., Paux, E., Le Gouis, J., Bastien, C., & Segura, V. (2018). Phenomic Selection Is a Low-Cost and High-Throughput Method Based on Indirect Predictions: proof of Concept on Wheat and Poplar. *G3 Genes | Genomes | Genetics*, g3.200760.2018. https://doi.org/10.1534/g3.118.200760

Rio, S., Mary-Huard, T., Moreau, L., & Charcosset, A. (2019). Genomic selection efficiency and a priori estimation of accuracy in a structured dent maize panel. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, *132*(1)pmid 30288553, 81–96. https://doi.org/10.1007/s00122-018-3196-1

Roach, M. J., Johnson, D. L., Bohlmann, J., van Vuuren, H. J. J., Jones, S. J. M., Pretorius, I. S., Schmidt, S. A., & Borneman, A. R. (2018). Population sequencing reveals clonal diversity and ancestral inbreeding in the grapevine cultivar Chardonnay. *PLOS Genetics*, *14*(11), e1007807. https://doi.org/10.1371/journal.pgen.1007807

Rogiers, S. Y., Coetzee, Z. A., Walker, R. R., Deloire, A., & Tyerman, S. D. (2017). Potassium in the Grape (Vitis vinifera L.) Berry: transport and Function. *Frontiers in Plant Science*, *8*. https://doi.org/10.3389/fpls.2017.01629

Roth, M., Muranty, H., Di Guardo, M., Guerra, W., Patocchi, A., & Costa, F. (2020). Genomic prediction of fruit texture and training population optimization towards the application of genomic selection in apple. *Horticulture Research*, *7*(1), 148. https://doi.org/10.1038/s41438-020-00370-5

Ruffner, H. P., Hawker, J. S., & Hale, C. R. (1976). Temperature and enzymic control of malate metabolism in berries of Vitis vinifera. *Phytochemistry*, *15*(12), 1877–1880. https://doi.org/10.1016/S0031-9422(00)88835-4

Runcie, D., & Cheng, H. (2019). Pitfalls and Remedies for Cross Validation with Multi-trait Genomic Prediction Methods. *G3 Genes | Genomes | Genetics*pmid 31511297, g3.400598.2019. https://doi.org/10.1534/g3.119.400598

Runcie, D. E., Qu, J., Cheng, H., & Crawford, L. (2021). MegaLMM: Mega-Scale Linear Mixed Models for Genomic Predictions with Thousands of Traits. *Genome Biology*, *22*(1), 213. https://doi.org/10.1186/s13059-021-02416-w

Rustioni, L., Cola, G., Maghradze, D., Abashidze, E., Argiriou, A., Aroutiounian, R., Brazão, J., Chipashvili, R., Cocco, M., Cornea, V., Dejeu, L., Eiras-Dias, J. E., Goryslavets, S.,

Ibáñez, J., Kocsis, L., Lorenzini, F., Maletic, E., Mamasakhlisashvili, L., Margaryan, K., … Bacilieri, R. (2019). Description of the Vitis vinifera L. phenotypic variability in eno-carpological traits by a Euro-Asiatic collaborative network among ampelographic collections. *VITIS - Journal of Grapevine Research*, 37–46 Pages. https://doi.org/10.5073/VITIS.2019.58.37-46

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L. G., Crossa, J., Reynolds, M., & Singh, R. (2016). Canopy Temperature and Vegetation Indices from High-Throughput Phenotyping Improve Accuracy of Pedigree and Genomic Selection for Grain Yield in Wheat. *G3 Genes | Genomes | Genetics*, *6*(9), 2799–2808. https://doi.org/10.1534/g3.116.032888

Sade, N., Gebremedhin, A., & Moshelion, M. (2012). Risk-Taking Plants. *Plant Signaling & Behavior*, *7*(7)pmid 22751307, 767–770. https://doi.org/10.4161/psb.20505
_eprint: https://doi.org/10.4161/psb.20505

Sadras, V. O., & Moran, M. A. (2012). Elevated temperature decouples anthocyanins and sugars in berries of Shiraz and Cabernet Franc. *Australian Journal of Grape and Wine Research*, *18*(2), 115–122. https://doi.org/10.1111/j.1755-0238.2012.00180.x
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1755-0238.2012.00180.x

Salmon, J.-M., Ojeda, H., & Escudier, J.-L. (2018). Disease resistant grapevine varieties and quality: the case of Bouquet varieties. *OENO One*, *52*(3), 225–230. https://doi.org/10.20870/oeno-one.2018.52.3.2139

Santos, J. A., Fraga, H., Malheiro, A. C., Moutinho-Pereira, J., Dinis, L.-T., Correia, C., Moriondo, M., Leolini, L., Dibari, C., Costafreda-Aumedes, S., Kartschall, T., Menz, C., Molitor, D., Junk, J., Beyer, M., & Schultz, H. R. (2020). A Review of the Potential Climate Change Impacts and Adaptation Options for European Viticulture. *Applied Sciences*, *10*(9), 3092. https://doi.org/10.3390/app10093092

Saponari, M., Boscia, D., Nigro, F., & Martelli, G. P. (2013). Identification of DNA sequences related to Xylella fastidiosa in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (southern Italy). *Journal of Plant Pathology*, *95*(3). Retrieved August 4, 2021, from https://www.cabdirect.org/cabdirect/abstract/20153279019

Sargolzaei, M., Maddalena, G., Bitsadze, N., Maghradze, D., Bianco, P. A., Failla, O., Toffolatti, S. L., & De Lorenzis, G. (2020). Rpv29, Rpv30 and Rpv31: three Novel Genomic Loci Associated With Resistance to Plasmopara viticola in Vitis vinifera. *Frontiers in Plant Science*, *11*pmid 33163011, 562432. https://doi.org/10.3389/fpls.2020.562432

Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, *15*(1), 478. https://doi.org/10.1186/1471-2164-15-478

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, *36*, 1627–1639. Retrieved August 25, 2021, from https://ui.adsabs.harvard.edu/abs/1964AnaCh..36.1627S
ADS Bibcode: 1964AnaCh..36.1627S

Savoi, S., Torregrosa, L., & Romieu, C. (2021). Transcripts switched off at the stop of phloem unloading highlight the energy efficiency of sugar import in the ripening V. vinifera fruit. *Horticulture Research*, *8*(1), 1–15. https://doi.org/10.1038/s41438-021-00628-6 Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Primary_atype: Research Subject_term: Metabolism;Plant development;Transcriptomics Subject_term_id: metabolism;plant-development;transcriptomics

Schneider, C., Onimus, C., Prado, E., Dumas, V., Wiedemann-Merdinoglu, S., Dorne, M., Lacombe, M., Piron, M., Umar-Faruk, A., Duchêne, E., Mestre, P., & Merdinoglu, D. (2019a). INRA-ResDur: the French grapevine breeding programme for durable resistance to downy and powdery mildew. *Acta Horticulturae*, (1248), 207–214. https://doi.org/10.17660/ActaHortic.2019.1248.30

Schneider, C., Spring, J.-L., Onimus, C., Prado, E., Verdenal, T., Lemarquis, G., Lorenzini, F., Ley, L., Duruz, P., Gindro, K., & Merdinoglu, D. (2019b). Programme de Collaboration Franco-Suisse Pour La Création de Nouvelles Variétés de Vigne Durablement Résistantes Au Mildiou et à l'oïdium (P. Roca, Ed.). *BIO Web of Conferences*, *15*, 01018. https://doi.org/10.1051/bioconf/20191501018

Schrag, T. A., Westhues, M., Schipprack, W., Seifert, F., Thiemann, A., Scholten, S., & Melchinger, A. E. (2018). Beyond Genomic Prediction: Combining Different Types of Omics Data Can Improve Prediction of Hybrid Performance in Maize. *Genetics*, *208*(4), 1373–1385. https://doi.org/10.1534/genetics.117.300374

Schulthess, A. W., Zhao, Y., Longin, C. F. H., & Reif, J. C. (2018). Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (Triticum aestivum L.) *Theoretical and Applied Genetics*, *131*(3), 685–701. https://doi.org/10.1007/s00122-017-3029-7

Schultz, H. R. (2003). Differences in hydraulic architecture account for near-isohydric and anisohydric behaviour of two field-grown Vitis vinifera L. cultivars during drought. *Plant, Cell & Environment*, *26*(8), 1393–1405. https://doi.org/10.1046/j.1365-3040.2003.01064.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-3040.2003.01064.x

Scutari, M., Mackay, I., & Balding, D. (2016). Using Genetic Distance to Infer the Accuracy of Genomic Prediction. *PLoS Genetics*, *12*(9)pmid 27589268. https://doi.org/10.1371/journal.pgen.1006288

Shahood, R., Torregrosa, L., Savoi, S., & Romieu, C. (2020). First quantitative assessment of growth, sugar accumulation and malate breakdown in a single ripening berry. *OENO One*, *54*(4), 1077–1092. https://doi.org/10.20870/oeno-one.2020.54.4.3787

Simonneau, T., Lebon, E., Coupel-Ledru, A., Marguerit, E., Rossdeutsch, L., & Ollat, N. (2017). Adapting plant material to face water stress in vineyards: which physiological targets for an optimal control of plant water status? *OENO One*, *51*(2), 167–179. https://doi.org/10.20870/oeno-one.2017.51.2.1870

Solberg, T. R., Sonesson, A. K., Woolliams, J. A., & Meuwissen, T. H. E. (2008). Genomic Selection Using Different Marker Types and Densities. *Journal of Animal Science*, *86*(10), 2447–2454. https://doi.org/10.2527/jas.2007-0010

Speed, D., & Balding, D. J. (2015). Relatedness in the post-genomic era: is it still useful? *Nature Reviews Genetics*, *16*(1), 33–44. https://doi.org/10.1038/nrg3821
Bandiera_abtest: a Cg_type: Nature Research Journals Primary_atype: Reviews Subject_term: Genetics;Genomic analysis Subject_term_id: genetics;genomic-analysis

Stevenson, I. (1980). The diffusion of disaster: the phylloxera outbreak in the département of the Hérault, 1862–1880. *Journal of Historical Geography*, *6*(1), 47–63. https://doi.org/10.1016/0305-7488(80)90043-2

Sun, X., Ma, P., & Mumm, R. H. (2012). Nonparametric Method for Genomics-Based Prediction of Performance of Quantitative Traits Involving Epistasis in Plant Breeding. *PLOS ONE*, *7*(11), e50604. https://doi.org/10.1371/journal.pone.0050604

Swarts, K., Li, H., Romero Navarro, J. A., An, D., Romay, M. C., Hearne, S., Acharya, C., Glaubitz, J. C., Mitchell, S., Elshire, R. J., Buckler, E. S., & Bradbury, P. J. (2014). Novel Methods to Optimize Genotypic Imputation for Low-Coverage, Next-Generation Sequence Data in Crop Plants. *The Plant Genome*, *7*(3). https://doi.org/10.3835/plantgenome2014.05.0023

Tardieu, F. (2012). Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. *Journal of Experimental Botany*, *63*(1), 25–31. https://doi.org/10.1093/jxb/err269

Tardieu, F., & Simonneau, T. (1998). Variability among Species of Stomatal Control under Fluctuating Soil Water Status and Evaporative Demand: Modelling Isohydric and Anisohydric Behaviours. *Journal of Experimental Botany*, *49*jstor 23695975, 419–432.

Tardieu, F., Simonneau, T., & Muller, B. (2018). The Physiological Basis of Drought Tolerance in Crop Plants: a Scenario-Dependent Probabilistic Approach. *Annual Review of Plant Biology*, *69*(1), 733–759. https://doi.org/10.1146/annurev-arplant-042817-040218

Tayeh, N., Klein, A., Le Paslier, M.-C., Jacquin, F., Houtin, H., Rond, C., Chabert-Martinello, M., Magnin-Robert, J.-B., Marget, P., Aubert, G., & Burstin, J. (2015). Genomic Prediction in Pea: effect of Marker Density and Training Population Size and Composition on Prediction Accuracy. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00941

Tello, J., Roux, C., Chouiki, H., Laucou, V., Sarah, G., Weber, A., Santoni, S., Flutre, T., Pons, T., This, P., Péros, J.-P., & Doligez, A. (2019). A novel high-density grapevine (Vitis vinifera L.) integrated linkage map using GBS in a half-diallel population. *Theoretical and Applied Genetics*, *132*(8), 2237–2252. https://doi.org/10.1007/s00122-019-03351-y

Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., Angiuoli, S. V., Crabtree, J., Jones, A. L., Durkin, A. S., DeBoy, R. T., Davidsen, T. M., Mora, M., Scarselli, M., y Ros, I. M., Peterson, J. D., Hauser, C. R., Sundaram, J. P., Nelson, W. C., … Fraser, C. M. (2005). Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". *Proceedings of*

the National Academy of Sciences, *102*(39)pmid 16172379, 13950–13955. https://doi.org/10.1073/pnas.0506758102

This, P., Boursiquot, J.-M., Laucou, V., Peros, J.-P., & Lacombe, T. (2021). Domestication et diversification des plantes: Le cas de la Vigne. *Jardins de France*, *661*, 42. Retrieved August 23, 2021, from https://hal.inrae.fr/hal-03272430

This, P., Lacombe, T., & Thomas, M. R. (2006). Historical origins and genetic diversity of wine grapes. *Trends in Genetics*, *22*(9), 511–519. https://doi.org/10.1016/j.tig.2006.07.008

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1)jstor 2346178, 267–288.

Töpfer, R., Hausmann, L., Harst, M., Maul, E., Zyprian, E., & Eibach, R. (2010). New Horizons for Grapevine Breeding. *New Horizons for Grapevine Breeding*, (1), 22.

Toro, M., García-Cortés, L., & Legarra, A. (2011). A note on the rationale for estimating genealogical coancestry from molecular markers. *Genetics Selection Evolution*, *43*(1), 27. https://doi.org/10.1186/1297-9686-43-27

Trenti, M., Lorenzi, S., Bianchedi, P. L., Grossi, D., Failla, O., Grando, M. S., & Emanuelli, F. (2021). Candidate genes and SNPs associated with stomatal conductance under drought stress in Vitis. *BMC Plant Biology*, *21*(1), 7. https://doi.org/10.1186/s12870-020-02739-z

Tripiana, V., & Mayoux, L. (2018). Étude sur l'âge du vignoble et des vignes arrachées dans le bassin viticole Languedoc-Roussillon. *Agreste Occitanie*, (8), 6. Retrieved August 24, 2021, from http://www.epsilon.insee.fr:80/jspui/handle/1/82737
Accepted: 2018-10-10T08:22:45Z

Tsuchikawa, S., & Kobori, H. (2015). A review of recent application of near infrared spectroscopy to wood science and technology. *Journal of Wood Science*, *61*(3), 213–220. https://doi.org/10.1007/s10086-015-1467-x

Ubbens, J., Parkin, I., Eynck, C., Stavness, I., & Sharpe, A. G. (2021). Deep neural networks for genomic prediction do not estimate marker effects. *The Plant Genome*pmid 34596363, e20147. https://doi.org/10.1002/tpg2.20147

van Leeuwen, C., & Destrac-Irvine, A. (2017). Modified grape composition under climate change conditions requires adaptations in the vineyard. *OENO One*, *51*(2), 147–154. https://doi.org/10.20870/oeno-one.2017.51.2.1647

van Leeuwen, C., Destrac-Irvine, A., Dubernet, M., Duchêne, E., Gowdy, M., Marguerit, E., Pieri, P., Parker, A. K., de Rességuier, L., & Ollat, N. (2019). An Update on the Impact of Climate Change in Viticulture and Potential Adaptations. *Agronomy*, *9*(9), 514. https://doi.org/10.3390/agronomy9090514

VanRaden, P. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Varona, L., Legarra, A., Toro, M. A., & Vitezica, Z. G. (2018). Non-additive Effects in Genomic Selection. *Frontiers in Genetics*, *9*. https://doi.org/10.3389/fgene.2018.00078

Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D. A., Cestaro, A., Pruss, D., Pindo, M., FitzGerald, L. M., Vezzulli, S., Reid, J., Malacarne, G., Iliev, D., Coppola, G., Wardell,

B., Micheletti, D., Macalma, T., Facci, M., Mitchell, J. T., Perazzolli, M., … Viola, R. (2007). A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety. *PLOS ONE*, *2*(12), e1326. https://doi.org/10.1371/journal.pone.0001326

Vezzulli, S., Doligez, A., & Bellin, D. (2019). Molecular Mapping of Grapevine Genes. In D. Cantu & M. A. Walker (Eds.), *The Grape Genome* (pp. 103–136). Cham, Springer International Publishing. https://doi.org/10.1007/978-3-030-18601-2_7

Viana, A. P., de Resende, M. D. V., Riaz, S., & Walker, M. A. (2016). Genome selection in fruit breeding: application to table grapes. *Scientia Agricola*, *73*(2), 142–149. https://doi.org/10.1590/0103-9016-2014-0323

Vitezica, Z. G., Varona, L., & Legarra, A. (2013). On the Additive and Dominant Variance and Covariance of Individuals Within the Genomic Selection Scope. *Genetics*, *195*(4), 1223–1230. https://doi.org/10.1534/genetics.113.155176

Vitezica, Z. G., Legarra, A., Toro, M. A., & Varona, L. (2017). Orthogonal Estimates of Variances for Additive, Dominance, and Epistatic Effects in Populations. *Genetics*, *206*(3), 1297–1307. https://doi.org/10.1534/genetics.116.199406

Vitezica, Z. G., Reverter, A., Herring, W., & Legarra, A. (2018). Dominance and epistatic genetic variances for litter size in pigs using genomic models. *Genetics Selection Evolution*, *50*(1). https://doi.org/10.1186/s12711-018-0437-3

Vivin, P., Lebon, É., Dai, Z., Duchêne, E., Marguerit, E., de Cortázar-Atauri, I. G., Zhu, J., Simonneau, T., van Leeuwen, C., Delrot, S., & Ollat, N. (2017). Combining ecophysiological models and genetic analysis: a promising way to dissect complex adaptive traits in grapevine. *OENO One*, *51*(2), 181–189. https://doi.org/10.20870/oeno-one.2017.51.2.1588

Wade, A. R., Durufle, H., Sanchez, L., & Segura, V. (2021). eQTLs are key players in the integration of genomic and transcriptomic data for phenotype prediction. *bioRxiv*, 2021.09.07.459279. https://doi.org/10.1101/2021.09.07.459279

Wang, X., Yang, Z., & Xu, C. (2015). A comparison of genomic selection methods for breeding value prediction. *Science Bulletin*, *60*(10), 925–935. https://doi.org/10.1007/s11434-015-0791-2

Wellenreuther, M., Mérot, C., Berdan, E., & Bernatchez, L. (2019). Going beyond SNPs: the role of structural genomic variants in adaptive evolution and species diversification. *Molecular Ecology*, *28*(6)pmid 30834648, 1203–1209. https://doi.org/10.1111/mec.15066

Werner, C. R., Gaynor, R. C., Sargent, D. J., Lillo, A., Gorjanc, G., & Hickey, J. M. (2020). Genomic selection strategies for clonally propagated crops. *bioRxiv*. https://doi.org/10.1101/2020.06.15.152017

Westhues, M., Schrag, T. A., Heuer, C., Thaller, G., Utz, H. F., Schipprack, W., Thiemann, A., Seifert, F., Ehret, A., Schlereth, A., Stitt, M., Nikoloski, Z., Willmitzer, L., Schön, C. C.,

Scholten, S., & Melchinger, A. E. (2017). Omics-based hybrid prediction in maize. *Theoretical and Applied Genetics*, *130*(9), 1927–1939. https://doi.org/10.1007/s00122-017-2934-0

Whalen, A., Gaynor, C., & Hickey, J. M. (2020, March 2). *Using high-throughput phenotypes to enable genomic selection by inferring genotypes* (preprint). Genetics. https://doi.org/10.1101/2020.02.28.969600

Wolfe, M. D., Chan, A. W., Kulakow, P., Rabbi, I., & Jannink, J.-.-L. (2021). Genomic mating in outbred species: predicting cross usefulness with additive and total genetic covariance matrices. *bioRxiv*, 2021.01.05.425443. https://doi.org/10.1101/2021.01.05.425443

Won, S., Park, J.-E., Son, J.-H., Lee, S.-H., Park, B. H., Park, M., Park, W.-C., Chai, H.-H., Kim, H., Lee, J., & Lim, D. (2020). Genomic Prediction Accuracy Using Haplotypes Defined by Size and Hierarchical Clustering Based on Linkage Disequilibrium. *Frontiers in Genetics*, *11*, 134. https://doi.org/10.3389/fgene.2020.00134

Xu, K., Riaz, S., Roncoroni, N. C., Jin, Y., Hu, R., Zhou, R., & Walker, M. A. (2008). Genetic and QTL analysis of resistance to Xiphinema index in a grapevine cross. *Theoretical and Applied Genetics*, *116*(2), 305–311. https://doi.org/10.1007/s00122-007-0670-6

Xu, S. (2003). Theoretical Basis of the Beavis Effect. *Genetics*, 10. https://doi.org/165:2259âĂŞ2268

Xu, S., Xu, Y., Gong, L., & Zhang, Q. (2016). Metabolomic prediction of yield in hybrid rice. *The Plant Journal*, *88*(2), 219–227. https://doi.org/10.1111/tpj.13242
_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/tpj.13242

Yamamoto, E., Kataoka, S., Shirasawa, K., Noguchi, Y., & Isobe, S. (2021). Genomic Selection for F1 Hybrid Breeding in Strawberry (Fragaria × ananassa). *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.645111

Yan, W., Kang, M. S., Ma, B., Woods, S., & Cornelius, P. L. (2007). GGE Biplot vs. AMMI Analysis of Genotype-by-Environment Data. *Crop Science*, *47*(2), 643–653. https://doi.org/10.2135/cropsci2006.06.0374
_eprint: https://acsess.onlinelibrary.wiley.com/doi/pdf/10.2135/cropsci2006.06.0374

Zhang, A., Wang, H., Beyene, Y., Semagn, K., Liu, Y., Cao, S., Cui, Z., Ruan, Y., Burgueño, J., San Vicente, F., Olsen, M., Prasanna, B. M., Crossa, J., Yu, H., & Zhang, X. (2017). Effect of Trait Heritability, Training Population Size and Marker Density on Genomic Prediction Accuracy Estimation in 22 Bi-Parental Tropical Maize Populations. *Frontiers in Plant Science*, *8*, 1916. https://doi.org/10.3389/fpls.2017.01916

Zhang, J., Hausmann, L., Eibach, R., Welter, L. J., Töpfer, R., & Zyprian, E. M. (2009). A framework map from grapevine V3125 (Vitis vinifera 'Schiava grossa' × 'Riesling') × rootstock cultivar 'Börner' (Vitis riparia × Vitis cinerea) to localize genetic determinants of phylloxera root resistance. *Theoretical and Applied Genetics*, *119*(6), 1039–1051. https://doi.org/10.1007/s00122-009-1107-1

Zhong, S., & Jannink, J.-L. (2007). Using Quantitative Trait Loci Results to Discriminate Among Crosses on the Basis of Their Progeny Mean and Variance. *Genetics*, *177*(1), 567–576. https://doi.org/10.1534/genetics.107.075358

Zhou, Y., Minio, A., Massonnet, M., Solares, E., Lv, Y., Beridze, T., Cantu, D., & Gaut, B. S. (2019). The population genetics of structural variants in grapevine domestication. *Nature Plants*, *5*(9), 965–979. https://doi.org/10.1038/s41477-019-0507-8

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Appendix A

Appendix Chapter 1

A.1 Book chapter phenomic selection

This book chapter has been written by Pauline Robert*, Charlotte Brault*, Vincent Segura and Renaud Rincent. It has been accepted for publication in a book entitled "Genomic prediction of complex traits", edited by Springer Nature, series Methods in Molecular Biology.

* equal contribution of these authors

Phenomic selection: a new and efficient alternative to genomic selection

Abstract

Recently, Rincent et al. ***(1)*** proposed to switch molecular markers to near-infrared (NIR) spectra for inferring relationships between individuals and further performing phenomic selection (PS), analogous to genomic selection (GS). The PS concept is similar to genomic-like omics-based (GLOB) selection, in which molecular markers are replaced by endophenotypes, such as metabolites or transcript levels, except that the phenomic information obtained for instance by near-infrared spectroscopy (NIRS) has usually a much lower cost than other omics. Though NIRS has been routinely used in breeding for several decades, especially to deal with end-product quality traits, its use to predict other traits of interest and further make selections is new. Since the seminal paper on PS, several publications have advocated the use of spectral acquisition (including NIRS and hyperspectral imaging) in plant breeding towards PS, potentially providing a scope of what is possible In the present chapter, we first come back to the concept of PS as originally proposed and provide a classification of selected papers related to the use of phenomics in breeding. We further provide a review of the selected literature concerning the type of technology used, the preprocessing of the spectra, and the statistical modelling to make predictions. We discuss the factors that likely affect the efficiency of PS and compare it to GS in terms of predictive ability. Finally, we propose several prospects for future work and application of PS in the context of plant breeding.

Key Words: plant breeding, phenomic selection (PS), genomic-like omics-based (GLOB) selection, genomic selection (GS), near infrared spectroscopy (NIRS), hyperspectral imaging

Introduction: the concept of phenomic selection and its relationship with other uses of spectra in breeding

In their recent publication, Rincent et al. ***(1)*** proposed to replace genomic information by phenomic information, such as near infrared spectra, to predict quantitative traits and further perform what they coined “phenomic selection” (PS). The use of spectroscopy in agriculture and plant breeding is not novel, however, its use as an alternative to molecular markers to build relationship matrices and further predict individual performances in the context of selection is new. PS is similar to genomic-like omics-based (GLOB) selection, in which molecular markers are replaced by endophenotypes such as transcriptomics, metabolomics, or any other omics ***(2-8)***. Endophenotypes are generally measured once and for all in controlled conditions, and used to build a predictive model for field agronomic traits. This approach is thus similar to genomic selection (GS) with genotyping in the lab replaced by endophenotypic characterization in controlled conditions. In these publications, models such as G-BLUP or multi-BLUP were generally used with the kinship matrix replaced by relationship matrices estimated with the omics data, and the predictive abilities obtained were generally similar and sometimes higher than those obtained with GS. Even if endophenotypic characterization remains costly, improved efficiency to capture non additive effects (epistasis, genotype by environment interaction - GEI) can be highly valuable. To decrease and scale both cost and throughput, Rincent et al. ***(1)*** proposed to replace genotyping or omics characterisation by NIRS. They illustrated that NIR sectra were indeed able to capture genetic similarities, and thus resulted in accurate predictions, even for traits unrelated to the tissue on which NIR spectra were measured. A GLOB selection approach based on NIR spectra acquired in one given environment was for instance able to accurately predict yield in other environments, as long as the calibration set was phenotyped in these environments. Note, that PS, and in particular GLOB selection, is radically different from the classical NIRS use. In the classical use, NIRS predicts the chemical

composition of the analyzed tissue. In PS and GLOB selection, NIR spectra (or other phenomics data) are used to capture the genetic similarities between the genotypes, which allows accurate predictions of any polygenic trait.

Spectroscopy techniques, such as NIRS, measure the emission or reflection of light on a sample for a given wavenumber range, e.g. for NIR from 780 nm to 2500 nm. Various chemical bonds absorb light at different wavelengths, and this can be used in a quantitative manner. NIRS provides a non-destructive and high-throughput measurement of living samples (where water absorbance bands do not overlap) as well as dried or crushed tissues. Absorption or reflectance values at a given wavelength are proportional to molecule concentrations, as depicted in the Beer-Lambert law. Thus, spectra variations are due to the combination of molecules in the tissue and their respective absorption bands. This chemical property has been widely used in agriculture and forestry in many species to predict traits of interest, such as those related to grain composition for cereals ***(9)***, milk composition for dairy cattle ***(10)***, wood properties for forest trees ***(11)***, quality traits for fruits and vegetables, ***(12)***.

Apart from this classical use of NIRS, which can be exploited in breeding to make selections on traits related to end-product composition, other use of NIRS in breeding have been reported in the litterature, which can be classified into four main categories as proposed in **Table 1**: NIR spectra as proxy of the target trait, NIR spectra as a secondary trait, PS within environments and PS across environments as a particular case of GLOB selection. In the following paragraphs, we provide a definition of each of these four categories.

The first category concerns the use of NIR spectraas a proxy of complex traits such as grain yield, water stress or chlorophyll content with vegetation indices (VIs), which are based on a few wavelength bands. Several indices were successively developed across years, the most famous

one being NDVI (normalized difference vegetation index). These indices are strongly correlated with photosynthetic activity and sometimes with yield ***(13)*** and are still widely used to reflect biomass or yield of the analyzed plants or plots ***(14-16,24,26)***. The advantage here of using NIR spectra as a proxy trait is to avoid expensive phenotyping and to enable indirect selection of target traits before harvest.

In the second category, we have gathered studies that combined NIRS measurements with molecular markers to increase the accuracy of genomic prediction. In that case, NIR spectrum was considered as a secondary trait to be associated with the target trait in a multivariate prediction model ***(17-19)***. It is worth mentioning that such studies report two distinct examples of application: one based on canopy reflectance **(17, 19)**, which relates to the overall plant health ; and the other based on grain NIRS **(18)**, which relates to the end product or cumulative energy accumulation of the plant. In any case, this approach is particularly valuable when the predicted set is phenotyped with NIRS (trait-assisted prediction). Other authors have used NIR measurements to specifically account for GEI in the genomic prediction model ***(20, 21)***, but this requires NIRS data for each environment to be measured.

The third category includes studies that have investigated using NIR spectra to build a (hyper)spectral relationship matrix between plots/individuals, referred to as H matrix, and integrated it in the prediction equation with or without the genomic relationship matrix (kinship), referred to as G matrix ***(20-22, 25)***. The NIRS-based similarity between two plots/individuals $H(i,j)$ can be estimated with the following formula:

$$H(i,j)=\frac{\sum_{k=1}^{n_w}\left[S(i,k)\times S(j,k)\right]}{n_w},$$

with $S(i,k)$ the preprocessed absorbance or reflectance (see hereafter) measured on the *i*th plot/individual for the *k*th wavelength. *S* is centered and scaled for each wavelength and has the

dimension *n* (number of plots/individuals) times n_w (number of wavelengths). In matrix notation $H = \frac{SS'}{n_w}$. This PS approach resulted in promising accuracies,

however, in these publications, predictions were still made for a specific site-year trial, and so NIR spectra have to be measured on each plot in each environment, with the strong limitation that all varieties have to be grown in each environment.

Thus, we propose here a fourth category to further make a distinction between PS applied to the plots on which NIR spectra were acquired (previous category) and PS predictions across-environments with NIR spectra measured in a reference site only, a particular case of what we called GLOB selection (scenario S2 in Rincent et al. **(1)** and **Figure 1**). In GLOB selection, we suppose that NIRS captures genetic similarities, which means that it is sufficient to acquire the spectra once and for all in one single environment. The derived H similarity matrix is then used in replacement of the G matrix in the classical GS models.

To date, we have found only two publications (Rincent et al., Lane et al. ***(23))***, that performed GLOB selection with NIRS , and they both resulted in accurate predictions, often more accurate than GS and with dramatically reduced costs. These publications and the results obtained with GLOB selection obtained with other kinds of omics prove that variations other than DNA markers can efficiently capture genetic similarities between genotypes and result in accurate predictions.

Literature review on the use of spectra in selection

Following previous definitions, we have selected a number of papers illustrating each of the four categories of use of NIRS in the context of breeding . The main features of these papers are summarized in **Table 1**, and further presented and detailed hereafter, with respect to the type of technology employed to obtain spectra, the statistical pretreatments of the spectra and the statistical model applied for phenotype prediction. We also provide in **Table 2** a comparison of the relative performances of phenomic and genomic predictions for the very few papers which enable such a comparison. Finally, we discuss the factors that affect the predictive ability of PS.

A. Types of technology

Traditionally, NIR measurements are conducted in laboratories under controlled conditions for either dried vegetative tissue (e.g. forages) or dried reproductive tissue (e.g. grain). This kind of data displays many advantages: measurements are robust, low-cost and routinely applied by breeders to predict quality traits. There are also a number of disadvantages, there is substantial extra effort needed to bring these materials from the field to the lab, and to dry them so that water absorbance (which overlaps other chemical bond absorbance) is minimized. In these laboratory conditions, spectra are constituted of many wavelengths possibly from the visible and near infrared (400 to 2500 nm approximately), constituting a dataset of hundreds of variables ***(9)***.

With the rise of high-throughput phenotyping, spectrum measurements have benefited from technological developments which enables the direct collection of spectra in the field possibly at several time points, like hyperspectral imaging from Unoccupied Aerial Vehicle (UAV) or direct measurements of fresh material with portable (micro-)spectrometers. Hyperspectral imaging takes images with several wavelengths for each pixel, possibly at multiple time points in the visible and in a small portion of the NIR spectrum. A reflectance measure is attributed to some groups of wavelengths (bands) or to individual wavelengths directly. The measurements at the pixel level can be integrated at the microplot level to characterize a unique variety ***(20, 21, 24,***

25). Portable (micro-)spectrometers have also been developed to measure the reflectance directly in fields on undestroyed fresh material covering the visible and NIR spectrum ***(26)***. Wavelengths can be used directly as variables in predictive models or they can be derived in several indexes, like VIs ***(20, 24)***. VIs describe vegetation properties by summarising the information of large amounts of data to facilitate processing of camera and satellite images. However, in Aguate et al. ***(24)*** the use of all the hyperspectral bands achieved better prediction than using VI individually.

Technologies used to collect NIR spectrum are numerous, each with advantages and disadvantages. On the one hand, the use of NIR spectrometers in laboratory conditions is a robust method but can be time consuming due to collection and possibly preparation of samples. On the other hand, UAV and portable (micro-)spectrometers are quick techniques to collect NIRS but the number of wavelengths available is usually reduced and measurements can be affected by environmental noise which, is harder to control in the field than in the laboratory. Depending on the application, trade-offs must be found between labor intensity, costs and spectrum quality.

To date, very few studies have tried to compare the predictive ability of different spectrum measurements, especially in the context of plant breeding. Recently, Zgouz et al. ***(27)*** have reported a dataset of spectra collected on 60 sugarcane samples with 8 visible/NIR spectrometers including handheld micro-spectrometers. Such a dataset is very useful to compare different tools, although results might be context dependent, *i.e.* the most accurate model for different traits and species might be obtained with different spectrometers. Still, quantifying the gain or the loss of predictive ability for each technique will be helpful to guide in using one technique rather than another for a specific objective. Other techniques could be used to facilitate measurement, for example to combine hyperspectral images and laboratory spectra. Instead of using a spectrometer to measure samples one by one, hyperspectral images can

measure several samples at the same time. This would enhance robustness of spectra collection and reduce time of measurements. Beyond technical issues, it is also important to consider practical organizational questions, such as the period at which spectra are measured, to make sure that the predictions are available before the sowing of the next season.

B. Preprocessing NIR spectra

In ideal conditions, NIR spectra are based on the Beer-Lambert law and the sample absorbance is directly linked to the concentration of chemical compounds of the sample. But in practice, many factors (independent from the sample composition) will influence the measured absorbance. This is the case for instance of temperature or granulometry, which will deform the final spectrum, biasing spectra comparison. To deal with external effects, a mathematical correction or preprocessing can be applied as illustrated in **Figure 2** for spectra collected on bread wheat grains. Mainly two external effects usually need to be corrected: additive and multiple effects. In additive effects, noise affects spectra irrespectively of the wavelength and usually yields a baseline shift which can be corrected with a detrend ***(28)*** or a derivation (**Figure 2C**) typically carried out through a Savitzky-Golay filter which consists in a polynomial smoothing ***(29)***. The baseline shift appears when the absorbance increases with the wavelength due to the increased light intensity. Multiplicative effects typically affect spectra differently depending on the wavelength and are usually linked to an increase of the distance crossed by the photons (due to different granulometry for example). They can be corrected by a normalization (**Figure 2B**). This effect is present when for low absorbances at a wavelength, the variability is also low and for high absorbances at a wavelength, the variability is high. Other preprocessing techniques have also been proposed to specifically deal with an external parameter known to bias spectra such as temperature or hygrometry. This is the case for instance of the method called External Parameter Orthogonalization (EPO, ***(30)***).

The preprocessing methods briefly introduced in the previous paragraph have been previously developed in the chemometrics literature. This preprocessing is routinely and widely used when applying NIRS in the classical way, *i.e.* to predict the composition of end-products. In the context of breeding and PS, further preprocessing taken from the breeding literature can be carried out to improve the ability of the spectra to predict genetic values. Such preprocessing includes building a model on the absorbance or reflectance at each wavelength taking into account the effects of the experimental design (e.g. blocks or spatial effects) together with genetic effects to further extract genotypic values **(23 , 25)**. Genotypic values may be BLUEs or BLUPs depending on whether the genotype effect is considered as fixed or random in the model. This preprocessing typically comes from the fact that PS is carried out at the genotype level rather than at the individual or plot level, and consequently one needs to obtain a unique NIRS matrix at the genotype level for model training and prediction. It is interesting to note that if the entire spectrum is considered rather than absorbance or reflectance at given wavelengths, such corrections are related to the orthogonalization approaches from the chemometrics literature. Indeed, recently Ryckeweart et al. ***(31)***proposed to make use of spectra replicates, typically obtained when characterizing plants under genetic trials, to reduce the repeatability error. They developed a new preprocessing technique based on orthogonalization after an ANOVA–simultaneous component analysis (REP-ASCA).

The filters mentioned previously are not an exhaustive list but have been the most commonly used in NIRS chemometric prediction. Preprocessing can be done in numerous ways, as shown across different studies, suggesting that no one standard preprocessing approach exists. We have noticed that PS predictions were influenced by the preprocessing applied on spectra, consequently we recommend testing different filters on a subset of data to cross-validate filters efficiency, before carrying out deeper analysis.

C. Statistical models for phenotype prediction

NIRS reflectances or absorbances are quantitative variables, like bi-allelic markers usually coded numerically with allelic dosages, basically all models developed or used in the frame of genomic selection can also be used for PS, from the “infinitesimal” model to Bayesian models with various prior distributions or machine learning methods.

One such reference model for PS is the H-BLUP, similar to G-BLUP but with a similarity matrix (H) estimated with NIRS ***(1, 21, 25)***. Different kernels can be used within such a framework, including Gaussian kernel or arc-cosine kernel ***(22)***. As with molecular markers, this model can be equivalent to a ridge-regression on the wavelengths, provided the H matrix is computed accordingly, as demonstrated hereafter. The predictive ability of the H-BLUP model can be measured with cross-validation, as with G-BLUP or other GS models.

The H-BLUP model is defined as: $y=\mu+u+e$, with $var(y)=H\sigma_u^2+I\sigma_e^2$, and where y is a vector of phenotypes, H is the NIR similarity matrix as defined above, μ is the intercept, u and e are random genetic and residual effects, respectively. The RRN-BLUP model (Ridge Regression NIRS BLUP) is defined as: $y=\mu+Sv+e$, with $var(y)=SS'\sigma_S^2+I\sigma_e^2$, and where S is the matrix of preprocessed, centered and scaled NIRS as defined above. The mean of y is equal to μ in both models, thus H-BLUP and RRN-BLUP are equivalent if $H\sigma_u^2=SS'\sigma_S^2$, which is for instance the case when $H=\frac{SS'}{n_w}$ and $\sigma_S^2=\frac{\sigma_u^2}{n_w}$.

Functional regression models seem particularly interesting for PS, as they model the linear trend of the spectra ***(20)***. Different kinds of functional regressions were proposed such as functional B-Spline, functional Fourier ***(20)*** and Bayesian functional ***(32)***. H-BLUP and functional regression models have proven to yield accurate predictions while reducing computational time by diminishing the number of parameters to estimate. This could be important if several spectra from different environments are available, resulting in a high number of predictors.

Partial Least Squares (PLS) regression, classically used in chemometrics, or variable selection approaches (such as LASSO or BayesB) can also be used to tackle multicollinearity and high dimensionality. PLS regression consists of condensing the information contained across all wavelengths into a few orthogonal variables that maximize the covariation between the predictor matrix and the response variable. In LASSO and Bayes B, it is assumed that only a portion of the variables has an effect on the trait. Variable selection seems promising for PS, because the spectrum could be restricted to its most heritable parts ***(25, 33)***. However, it should be noted that the preselection of wavelengths using vegetation indices or with knowledge on the genomic heritability of the wavelengths generally result in lower prediction accuracies than when using the full spectrum ***(24-25)***.

In GSthe choice of the prediction model can be guided by the expected genetic architecture of the predicted trait. The choice of a PS model adapted to a given trait cannot yet rely on such assumptions, and it is not clear how the optimal prediction model can be related to the trait characteristics. The various models tested in the literature sometimes resulted in contrasted prediction accuracies, but in general sophisticated models were not better than a simple H-BLUP. Models relying on a mixture of distributions such as BayesR ***(34)*** are accurate for contrasted genetic architecture in GS, it would be interesting to test them in PS. In any case, alternative prediction models should be compared using cross-validations within the calibration set.

Contrary to molecular markers in GS, in PS several spectra corresponding to different replicates of genotypes possibly across different environments can be available to build predictive models. In this case, one possibility for calibration is to test each spectrum in order to determine the one which yields the most accurate predictions. Another possibility is to make use of all the available spectra. Lane *et al.* **(23)** proposed in the frame of the H-BLUP model to compute the mean of the relationship matrices calculated from each spectrum individually. It is noteworthy that this

proposition is equivalent to computing the relationship matrix from a large combined spectra matrix, providing that the individual spectra matrices have the same number of wavelengths, as shown hereafter.

The similarity matrix $H_T(i,j)$computed with the combined spectra matrix S_T(in which all spectra matrices are included one next to the other) is given by:

$$H_T(i,j)=\frac{\sum_{k=1}^{n_t}\left[S_T(i,k)\times S_T(j,k)\right]}{n_t}$$

$$H_T(i,j)=\frac{\sum_{p=1}^{n_w}\left[S_1(i,p)\times S_1(j,p)\right]+\ldots+\sum_{p=1}^{n_w}\left[S_{n_l}(i,p)\times S_{n_l}(j,p)\right]}{n_l\times n_w}$$

$$H_T(i,j)=\frac{1}{n_l}\times$$

$$H_T(i,j)=\frac{1}{n_l}\times$$

S_T has dimension n (number of individuals) times $n_t=n_l\times n_w$, with $n_l$the number of spectra (e.g. number of environments in which NIRS was acquired) and $n_w$the number of wavelengths of each spectrum (we consider that all spectra have the same wavelengths). $S_{ud}(i,k)$ is the absorbance or reflectance measured on the *i*th individual for the *k*th wavelength in the *u*th NIR preprocessed spectrum. S_{ud} has the dimension *n* (number of individuals) times n_w(number of wavelengths). $H_u(i,j)$is the similarity between individuals *i* and *j* estimated with one given *u* spectrum.

D. Relative performance of PS versus GS

There are very few studies that compare PS (and in particular GLOB selection) with GS (**Table 2)**. Although Lane et al. ***(23)*** was one of the two studies that implemented GLOB prediction with spectra following our definition, it could not be included in this comparison because they did not apply GS. **Table 2** illustrates that PS and GLOB selection have been mainly implemented on cereal species, probably because of the widespread and routine use of NIR measurements on grains to predict protein content. Krause et al. **(21)** and Galan et al. **(25)** reported similar accuracies for GS and PS, while Cuevas et al. **(22)** showed lower accuracies for PS compared to GS (0.37 and 0.46, respectively). The highest PS accuracy compared with GS was observed in Rincent et al. ***(1)*** in wheat. The lowest PS accuracy compared to GS was observed in Rincent et al. ***(1)*** in poplar for which NIR spectra were collected on wood for a reduced range of wavelengths. From these data, it is apparent that PS had comparable or higher accuracies than GS in most cases. Even in cases where PS is less accurate than GS, as NIR measurements are high-throughput and low-cost compared to genotyping, PS could still provide higher genetic gains than GS, as demonstrated in Rincent et al. ***(1)***. In our ongoing research we compared GS and PS at different generations of elite bread wheat selection. We found that PS could be as accurate as GS and even better when applied to early generations. Further work on other species is clearly needed to deepen this comparison and provide valuable information on the factors and conditions (e.g. tissue, environment) that determine the predictive ability of NIRS.

By considerably reducing the costs of implementation, PS is a tool of choice to improve the balance between costs and benefits in comparison with GS. PS would be particularly valuable for orphan crops for which genotyping is expensive, or for major crops for which phenomic data are already routinely collected (e.g. maize and wheat). In the latter case, phenomic prediction already opens new possibilities in existing breeding programs without any additional cost, and with predictive abilities similar to those obtained with genomic prediction ***(1)***.

E. Factors affecting PS predictive ability

In the past, several kinds of omics were used to make genomic-like predictions with promising results ***(8)***. NIRS captures an integrative signal, and is biologically more difficult to interpret than other omics, which describe each molecule individually (e.g. transcriptomics, proteomics, metabolomics). However, because prediction models do not necessarily need to be interpreted biologically, NIRS can be used to make predictions using “black-box” models.

There are two factors that contribute to the predictive ability and consequently to the success of PS: (i) the ability to capture target trait proxies, and (ii) the ability to infer genetic relatedness. The former depends on the physiological connectedness between the target trait and the composition and features of the tissue analyzed with NIRS. This is for example the case when NIR spectra collected on wood powder is used to predict wood properties, or when NIR spectra collected on fruits is used to predict fruit composition. In these cases, PS should be nearly equivalent or superior to the traditional way of using NIRS (prediction of the tissue composition), the only difference being that when doing PS we usually work at the genotype level because we aim at ranking and selecting the best genotypes while in the traditional use of NIRS we make predictions at the plot or plant level ***(23)***. But we could think of more indirect relationships between the target trait and NIR spectra to explain its predictive ability, for instance in wheat the good predictive ability of PS for yield could be due to the fact that NIRS is a very good predictor of grain composition, that is often negatively correlated to yield. This could also be the case for maturity: the spectra are influenced by the maturity of the plants, and this maturity is sometimes correlated to yield ***(21)***. However, it is important to stress that even in the absence of any direct relationship between the predicted trait and the tissue analyzed with NIRS, PS can still be accurate. It was for instance shown in Rincent et al. ***(1)*** that NIR spectra collected on leaves in one environment could be used to estimate a covariance matrix resulting in accurate prediction of yield in a completely independent environment. In this particular example, the correlation between yield in the environment in which NIR spectra were obtained and yield in the predicted

environment (by cross-validation within the predicted environment) was as low as 0.16, whereas PS predictive ability was above 0.5. This means that NIRS derived relationship matrices were able to capture genetic relatedness between lines valuable for predicting yield. This was further demonstrated by the fact that genomic heritability was significant for many wavelengths. A further demonstration could be done with a simulation study, by estimating the predictive ability of PS for traits simulated with genotype data. In this case, the predictive ability of PS averaged over a large number of simulated traits would provide an evaluation of the ability of NIRS to infer genetic relationships for predicting quantitative traits unrelated to the tissue composition.

PS is a recent research topic, and further investigations are required to use it in an optimal way. One can expect that, as for genomic selection, prediction accuracy will be strongly dependent on the target trait and its heritability, as well as the size and composition of the training set. In comparison to GS, prediction accuracy obtained with PS can also be affected by the origin of the spectrum (tissue, environments, kind of sensors). This is similar to the choice of a SNP array and marker filtering in GS, but the effect of the origin of the spectrum appears to be more pronounced. First results suggest that NIR spectra collected under plant stress conditions are more efficient ***(1, 23)***, but other experiments are required before it can be understood if this is the rule or the exception. An interesting result is that, in practice, the combination of different NIR spectra (collected on different tissues or different environments) leads to predictive abilities at least as good as those obtained with the best NIR spectrum taken alone ***(1***, unpublished results***)***. This means that in some cases it is not necessary to identify the best conditions to obtain NIR spectra, but simply to aggregate all the spectra collected (e.g. spectra obtained on the same genotypes at the different steps of the breeding program). As shown in the present review, aggregating NIRS matrices prior to computing the H matrix is equivalent to averaging H matrices estimated with individual NIRS matrices and is thus quite straightforward. But, in any case, the choice of the tissue, timing and sensors could and should also be optimized. We can

think that NIR spectra collected on homogeneous, representative samples (leaf powder, seed sample or flour) are more useful than NIR spectra obtained on a tiny area of raw material. The way NIR spectra are collected should also be optimized in terms of practical feasibility. For instance in wheat, it would be much more feasible to measure NIR spectra during the growing season, than on seeds after harvest, because the few weeks between harvest and sowing of the next generation is labor intensive and so NIR spectra acquisition would be difficult during this period.

Prospects

As introduced above, there are numerous ways of using PS in breeding. We particularly foresee several applications to be addressed with PS and that we detail hereafter. Some of them are quite direct applications, which can already be deployed in breeding programs (A-F), while others are prospects which deserve research investments prior to their adoption in breeding in a relatively longer term (G). Although many of the prospects presented here are also shared with GS, we have tried in what follows to underline their specificity with respect to PS.

A. Prebreeding: screening diversity collections at low cost

Gene banks are a reservoir of genetic diversity in which genes of tolerance to biotic or abiotic stress can be discovered. These collections will become crucial as the genetic diversity in the breeding program will not be sufficient to face upcoming changes due to the evolution of management practices or climate change. However, the identification of promising genes or individuals require the phenotyping of the collection for the target trait, which could be too expensive given the number of accessions stored in these collections. Yu et al. ***(35)*** and Crossa et al. ***(36)***, have proposed to run GS to screen these collections. However, this requires genotyping the full collection, which is also expensive considering the large size of the gene banks. Another option would be to measure NIR spectra on each accession, phenotype a

subset, and predict the remaining accessions using PS. In most genebanks, accessions are regularly sown in nurseries to produce new seeds, as the germination rate decreases with time. NIRS could thus either be collected on seeds directly in gene banks or on other tissues in these nurseries. The same approach could also be extended to other species like perennials which are usually not conserved as seeds but as living plants in the field.

B. Sparse testing: experimental design optimization in breeding programs

For most species, selection candidates are evaluated in multi-environment trials to estimate their stability and productivity in contrasted environments. This is an expensive step as the number of variety/environment combinations can rapidly become very high. To increase the number of environments or varieties with the same costs, one option would be to run a sparse testing design, in which all varieties are evaluated at least in one environment, but with a given proportion of varieties/environments not tested. Sparse testing is sometimes imposed to breeders because part of a trial is accidentally destroyed, or not harvested for some reason. Sparse testing is a scenario for which genomic predictions are particularly accurate, in comparison to predicting a completely new variety or a new environment ***(37-39)***. Our proposition is to use PS to predict the unobserved variety/environment combinations. In wheat, this approach is already applicable as NIR spectra are usually collected in nurseries the year before the multi-environment trials. Our first unpublished results on sparse testing show that the prediction accuracy of PS under this scenario can be as high as to the one obtained with GS.

C. Combining reduction of generation time (speed breeding) and performance prediction (PS) to increase genetic progress

One challenge in breeding is to accelerate programs to quickly release new varieties. Breeding is often constrained to one to two generations per year for annual crops, which limits genetic progress. Several methods have been proposed to reduce generation time, including recently "speed breeding" ***(40)***. Speed breeding consists in controlling photoperiod and temperature to get optimal growing conditions and accelerate the time elapsed from seed to seed, allowing up to six spring wheat generations per year. Thus, a great number of segregation and recombination events can occur in a short time, allowing to rapidly produce varieties combining favorable alleles. PS could have two applications that could work particularly well if combined

with speed breeding. The first application could be implemented in the speed breeding process itself. During this process, each plant is unique and is phenotypically quite different from what it would look in the field, which makes direct selection impossible except for phenology, height and some disease resistance traits. But molecular markers or NIR spectra could be measured on the plants or on the seeds which would allow predicting performance traits using GS ***(41)***, or at a lower cost PS. This would considerably increase genetic progress by reducing both the generation interval (speed breeding) and the phenotyping process (partially replaced by predictions). The second application would be to apply PS to reduce the field trials size after the speed breeding process. Thousands of genotypes can be produced by speed breeding, and it would be difficult to phenotype all of them in field experiments (after multiplication in the nurseries). As NIR spectra can be collected at the end of the process or in the nurseries, PS could be used to predict genotype performance in the nurseries or with sparse testing in the following field experiments as described above.

D. GEI prediction

Unlike molecular markers, NIR spectra are directly influenced by the response of the plants to the environmental conditions. This seems likely to result in a lack of stability of the spectra in different environments. It also may mean that the spectra are able to capture the genetic responses to a given environment (GEI), as illustrated in Rincent et al. ***(1)***, which opens new perspectives of application for PS. This GEI variance could be exploited to enhance predictive ability of local adaptation. One possibility to improve the GEI prediction models is to use NIRS collected in each environment to estimate environment specific covariance matrices. Krause et al. ***(21)*** compared different models (single and multi-kernel) using molecular marker, pedigree or NIR spectra, to predict wheat grain yield in multi-environment trials. They found that the best multi-kernel integrating GEI was the one with the hyperspectral matrix. Using NIRS enhanced predictive accuracy of GEI compared to models that use molecular markers or pedigree. Similar

results were observed by Montesinos et al. ***(20)*** and Lane et al. ***(23)*** where the interaction between NIR bands and environments was integrated in the models. The introduction of NIR information in the prediction model allowed increasing GEI prediction accuracy, which was not the case with molecular markers.

This multi-kernel method requires collecting the spectra of all genotypes in all the environments of the multi-environment trials, which is not possible with classical sparse testing designs. One possibility would be to grow a nursery in parallel of each trial, which would be much cheaper than a real trial. In this case, the trial will be dedicated to phenotype the training population, and the nursery to collect spectra on both training and predicted varieties. Krause et al. ***(21)*** proposed to reduce the size of the microplots, to measure NIR spectra on all varieties in all environments using a UAV. The objective is to find a compromise between the number of candidate varieties predicted with PS and the number of training varieties phenotyped for the target traits.

Another potential application would be to use NIR spectra collected in the different environments to estimate similarities between them, as proposed in Jarquin et al. ***(37)*** and Heslot et al. ***(42)*** with environmental covariates. In comparison to classical environmental covariates (e.g. temperature, hydric balance, radiation), NIRS has the advantage of capturing signals from the plant that could be influenced by the experienced stresses. GEI models enriched with NIRS would be particularly useful to make predictions in multi-environment trials. This application of PS could potentially allow making predictions in new environments (no other phenotype than NIR spectra are collected in these environments) by estimating the similarity between the new environment and the calibration environments. NIR spectra can also be used as a covariate in the predictive model to characterize the GEI. Lane et al. ***(23)*** proposed to take into account the GEI by using wavelengths as covariates in the predictive model through functional regression. They predicted yields of known hybrids in an unknown environment and

found that taking into account GEI with covariates worked better than regular H-BLUP. The different models tested in these studies underline that NIRS can be a good predictor of GEI, and in many different ways. These results are very promising to enhance predictive ability in the context of multi-environment trials.

E. Making use of historical NIRS data in prediction

In GS the enrichment of the training set with historical data from multiple environments can improve predictive ability by increasing the size of the training set and by limiting the effect of atypical years that are difficult to predict with reduced and traditional datasets ***(43)***. The use of historical data might be more complicated with PS, as NIR spectra are likely to be more specific to the environment in which they are measured. To estimate the NIR similarity matrix (H) between historical varieties characterized with NIRS over successive years (the varieties changing from one year to another), the effect of environment and GEI should be accounted for and corrected from the spectra. One option would be to use check varieties for which NIR spectra would be collected each year to determine the transformation from the spectra collected in one environment to the spectra collected in another environment.

F. The case of perennials

Genetic resources for perennials are usually conserved as collections in nurseries. These collections typically include individuals from different species, populations, genotypes as well as clones. In this case, we foresee several potential applications of PS. First, as already mentioned in the case of genebanks, NIRS could be used to rapidly screen these resources for target traits that are typically difficult to evaluate on a large number of individuals, such as fruit quality or phenology. Second, in a more advanced breeding context, one could imagine that progenies from controlled crosses could be planted near a well characterized reference panel of widely cultivated varieties. NIRS data could be collected at the same time on both reference and

candidate individuals to limit environmental heterogeneity between NIRS matrices from training and test sets, and a model could be trained by using phenotypic data potentially obtained in other sites for the reference panel in order to predict the performance of the candidates in these particular sites. This scenario corresponds to GLOB prediction, which proved to be quite efficient for some traits in wheat and poplar in Rincent et al. **(1)**.

G. Other applications

G.a. Genotype inference

Whalen et al. ***(44)*** described a new application of NIRS for breeding beyond PS: the use of High-Throughput Phenotypes (HTP) such as NIRS to infer genotypes. They illustrated their approach with simulations, in which spectra resulted from an additive genetic model, with 100 QTLs per chromosome and heritability of wavelength ranging between 0.1 and 0.7. From that, they fit a model which links the segregation states at each locus with the HTP, for the training population. Then, this model was applied to predict the segregation states of non genotyped varieties. They concluded that under certain conditions it was possible to infer the genotypes of individuals derived from biparental crosses. The HTP-enabled genomic prediction (with genotypes inferred from HTP data) yielded higher accuracies than PS, the best accuracies being from classic genomic prediction models with real genotypes. However, it is important to note that PS does not only capture additive genetic effects, but also epistatic and GEI. This advantage of PS over GS could not be illustrated here, as only additive effects were simulated. Nevertheless this application seems very promising for breeding, especially for species with high genotyping cost.

Furthermore, as PS relies on global relatedness between genotypes derived from NIRS, it could be difficult to predict traits with mono- or oligogenic genetic architecture. Hence this transition proposed by Whalen et al. **(44)** from NIRS to marker data allows to apply GS but also QTL

detection. This shows that the use of NIRS instead of genotyping in breeding may be adapted to contrasted genetic architectures. A demonstration of its usefulness with real data remains to be conducted.

G.b. Hybrid prediction

As NIR spectrum is a phenotype, it captures both additive and interaction effects. For this reason, it is possible that the NIRS covariance matrix could be used to predict hybrid performances, taking into account both general and specific combining abilities. The idea would be to collect NIR spectra on the hybrids to estimate a covariance matrix tacking interaction effects into account. One option would be to collect NIRS data on large collections of hybrids in nurseries, and phenotype only part of the hybrids in classical multi-environment trials, possibly with the sparse testing approach described above.

G.c. Progeny sorting

In some programs, breeders are interested in quickly characterizing progenies from controlled crosses with respect to their resemblance to their parents. This is typically the case when crossing an established widely-used variety with a donor genotype with the aim to introgress particular features from the donor to the variety. An example of this would be the case of grapevine for which a current challenge is to quickly breed varieties with resistance to biotic factors while maintaining some established quality for wine making. In this case when a given variety is crossed with a disease resistant genotype, it is usually quite straightforward using marker assisted selection to select with molecular markers the progenies that carry disease resistant genes, but for wine quality it is much more complicated. The goal would be to select among the resistant progenies those that are more similar to the parental wine-making variety, which using molecular markers is a very difficult task without considering prior knowledge on the

genetic architecture of the traits. In this particular case NIRS could be useful to provide a distance between the resistant progenies and their parents.

Conclusion

We have reviewed the different approaches that have so far been proposed in literature to predict agronomic traits with NIRS, from prediction at the plot/individual level to PS and GLOB selection. NIRS has been intensively used to make predictions at the plot or genotype level, considering NIRS as a secondary trait or as a yield proxy. The originality of PS and GLOB selection as defined in Rincent et al. ***(1)*** is that NIR spectra are used in a similar way as molecular markers in GS. They indeed supposed that NIR spectra were able to capture genetic relationships between individuals. PS resulted in good predictive abilities (often similar or even higher than those obtained with GS), even when the predicted trait was completely independent from the tissue analyzed with NIRS (e.g. different environments). The high-throughput and low-cost of this approach makes it interesting to increase breeding efficiency in comparison to GS, particularly for species for which genotyping is expensive, or for crops for which NIRS data are already routinely collected in the breeding programs. We have also proposed different promising applications of PS in breeding and prebreeding among which some can readily be applicable, while others require further work in order to test and optimize this approach.

Funding

The preparation of this chapter was partly supported by the project OASIs funded by the French Ministry of Agriculture and Food (“fond CASDAR, appel à projets Semences et sélection végétale 2020, projet n° C-2020-5”) and by the BreedWheat project funded by the French National Research Agency under Investment for the Future (ANR-10-BTBR-03).

Partial funding of Charlotte Brault's PhD was provided by the Association Nationale de la Recherche et de la Technologie (ANRT, grant number 2018/0577), IFV and Inter-Rhône. Pauline Robert's PhD was jointly funded by Agri-Obtentions, Florimond-Desprez, and the Association Nationale de la Recherche et de la Technologie (ANRT, grant number 2019/0060).

Aknowledgements

We would like to thank the reviewer and editors for their suggestions and useful remarks on a previous version of this chapter. We also thank Jacques Le Gouis for proposing the "GLOB" denomination.

References

1. Rincent R, Charpentier J-P, Faivre-Rampant P, Paux E, Le Gouis J, Bastien C, Segura V (2018) Phenomic Selection Is a Low-Cost and High-Throughput Method Based on Indirect Predictions: Proof of Concept on Wheat and Poplar. G3 Genes Genomes Genet 8: 3961–3972. https://doi.org/10.1534/g3.118.200760

2. Riedelsheimer C, Czedik-Eysenberg A, Grieder C, Lisec J, Technow F, Sulpice R, Altmann T, Stitt M, Willmitzer L, Melchinger AE (2012) Genomic and metabolic prediction of complex heterotic traits in hybrid maize. Nat Genet 44: 217–220. https://doi.org/10.1038/ng.1033

3. Feher K, Lisec J, Römisch-Margl L, Selbig J, Gierl A, Piepho H-P, Nikoloski Z, Willmitzer L, (2014) Deducing Hybrid Performance from Parental Metabolic Profiles of Young Primary Roots of Maize by Using a Multivariate Diallel Approach. PLOS ONE 9: e85435. https://doi.org/10.1371/journal.pone.0085435

4. Ward J, Rakszegi M, Bedő Z, Shewry PR, Mackay I (2015) Differentially penalized regression to predict agronomic traits from metabolites and markers in wheat. BMC Genet 16: 19. https://doi.org/10.1186/s12863-015-0169-0

5. Fernandez O, Urrutia M, Bernillon S, Giauffret C, Tardieu F, Le Gouis J, Langlade N, Charcosset A, Moing A, Gibon Y (2016) Fortune telling: metabolic markers of plant performance. Metabolomics 12: 158. https://doi.org/10.1007/s11306-016-1099-1

6. Xu S, Xu Y, Gong L, Zhang Q (2016) Metabolomic prediction of yield in hybrid rice. Plant J 88: 219–227. https://doi.org/10.1111/tpj.13242

7. Guo Z, Magwire MM, Basten CJ, Xu Z, Wang D (2016) Evaluation of the utility of gene expression and metabolic information for genomic prediction in maize. Theor Appl Genet 129: 2413–2427. https://doi.org/10.1007/s00122-016-2780-5

8. Schrag TA, Westhues M, Schipprack W, Seifert F, Thiemann A, Scholten S, Melchinger AE (2018) Beyond Genomic Prediction: Combining Different Types of omics Data Can Improve Prediction of Hybrid Performance in Maize. Genetics 208: 1373–1385. https://doi.org/10.1534/genetics.117.300374

9. Osborne BG (2006) Applications of near Infrared Spectroscopy in Quality Screening of Early-Generation Material in Cereal Breeding Programmes. J Near Infrared Spectrosc 14, 93–101. https://doi.org/10.1255/jnirs.595

10. Holroyd SE (2013) The Use of near Infrared Spectroscopy on Milk and Milk Products. J Near Infrared Spectrosc 21: 311–322. https://doi.org/10.1255/jnirs.1055

11. Tsuchikawa S, Kobori H (2015) A review of recent application of near infrared spectroscopy to wood science and technology. J Wood Sci 61: 213–220. https://doi.org/10.1007/s10086-015-1467-x

12. Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI, Lammertyn J (2007) Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A

review. Postharvest Biol Technol 46: 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024

13. Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ 8: 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

14. Labus MP, Nielsen GA, Lawrence RL, Engel R, Long DS (2002) Wheat yield estimates using multi-temporal NDVI satellite imagery. Int J Remote Sens 23: 4169–4180. https://doi.org/10.1080/01431160110107653

15. Ferrio JP, Bertran E, Nachit MM, Català J, Araus JL (2004) Estimation of grain yield by near-infrared reflectance spectroscopy in durum wheat. Euphytica 137: 373–380. https://doi.org/10.1023/B:EUPH.0000040523.52707.1e

16. Babar MA, Reynolds MP, van Ginkel M, Klatt AR, Raun WR, Stone ML (2006) Spectral Reflectance to Estimate Genetic Variation for In-Season Biomass, Leaf Chlorophyll, and Canopy Temperature in Wheat. Crop Sci 46: 1046–1057. https://doi.org/10.2135/cropsci2005.0211

17. Rutkoski J, Poland J, Mondal S, Autrique E, Pérez LG, Crossa J, Reynolds M, Singh R (2016) Canopy Temperature and Vegetation Indices from High-Throughput Phenotyping Improve Accuracy of Pedigree and Genomic Selection for Grain Yield in Wheat. G3 Genes Genomes Genet 6: 2799–2808. https://doi.org/10.1534/g3.116.032888

18. Hayes BJ, Panozzo J, Walker CK, Choy AL, Kant S, Wong D, Tibbits J, Daetwyler HD, Rochfort S, Hayden MJ, Spangenberg GC (2017) Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. Theor Appl Genet 130: 2505–2519. https://doi.org/10.1007/s00122-017-2972-7

19. Sun J, Rutkoski JE, Poland JA, Crossa J, Jannink J-L, Sorrells ME (2017) Multitrait, Random Regression, or Simple Repeatability Model in High-Throughput Phenotyping Data Improve Genomic Prediction for Wheat Grain Yield. Plant Genome 10. https://doi.org/10.3835/plantgenome2016.11.0111

20. Montesinos-López OA, Montesinos-López A, Crossa J, de los Campos G, Alvarado G, Suchismita M, Rutkoski J, González-Pérez L, Burgueño J (2017) Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. Plant Methods 13: 4. https://doi.org/10.1186/s13007-016-0154-2

21. Krause MR, González-Pérez L, Crossa J, Pérez-Rodríguez P, Montesinos-López O, Singh RP, Dreisigacker S, Poland J, Rutkoski J, Sorrells M, Gore MA, Mondal S (2019) Hyperspectral Reflectance-Derived Relationship Matrices for Genomic Prediction of Grain Yield in Wheat. G3 Genes Genomes Genet 9: 1231–1247. https://doi.org/10.1534/g3.118.200856

22. Cuevas J, Montesinos-López O, Juliana P, Guzmán C, Pérez-Rodríguez P, González-Bucio J, Burgueño J, Montesinos-López A, Crossa J (2019) Deep Kernel for Genomic and Near Infrared Predictions in Multi-environment Breeding Trials. G3 Genes Genomes Genet 9: 2913–2924. https://doi.org/10.1534/g3.119.400493

23. Lane HM, Murray SC, Montesinos-López OA, Montesinos-López A, Crossa J, Rooney DK, Barrero-Farfan ID, Fuente GNDL, Morgan CLS (2020) Phenomic selection and prediction of maize grain yield from near-infrared reflectance spectroscopy of kernels. Plant Phenome J 3: e20002. https://doi.org/10.1002/ppj2.20002

24. Aguate FM, Trachsel S, Pérez LG, Burgueño J, Crossa J, Balzarini M, Gouache D, Bogard M, de los Campos G (2017) Use of Hyperspectral Image Data Outperforms Vegetation

Indices in Prediction of Maize Yield. Crop Sci 57: 2517–2524. https://doi.org/10.2135/cropsci2017.01.0007

25. Galán RJ, Bernal-Vasquez A-M, Jebsen C, Piepho H-P, Thorwarth P, Steffan P, Gordillo A, Miedaner T (2020) Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye. Theor Appl Genet 133: 3001–3015. https://doi.org/10.1007/s00122-020-03651-8

26. Hernandez J, Lobos G, Matus I, del Pozo A, Silva P, Galleguillos M (2015) Using Ridge Regression Models to Estimate Grain Yield from Field Spectral Data in Bread Wheat (Triticum Aestivum L.) Grown under Three Water Regimes. Remote Sens 7: 2109–2126. https://doi.org/10.3390/rs70202109

27. Zgouz A, Héran D, Barthès B, Bastianelli D, Bonnal L, Baeten V, Lurol S, Bonin M, Roger J-M, Bendoula R, Chaix G (2020) Dataset of visible-near infrared handheld and micro-spectrometers – comparison of the prediction accuracy of sugarcane properties. Data in Brief 31: 106013. https://doi.org/10.1016/j.dib.2020.106013

28. Barnes RJ, Dhanoa MS, Lister SJ (1989) Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. Appl Spectrosc 43: 772–777. https://doi.org/10.1366/0003702894202201

29. Savitzky A, Golay MJ (1964) Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry 36: 1627-1639. https://doi.org/10.1021/ac60214a047

30. Roger J-M, Chauchard F, Bellon-Maurel V (2003) EPO–PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. Chemom Intell Lab Syst 66: 191–204. https://doi.org/10.1016/S0169-7439(03)00051-0

31. Ryckewaert M, Gorretta N, Henriot F, Marini F, Roger J-M (2020) Reduction of repeatability error for analysis of variance-Simultaneous Component Analysis (REP-ASCA): Application to NIR spectroscopy on coffee sample. Anal Chim Acta 1101: 23–31. https://doi.org/10.1016/j.aca.2019.12.024

32. Montesinos-López A, Montesinos-López OA, Cuevas J, Mata-López WA, Burgueño J, Mondal S, Huerta J, Singh R, Autrique E, González-Pérez L, Crossa J (2017) Genomic Bayesian functional regression models with interactions for predicting wheat grain yield using hyper-spectral image data. Plant Methods 13: 62. https://doi.org/10.1186/s13007-017-0212-4

33. Ferragina A, de los Campos G, Vazquez AI, Cecchinato A, Bittante G (2015) Bayesian regression models outperform partial least squares methods for predicting milk components and technological properties using infrared spectral data. J Dairy Sci 98: 8133–8151. https://doi.org/10.3168/jds.2014-9143

34. Erbe M, Hayes BJ, Matukumalli LK, Goswami S, Bowman PJ, Reich CM, Mason BA, Goddard ME (2012) Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J Dairy Sci 95: 4114–4129. https://doi.org/10.3168/jds.2011-5019

35. Yu X, Li X, Guo T, Zhu C, Wu Y, Mitchell SE, Roozeboom KL, Wang D, Wang ML, Pederson GA, Tesso TT, Schnable PS, Bernardo R, Yu J (2016) Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nat Plants 2: 16150. https://doi.org/10.1038/nplants.2016.150

36. Crossa J, Jarquín D, Franco J, Pérez-Rodríguez P, Burgueño J, Saint-Pierre C, Vikram P, Sansaloni C, Petroli C, Akdemir D, Sneller C, Reynolds M, Tattaris M, Payne T, Guzman C, Peña RJ, Wenzl P, Singh S (2016) Genomic Prediction of Gene Bank Wheat Landraces. G3 Genes Genomes Genet 6: 1819–1834. https://doi.org/10.1534/g3.116.029637

37. Jarquín D, Crossa J, Lacaze X, Du Cheyron P, Daucourt J, Lorgeou J, Piraux F, Guerreiro L, Pérez P, Calus M, Burgueño J, de los Campos G (2014) A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor Appl Genet 127: 595–607. https://doi.org/10.1007/s00122-013-2243-1

38. Jarquin D, Howard R, Crossa J, Beyene Y, Gowda M, Martini JWR, Covarrubias Pazaran G, Burgueño J, Pacheco A, Grondona M, Wimmer V, Prasanna BM (2020) Genomic Prediction Enhanced Sparse Testing for Multi-environment Trials. G3 Genes Genomes Genet 10: 2725–2739. https://doi.org/10.1534/g3.120.401349

39. Ly D, Huet S, Gauffreteau A, Rincent R, Touzy G, Mini A, Jannink J-L, Cormier F, Paux E, Lafarge S, Le Gouis J, Charmet G (2018) Whole-genome prediction of reaction norms to environmental stress in bread wheat (Triticum aestivum L.) by genomic random regression. Field Crops Res 216: 32–41. https://doi.org/10.1016/j.fcr.2017.08.020

40. Watson A, Ghosh S, Williams MJ, Cuddy WS, Simmonds J, Rey M-D, Asyraf Md Hatta M, Hinchliffe A, Steed A, Reynolds D, Adamski NM, Breakspear A, Korolev A, Rayner T, Dixon LE, Riaz A, Martin W, Ryan M, Edwards D, Batley J, Raman H, Carter J, Rogers C, Domoney C, Moore G, Harwood W, Nicholson P, Dieters MJ, DeLacy IH, Zhou J, Uauy C, Boden SA, Park RF, Wulff BBH, Hickey LT (2018) Speed breeding is a powerful tool to accelerate crop research and breeding. Nat Plants 4: 23–29. https://doi.org/10.1038/s41477-017-0083-8

41. Watson A, Hickey LT, Christopher J, Rutkoski J, Poland J, Hayes BJ (2019) Multivariate Genomic Selection and Potential of Rapid Indirect Selection with Speed Breeding in Spring Wheat. Crop Sci 59: 1945–1959. https://doi.org/10.2135/cropsci2018.12.0757

42. Heslot N, Akdemir D, Sorrells ME, Jannink J-L (2014) Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor Appl Genet 127: 463-480. https://doi.org/10.1007/s00122-013-2231-5

43. Sweeney DW, Sun J, Taagen E, Sorrells ME (2019) Genomic selection in wheat. In Miedaner T, Korzun V (eds) Applications of Genetic and Genomic Research in Cereals. Woodhead Publishing. https://doi.org/10.1016/B978-0-08-102163-7.00013-2

44. Whalen A, Gaynor C, Hickey JM (2020) Using high-throughput phenotypes to enable genomic selection by inferring genotypes. bioRxiv 2020.02.28.969600. https://doi.org/10.1101/2020.02.28.969600

Tables

Table 1. Selection of work using NIRS information in the context of breeding. The papers are sorted according to the type of prediction made with NIRS information from NIRS as a proxy trait to phenomic selection (PS). For details see the text. PLS-R: Partial Least Squares Regression, VI: Vegetation Index, UAV: Unmanned Aerial Vehicle, OLS: Ordinary Least Squares, G-BLUP: Genomic Best Linear Unbiased Prediction, LASSO: Least Absolute Shrinkage and Selection Operator, GLOB: Genomi-Like Omics-Based.

Reference	**Type of technology**	**Statistical method**	**Type of prediction**	**Prediction setting**
Ferrio et al. ***(15)***	NIR spectroscopy in the laboratory on milled grains, wavelength range: 1100-2500nm.	PLS-R	NIRS as a proxy trait	Plot/individual level
Hernandez et al. ***(26)***	NIR spectroscopy in the field on plant canopy, wavelength range: 350-2500nm.	VIs, Ridge regression	NIRS as a proxy trait	Plot/individual level
Aguate et al. ***(24)***	Hyperspectral imaging in the field (UAV), wavelength range: 392-850nm	VIs, PLS-R, OLS,	NIRS as a proxy trait	Plot/individual level

		BayesB		
Hayes et al. ***(18)***	NIR spectroscopy in the laboratory on grains, wavelength range: 400-2498nm.	Multi-trait G-BLUP	NIRS as a secondary trait	Plot/individual level
Rutkoski et al. ***(17)***	Hyperspectral imaging in the field (UAV)	Multi-trait G-BLUP	NIRS as a secondary trait	Entry/genotype level within each site-year
Sun et al. ***(19)***	Hyperspectral imaging in the field (UAV)	Multi-trait G-BLUP	NIRS as secondary trait	Entry/genotype level within each site-year
Montesinos-Lopez et al. ***(20)***	Hyperspectral imaging in the field (UAV), wavelength range: 392.03-850.69nm	VIs, PLS-R, OLS, BAyes B, functional regressions	PS	Entry/genotype level within each site-year
Krause et al. ***(21)***	Hyperspectral imaging in the field (UAV), wavelength range: 380-850nm	G-BLUP Single & multikernel	PS	Entry/genotype level within each site-year
Galan et al. ***(25)***	Hyperspectral imaging in the field (UAV), wavelength	G-BLUP Single &	PS	Entry/genotype level within each

	range: 410-993nm	multikernel, LASSO and EN for feature selection		site-year
Cuevas et al. ***(22)***	NIR spectroscopy in the laboratory on grains, wavelength range: 400-2500nm.	G-BLUP Gaussian Kernel Arc-cosine kernel	PS	Entry/genotype level within a single environment
Lane et al. ***(23)***	NIR spectroscopy in the laboratory on grains, wavelength range: 1000-2500nm.	PLSR, G-BLUP, Functional Regression	PS & GLOB	Plot/individual level & entry/genotype level within and across environments
Rincent et al. ***(1)***	NIR spectroscopy in the laboratory on milled leaves and grains for wheat (wavelength range: 400-2500nm) and milled wood for poplar (wavelength range: 1250-2500nm)	G-BLUP, Bayesian LASSO	PS & GLOB	Entry/genotype level within and across environments

Table 2. Comparative predictive ability reported for prediction based on G and H matrices. GY: grain yield, DTHD: days to heading date, DMY: dry matter yield, HD: heading date, HT: height, CIRC: circumference, BF: bud flush, BS: bud set, RUST: resistance to rust. ORL: Orléans site (France), SAV: Savigliano site (Italy). Several predictive abilities were available in the study of Krause et al. (2019), one per flying date. When several methods were tested, the one with the highest predictive ability was kept for each modality. When the prediction was made in several environments, years or on different tissues for PS, the range of predictive ability is indicated in brackets.

References	Species	Traits	Mean predictive ability	
			G matrix	H matrix
Krause et al. ***(21)***	Wheat	GY (without DTHD correction)	0.41 (0.19-0.6)	**0.42 (0-0.75)**
Galan et al. ***(25)***	Winter rye	DMY	**0.6**	0.59
Cuevas et al. ***(22)***	Wheat	GY	**0.46**	0.37
Rincent et al. ***(1)***	Wheat S1	HD GY	0.57 0.38	**0.84 (0.78-0.88)** **0.51 (0.37-0.62)**
	Wheat S2	HD	0.61	**0.78 (0.67-0.83)**

		GY	**0.46**	0.44 (0.28-0.53)
	Poplar	HT.ORL	0.56	**0.64 (S1)** / 0.49 (S2)
		CIRC.ORL	0.61	**0.72 (S1)** / 0.48 (S2)
		CIRC.SAV	0.77	**0.80 (S1)** / 0.46 (S2)
		BF.ORL	**0.72**	0.15 (S1) / 0.33 (S2)
		BF.SAV	**0.73**	0.32 (S1) / 0.09 (S2)
		BS.ORL	**0.73**	0.53 (S1) / 0.44 (S2)
		BS.SAV	**0.59**	0.34 (S1) / 0.45 (S2)
		RUST.ORL	**0.61**	0.45 (S1) / 0.33 (S2)

Figures

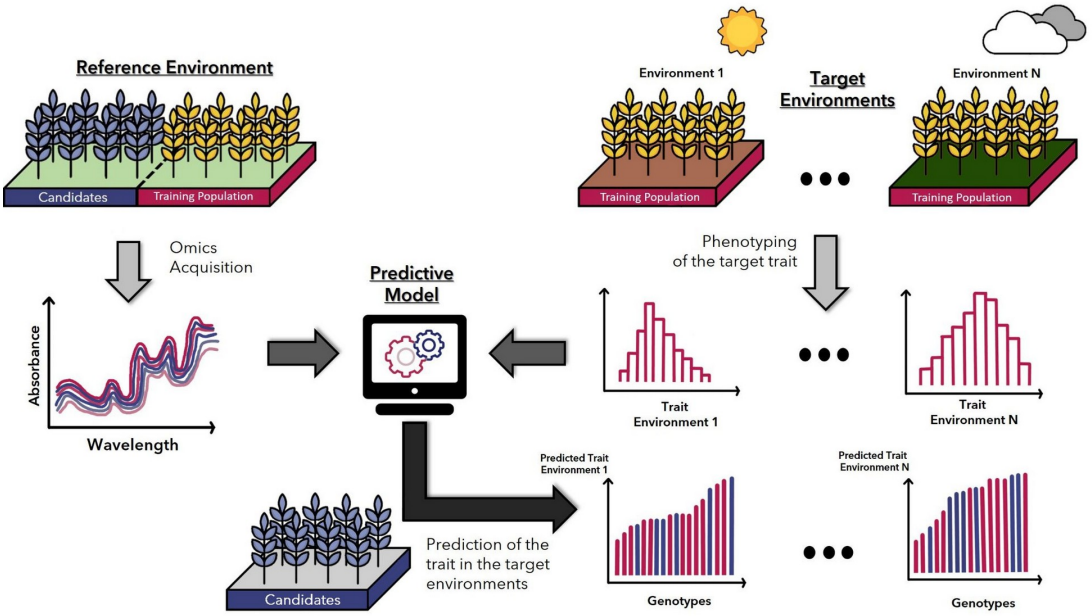


Figure 1.

Prediction of a target trait for the selection candidates in different environments with GLOB selection. Training population is phenotyped for the target traits (e.g. productivity) in the target environments. Omics or phenomic data (e.g. NIR spectra) are collected on both the candidate and the training individuals in a same reference environment, for capturing a genetic similarity between individuals. Genotypic values of the selection candidates are predicted for the target traits in each target environment.

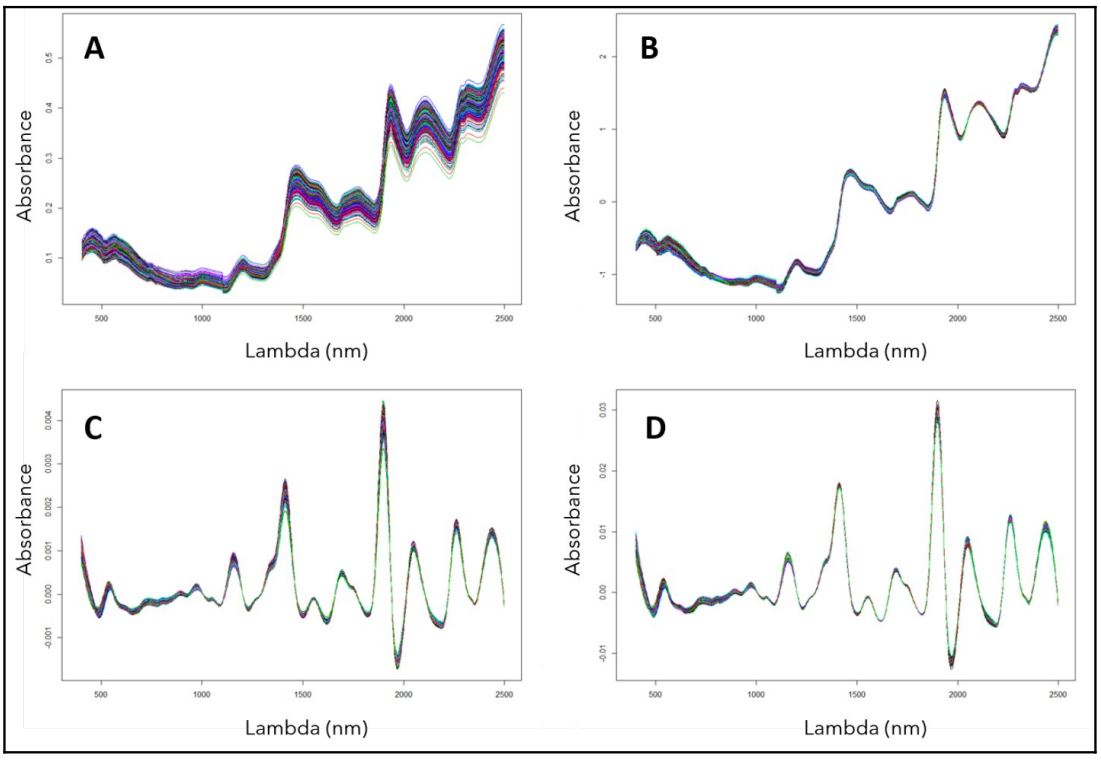


Figure 2. Visualization of different filters applied on a spectrum dataset. Each color represents a bread wheat variety. Spectra were collected on grains with a lab spectrometer NIRS 6500 FOSS. A. raw spectra (no preprocessing) ; B. normalization (standard normal variate) ; C. first derivative of raw spectra ; D. first derivative on normalized spectra.

Appendix B

Appendix Chapter 2

B.1 Supplementary information for article I

Supplemental information

Harnessing multivariate, penalized regression methods for genomic prediction and QTL detection of drought-related traits in grapevine

Authors: Brault C, Doligez A, Le Cunff L, Coupel-Ledru A, Simonneau T, Chiquet J, This P, Flutre T

Table of content:

segreg	coding_ab	phase	coding_phased	AC	AD	BC	BD	gene.dose.AC	gene.dose.AD	gene.dose.BC	gene.dose.BD
lmxll	abxaa	{0-}	abxaa	aa	aa	ba	ba	2	2	1	1
lmxll	abxaa	{1-}	baxaa	ba	ba	aa	aa	1	1	2	2
nnxnp	aaxab	{-0}	aaxab	aa	ab	aa	ab	2	1	2	1
nnxnp	aaxab	{-1}	aaxba	ab	aa	ab	aa	1	2	1	2
hkxhk	abxab	{00}	abxab	aa	ab	ba	bb	2	1	1	0
hkxhk	abxab	{01}	abxba	ab	aa	bb	ba	1	2	0	1
hkxhk	abxab	{10}	baxab	ba	bb	aa	ab	1	0	2	1
hkxhk	abxab	{11}	baxba	bb	ba	ab	aa	0	1	1	2

Table S1: Conversion from genotypic class (in ab x cd segregation) to biallelic gene-dose.
Table used to convert imputed genotypic class from Lep-MAP3 to additive gene dose according to the initial segregation of markers.

Grenache is considered as the father and Syrah the mother for the whole progeny for map construction

FigureS2_BLUPs_comparison_old-new.pdf

Figure S2: Comparison between published and re-computed genotypic BLUPs
For each trait, we plotted the new BLUPs with the BLUPs from Coupel-Ledru *et al.* (2014, 2016). The grey dashed line is the bisector and the red line is the regression line. Correlation between new and previous BLUPs as well as *p-value* are indicated.

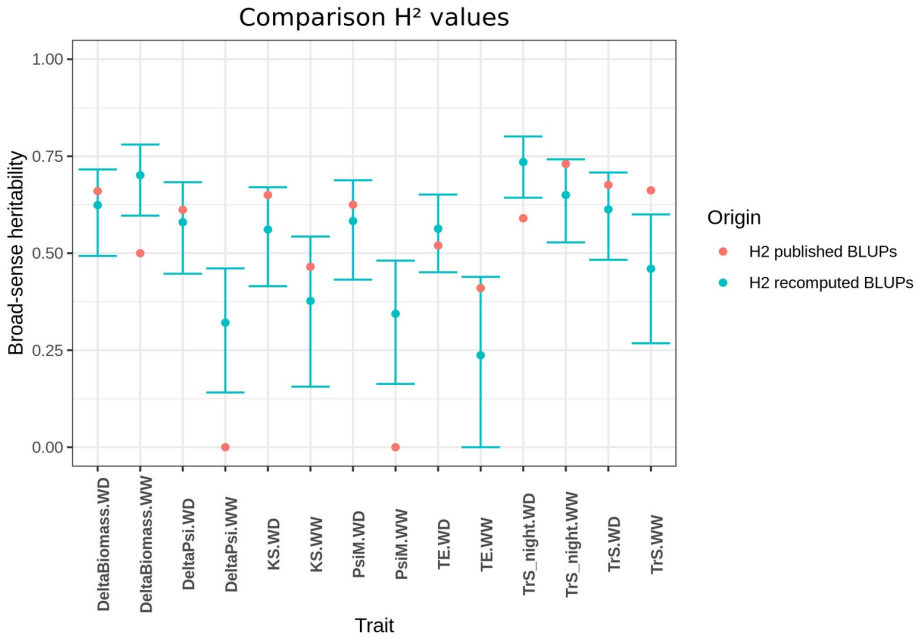


Figure S3: Comparison between published and re-computed broad-sense heritability estimates
Broad-sense heritability values estimated by Coupel-Ledru *et al.* (2014, 2016) are reported in red. Recomputed

heritability values are reported in blue with the 95% confidence interval.

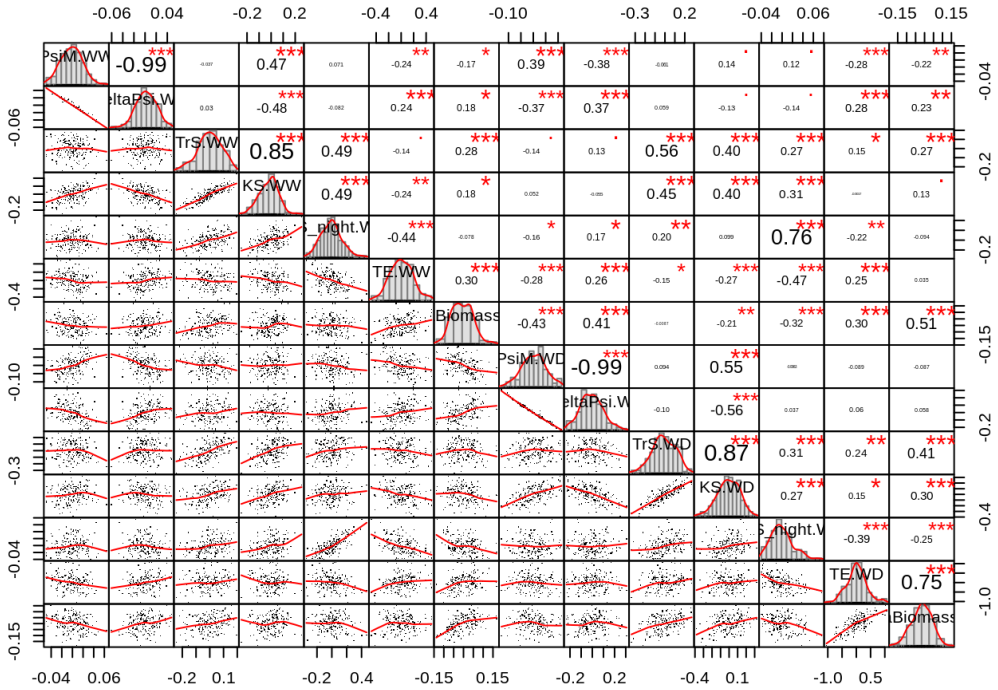


Figure S4: Correlation and distribution of BLUPs for experimental data
Lower panel: pairwise comparison of genotypic BLUP distributions with bivariate scatterplots with fitted lines (with local polynomes).
Upper panel: Pearson's correlation values and significant *p-values* (“.” < 0.1, “*” < 0.05, “**” < 0.01, “***” < 0.001).
Diagonal panel: distributions of genotypic BLUPs for each trait
Trait order: *PsiM.WW*, *DeltaPsi.WW*, *TrS.WW, KS.WW, TrS_night.WW, TE.WW, DeltaBiomass.WW, PsiM.WD, DeltaPsi.WD, TrS.WD, KS.WD, TrS_night.WD, TE.WD, DeltaBiomass.WD.*

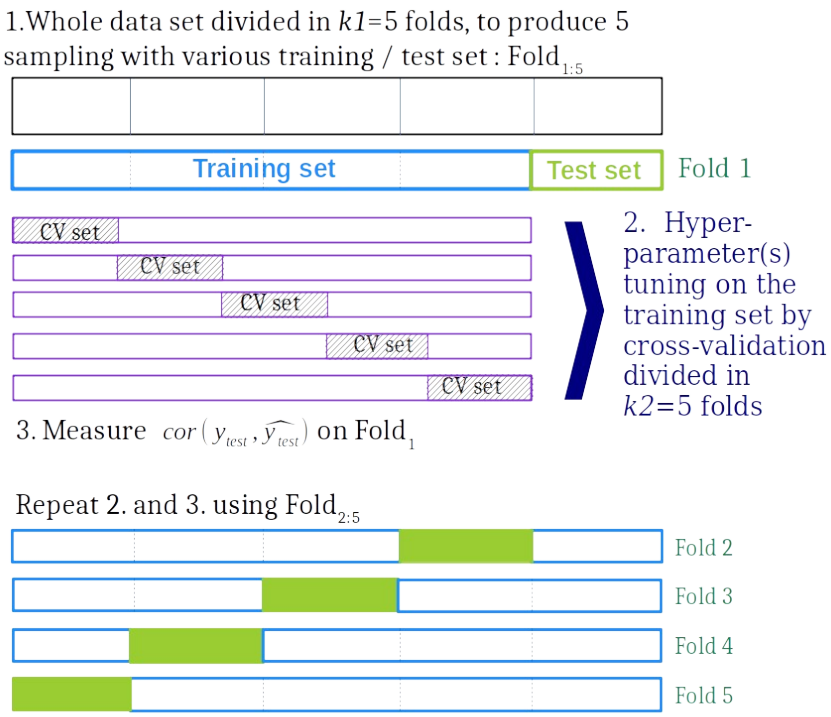


Figure S5: Nested cross-validation scheme

Procedure for cross-validation we applied on simulated and real data for prediction.

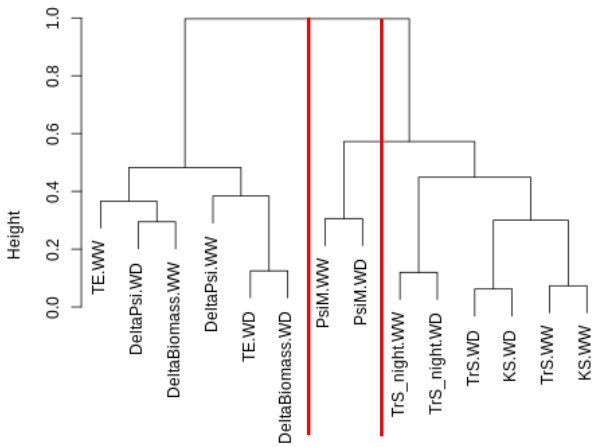


Figure S6: Hierarchical clustering of PhenoArch traits

Clusterisation was done with R/hclust function from stats package. Red bars were added and represent the three cluster of traits we have done.

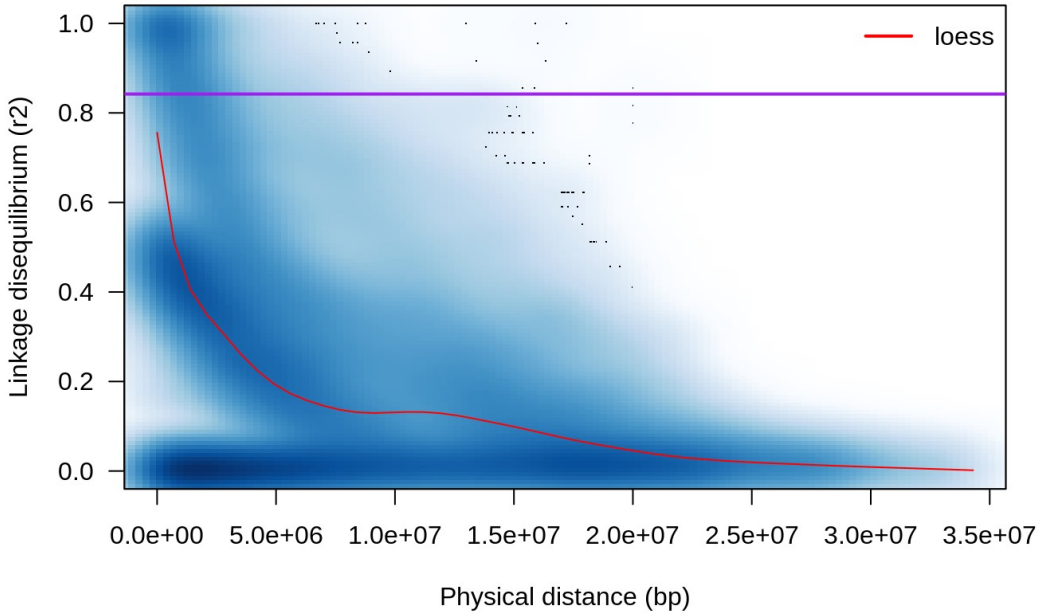


Figure S7: Linkage Disequilibrium pattern in Syrah x Grenache progeny

Linkage disequilibrium was calculated for each chromosome separately with the 3,961 SNP markers with estimLd function and plotted with plotLd function (both from rutilstimflutre R package). The purple line corresponds to the 95% quantile of the LD distribution (where r^2=0.84).

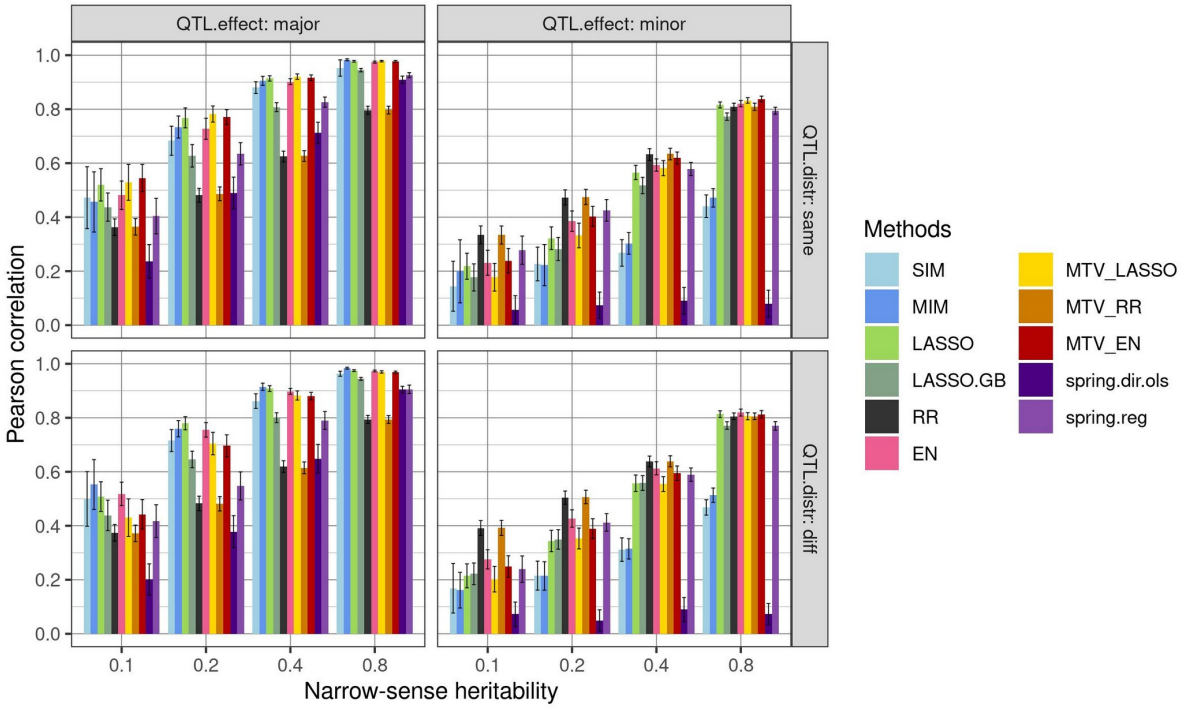


Figure S8: Prediction accuracy for all methods on simulated data

Genomic prediction accuracy (Pearson's correlation between predicted and simulated genotypic values) of 11 methods applied to 3,961 markers and two simulated traits in a bi-parental population, with four heritability values and four QTL configurations (effect x distribution among traits). **Major**: 2 QTLs; **minor**: 50 QTLs; **same**:

QTLs at the same positions for both traits; **diff**: QTLs at different positions between traits. For each heritability value and configuration, prediction accuracy was averaged over 100 values (2 traits x 10 simulation replicates x 5 cross-validation folds). The error bar corresponds to the 95% confidence interval around the mean. In comparison with Figure 1, four methods are added: SIM, LASSO.GB, MTV_RR and *spring.dir.ols*.

TableS9_results_all-methods.tsv

Table S9: Genomic prediction performance on simulated data for all methods and configurations

Genomic prediction performance, as measured by different metrics, of 11 methods applied to 3,961 markers and two simulated traits in a bi-parental population, with four heritability values and four QTL configurations (effect x distribution among traits). **Major**: 2 QTLs; **minor**: 50 QTLs; **same**: QTLs at the same positions for both traits; **diff**: QTLs at different positions between traits. Mean results are averaged over 2 traits, 10 simulation replicates x 5 cross-validation folds. For each metric measured, the standard error is also given (last columns, suffix "SE").

This table can be downloaded at: supplementary.

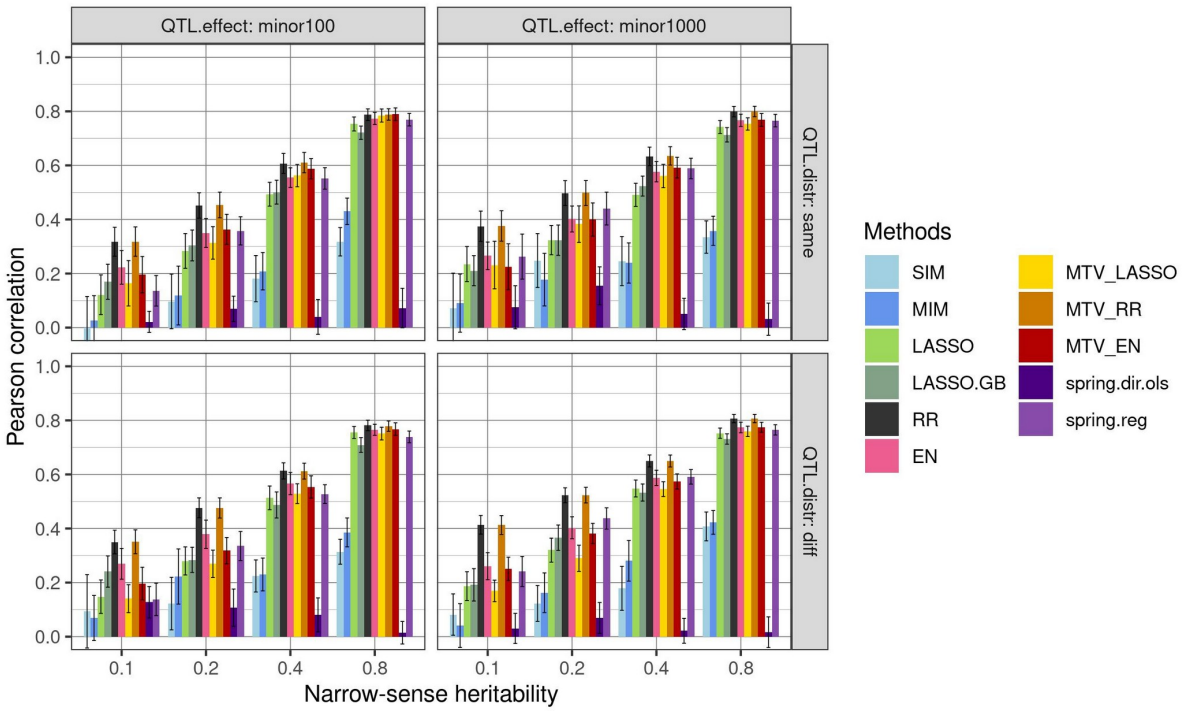


Figure S10: Prediction accuracy with 100 and 1000 simulated QTLs

Genomic prediction accuracy (Pearson's correlation between predicted and true genotypic values) of 11 methods applied to 3,961 markers and phenotypic data simulated in a bi-parental population with four heritability values and QTL configurations (effect x distribution among traits). **minor100**: 100 QTLs; **minor1000**: 1000 QTLs; **same**: QTLs at the same positions for both traits; **diff**: QTLs at different positions between traits. Results are averaged over 2 traits x 5 simulation replicates x 5 cross-validation folds. The error bar corresponds to the 95% confidence interval around the mean.

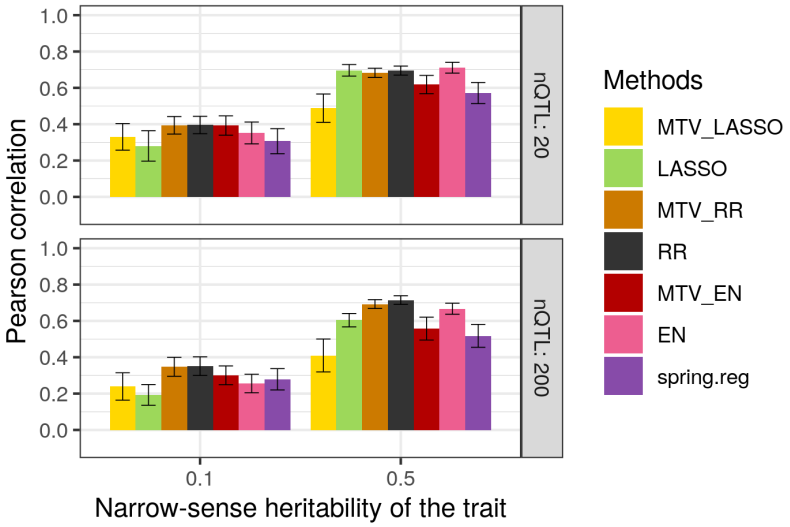


Figure S11: Prediction accuracy at different heritability values among traits

Genomic prediction accuracy (Pearson's correlation between predicted and simulated genotypic values) of seven methods applied to 3,961 markers and two simulated traits in a bi-parental population, with different heritability values (0.1 and 0.5), 20 or 200 QTLs, a genetic correlation of 0.5 and a genetic variance of 1. QTLs were the same for both traits. For each trait, results are averaged over 10 simulation replicates x 5 cross-validation folds. The error bar corresponds to the 95% confidence interval around the mean.

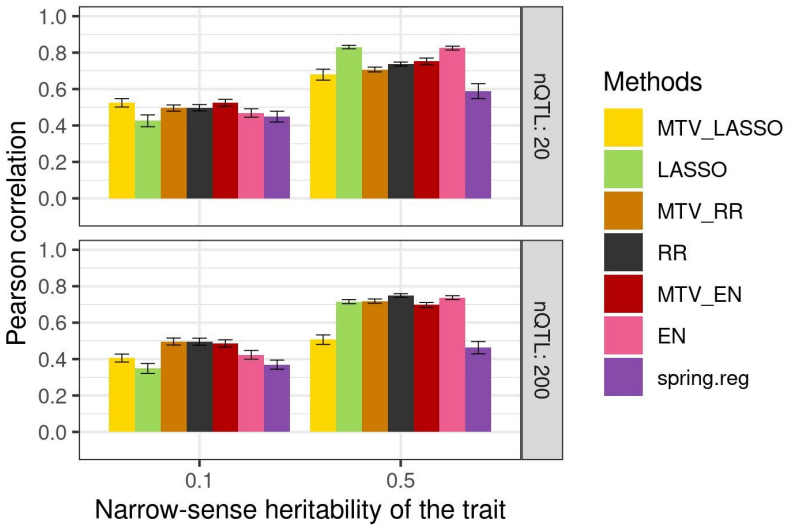


Figure S12: Prediction accuracy with simulated data from Jia & Jannink (2012)

Genomic prediction accuracy (Pearson's correlation between predicted and simulated genotypic values) of seven methods applied to simulated data from Jia and Jannink (2012). The simulated data comprised two traits, with 500 observations for 2000 bi-allelic markers, 20 or 200 QTLs and 24 simulation replicates for each simulation parameter set. The two simulated traits had different heritability values (0.1 and 0.5), a genetic correlation of 0.5, and no error correlation.

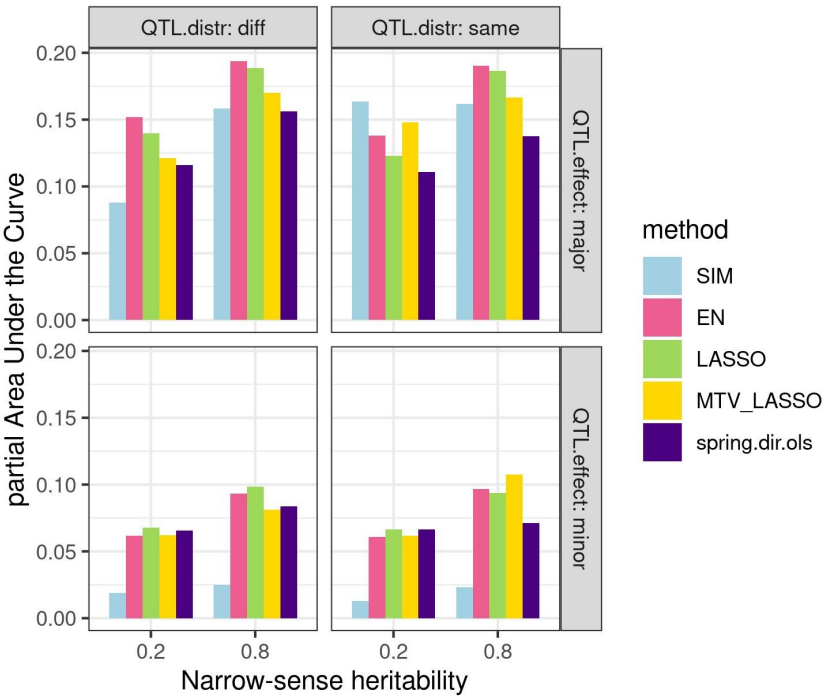


Figure S13: partial Area Under the ROC Curve for FPR up to 0.1

Partial area under the curve (pAUC), calculated from the ROC curve (Figure 2) for FPR ranging from 0 to 0.1, for five methods with FPR reaching 0.1 applied to 3,961 markers and two simulated traits in a bi-parental population, with two heritability values and four QTL configurations (effect x distribution among traits).
major: 2 QTLs; **minor**: 50 QTLs; **same**: QTLs at the same positions for both traits; **diff**: QTLs at different positions between traits.

TableS14_all_results_mean_real-data-pharch.tsv

Table S14: Genomic prediction performance on experimental grapevine data for all traits, and methods

Genomic prediction performance, as measured by different metrics, of 11 methods for 14 traits in a bi-parental grapevine population. Mean results are averaged over 10 cross-validation replicates x five cross-validation folds. As in **Table S9**, the standard error is also given (last columns, suffix "SE") for each metric measured.

This table can be downloaded at: supplementary.

TableS15_selected_markers_pharch_all-resp_PVE.tsv

Table S15: Results of marker selection for all traits and methods

genetic.position : genetic position on the SNP map (in cM)
PVE: percentage of variance explained. Marker selection is not extended to LD.

This table can be downloaded at: supplementary.

FigureS16_selection_all-resp.pdf

Figure S16: Genetic position of selected markers per trait and for all methods

Each marker selected by a given method is represented by a colored point, the color indicating the number of methods that have selected that specific marker. The boxes correspond to chromosomes and the *x*-axis to the position along the genetic map (in cM). Marker selection is not extended to LD.

This figure can be downloaded at: supplementary.

FigureS17_MapGBS_most-reliable.pdf

Figure S17: Most reliable QTL intervals displayed along the SNP genetic map

SNP genetic map of the Syrah x Grenache progeny with QTL intervals represented with colored rectangles, hatched for traits under water deficit condition. Trait name and percentage of variance explained are indicated next to intervals. New QTLs are indicated by an asterisk before trait name.

This figure can be downloaded at: supplementary.

TableS18_reliable_QTLs_information.tsv

Table S18: Complementary information about most reliable QTLs

For each reliable QTL, the table indicates the genetic and physical position from which the QTL interval has been defined, as well as physical and genetic QTL intervals. Methods which have selected the QTL and percentage of variance explained are indicated and we added a logical column (CIM SSR found) which report if the QTL has been published in Coupel-Ledru *et al.* (2014, 2016). If several markers were considered as highly reliable for the same QTL, the one with the largest PVE (from **Table S15**) was kept, then if several markers were at the same PVE, the first one was kept. An interval of physical position may be indicated corresponding to minimum and maximum of markers at the same genetic position.

This table can be downloaded at: supplementary.

FigureS19_boxplots_most_reliable_QTLs.pdf

Figure S19: Distribution of genotypic BLUPs according to genotypic class for each highly reliable QTL

Boxplot of genotypic BLUPs according to genotypic class in Syrah x Grenache progeny for the 25 most reliable QTLs. If several markers were considered as highly reliable for the same QTL, the one with the largest PVE (from **Table S15**) was kept, then if several markers were at the same PVE, the first one was kept. QTL position in centiMorgan and percentage of variance explained are indicated. The red line represent the mean of the genotypic BLUP.

This figure can be downloaded at: supplementary.

TableS20_genes_highly_reliable_QTLs.csv

Table S20: List of genes underlying the most highly reliable QTLs with their annotations

qtl.id: unique identifier of the QTL defined its location chromosome, start and end position in bp.
traits.id: list of traits overlapping in this QTL.
trait.of.interest: trait for which the QTL was determined as highly reliable within the overlapping traits.
gene.loc.in.QTL.peak: yes if the gene is located into the peak of the QTL, no if not.
vitvi.geneid: gene name retrieved from the reference Vitis genome 12X.v2
refseq.geneid: **ncbi.gene.id**, **other.alias**: other identifiers for this gene, retrieved from the VCost.v3 annotation (Canaguier *et al.* 2017).
chr, **start.IGGP12Xv2** and **end.IGGP12Xv2**: chromosome, start position (in bp) and end position (in bp) of the gene on the reference *Vitis* genome 12X.v2.
refseq.genedesc: putative function of the gene retrieved from NCBI, when available.
gene.name: gene name retrieved from TAIR and UniProt, when available.
TAIR.locus: locus in Arabidopsis for the gene's homolog retrieved from TAIR and UniProt, when available.
full.description: complete description of the gene function, retrieved from TAIR and UniProt, when available.
mean.expression.leaves, **min.expression.leaves**, **max.expression.leaves**: mean, min and max expression of the gene in the leaves retrieved from GREAT. Mean, max and min were calculated from all RNA-seq data available on the platform for leaves samples. They are expressed in log(RPKM+1) expression of the gene in the shoots retrieved from GREAT. Mean, max and min were calculated from all RNA-seq data available on the platform for shoots samples.

This table can be downloaded at: supplementary.

TableS21_genes_QTL_chr4.csv

Table S21: List and functional classification of annotated genes underlying the highly reliable QTL detected on chromosome 4 for TrS_night.WD_LG4, TE.WW_LG4, TE.WD_LG4, DeltaBiomass.WW_LG4, DeltaBiomass.WD_LG4

qtl.id: unique identifier of the QTL defined its location chromosome, start and end position in bp.
traits.id: list of traits overlapping in this QTL.
vitvi.geneid: gene name retrieved from the reference Vitis genome 12X.v2
refseq.geneid: ncbi.gene.id, other.alias: other identifiers for this gene, retrieved from the VCost.v3 annotation (Canaguier *et al* .2017).
chr, start.IGGP12Xv2 and end.IGGP12Xv2: chromosome, start position (in bp) and end position (in bp) of the gene on the reference Vitis genome 12X.v2.
peak.info: information about the location of the peak of the QTL; "in peak" means that the gene falls into the peak of the QTL; "closest to peak" means that the gene is outside the peak of the QTL but is the closest gene to the peak.
refseq.genedesc: putative function of the gene retrieved from NCBI, when available.
gene.name: gene name retrieved from TAIR and UniProt, when available.
TAIR.locus: locus in Arabidopsis for the gene's homolog retrieved from TAIR and UniProt, when available.
full.description: complete description of the gene function, retrieved from TAIR and UniProt, when available.
mean.expression.leaves, min.expression.leaves, max.expression.leaves: mean, min and max expression of the gene in the leaves retrieved from GREAT. Mean, max and min were calculated from all RNA-seq data available on the platform for leaves samples.
expression of the gene in the shoots retrieved from GREAT. Mean, max and min were calculated from all RNA-seq data available on the platform for shoots samples.
integrated.function: integrated function at the organ or plant level, if it is explicitly quoted in the column "full.description".
cell.or.mol.function: general cellular or molecular function of the gene, when relevant.
constituant.or.specific.molecule: when relevant, name of the constituent or specific molecule related to the cellular or molecular function.
activation.condition: condition triggering the gene activation, when explicitly quoted in the column "full.description".

relation.to.traits.of.interest: « related » if related to the traits in this QTL (*TrS_night, TE, DeltaBiomass*), "unrelated" if not.

This table can be downloaded at: supplementary.

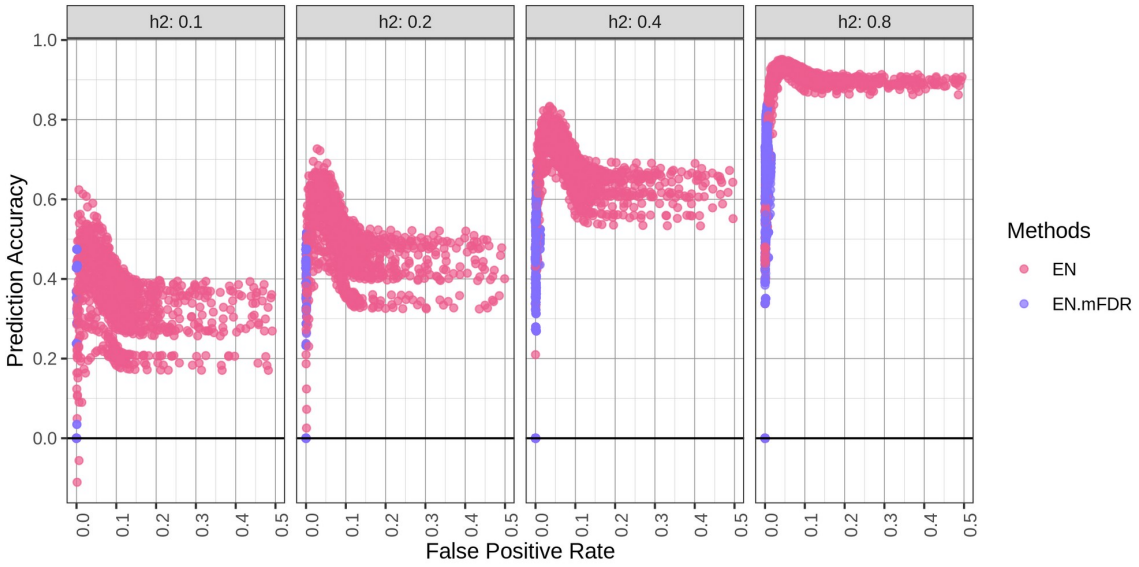


Figure S22: Prediction accuracy according to False Positive Rate for EN and EN.mFDR

For each varying parameter for both methods (i.e. lambda for EN and mFDR for EN.mFDR), we reported the

False Positive Rate according to the prediction accuracy, calculated on simulated data, for the two **minor** configurations and the four heritability values.

Appendix C

Appendix Chapter 3

C.1 Supplementary information for article II

Supplementary information

Across-population genomic prediction in grapevine opens up promising prospects for breeding

Authors: Charlotte Brault, Vincent Segura, Patrice This, Loïc Le Cunff, Timothée Flutre, Pierre François, Thierry Pons, Jean-Pierre Péros, Agnès Doligez

Table of contents

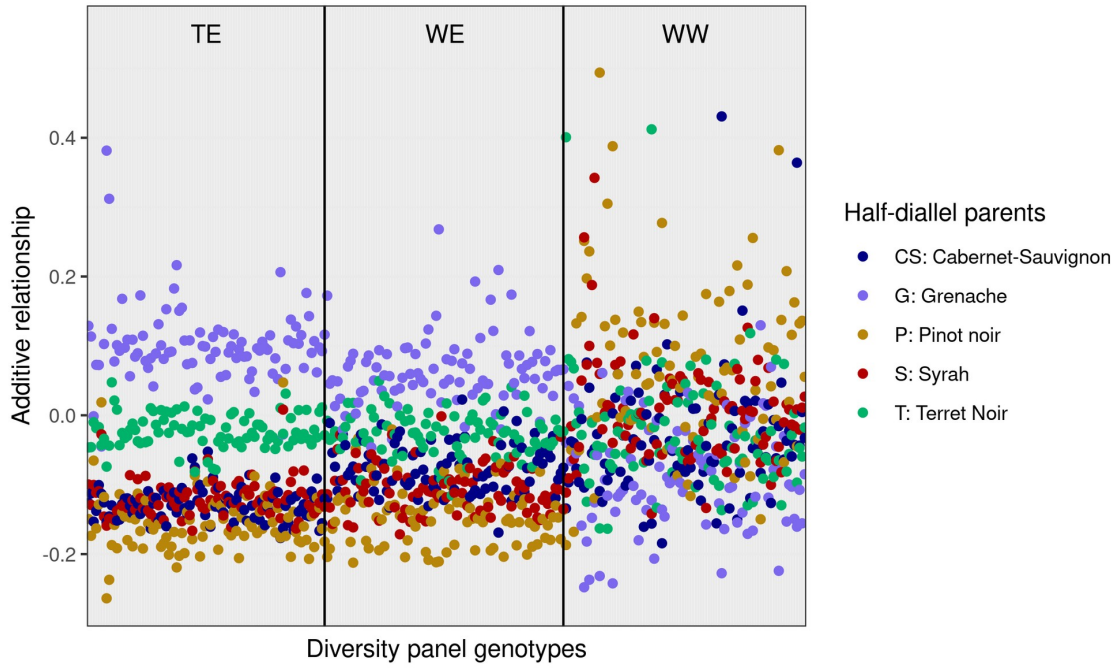


Figure S1 Additive relationship between half-diallel parents and diversity panel cultivars.

Additive relationship was calculated with the VanRaden (2008) method, with 32,894 SNPs. For each diversity panel genotype, its additive relationship with the five half-diallel parents was computed. Each point corresponds to one cultivar in the diversity panel, ordered by on the *x*-axis subpopulation (TE: Table East, WE: Wine East, WW: Wine West), and colored according to the corresponding half-diallel parent genotype.

trait	years	missing.data.percent	transformation	fixed.effects	random.effects
mal.ripe	2015	38	NA	none	geno, cross, x
tar.ripe	2015	38	NA	none	geno, cross, x
shik.ripe	2015	42	log	block	geno, cross, x
shiktar.ripe	2015	42	log	block	geno, cross, x
maltar.ripe	2015	38	log	none	geno, cross, x
verday	2013, 2014, 2017	36	NA	block, year, block:year	geno, cross, x, geno:year, cross:year
samplday	2013, 2014, 2015	28	NA	block, year	geno, cross, geno:year, year:x
vermatu	2013	33	NA	block	geno, cross, x
clucomp	2013, 2014, 2015	30	NA	block, year, block:year	geno, cross, geno:year, cross:year, year:x
nbclu	2013, 2014, 2015	28	sqrt	block, year, block:year	geno, cross, x:y, geno:year, cross:year, year:x
mcl	2013, 2014,	28	NA	year	geno, cross, x:y, geno:year, cross:year,

	2015				year:x, year:y
mcwi	2013	32	NA	block	geno, cross
mcw	2013, 2014, 2015	29	log	block, year, block:year	geno, cross, x:y, geno:year, cross:year, year:x
mbw	2013, 2014, 2015	29	sqrt	block, year, block:year	geno, cross, x:y, geno:year, cross:year
vigour	2014, 2015	16	log	block, year, block:year	geno, cross, x:y, geno:year, cross:year, year:x, year:y

trait	var.geno	var.cross	H2	H2.low	H2.high	CV.geno	CV.geno.low	CV.geno.high	var.cross.geno
mal.ripe	460	85.94	0.58	0.55	0.61	0.17	0.14	0.19	0.16
tar.ripe	230	50.68	0.70	0.68	0.73	0.16	0.14	0.18	0.18
shik.ripe	0.379	0.40	0.80	0.78	0.82	0.26	0.22	0.31	0.51
shiktar.ripe	0.377	0.43	0.84	0.82	0.86	0.09	0.08	0.10	0.53
maltar.ripe	0.042	0.02	0.74	0.71	0.76	0.73	0.55	1.00	0.31
verday	13.1	3.84	0.8	0.79	0.81	0.02	0.01	0.02	0.23
samplday	36.2	30.63	0.82	0.81	0.84	0.02	0.02	0.03	0.46
vermatu	48.0	19.30	0.65	0.63	0.68	0.18	0.15	0.20	0.29
clucomp	1.53	0.16	0.81	0.80	0.83	0.23	0.21	0.24	0.09
nbclu	0.742	0.15	0.84	0.83	0.85	0.20	0.18	0.22	0.17
mcl	2.25	0.53	0.80	0.78	0.81	0.12	0.11	0.13	0.19
mcwi	1.05	0.58	0.50	0.46	0.52	0.12	0.10	0.15	0.35
mcw	0.083	0.05	0.83	0.82	0.84	0.05	0.05	0.06	0.36
mbw	0.019	0.02	0.92	0.91	0.93	0.1	0.09	0.11	0.51
vigour	0.096	0.01	0.77	0.76	0.80	0.12	0.11	0.13	0.06

Table S1 Information on mixed model selection and genotypic BLUP estimation for 15 traits in the half-diallel population.

trait: see abbreviations meaning in Methods section.

years: years in which the trait was phenotyped.

missing.data.percent: percentage of missing raw phenotypic data, relative to the full initial design.

transformation: transformation applied to raw phenotypic data, before model selection and BLUP estimation (NA: none; sqrt: square root; log: neperian logarithm).

fixed.effects / random.effects: fixed and random effects kept in the final selected model, as defined in Methods section.

var.geno: intra-cross genotypic variance estimate (534 to 624 levels, depending on the trait).

var.cross: cross variance estimate (10 levels)

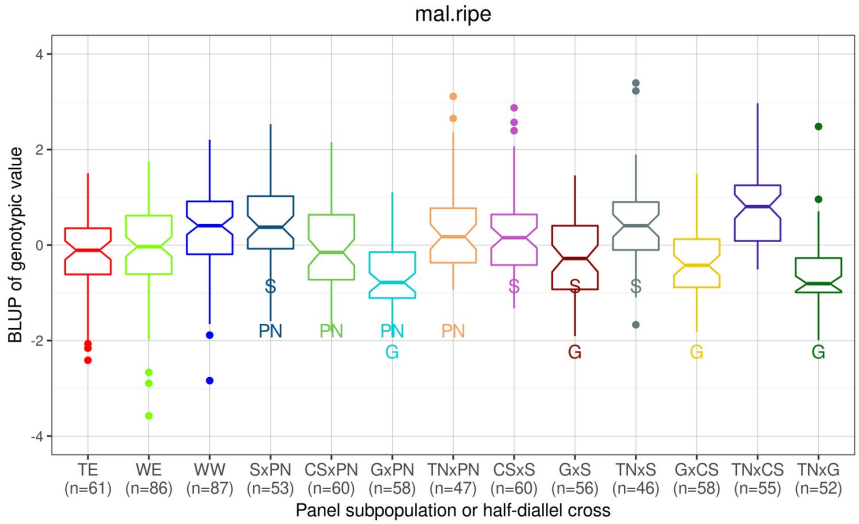
H2 / H2.low / H2.high: broad-sense heritability estimate and its confidence interval bounds computed through bootstrapping.

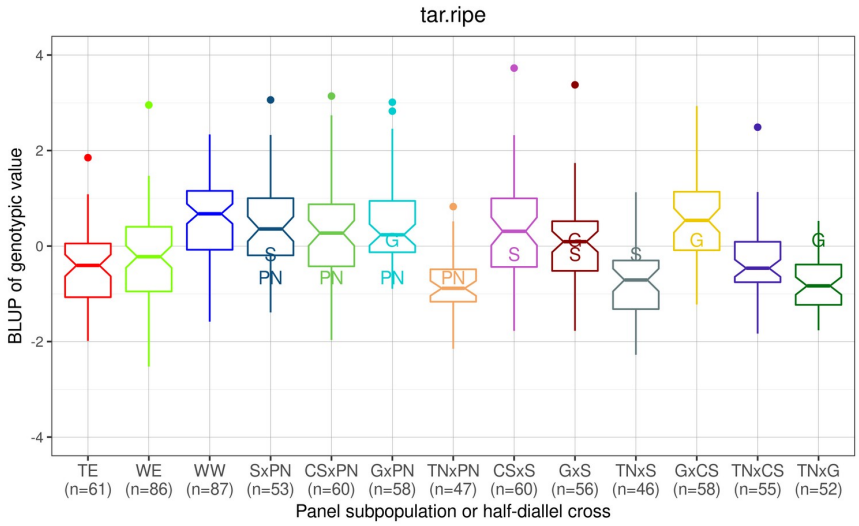
CV.geno / CV.geno.low / CV.geno.high: estimated coefficient of variation of genotypic effect and its confidence interval bounds.

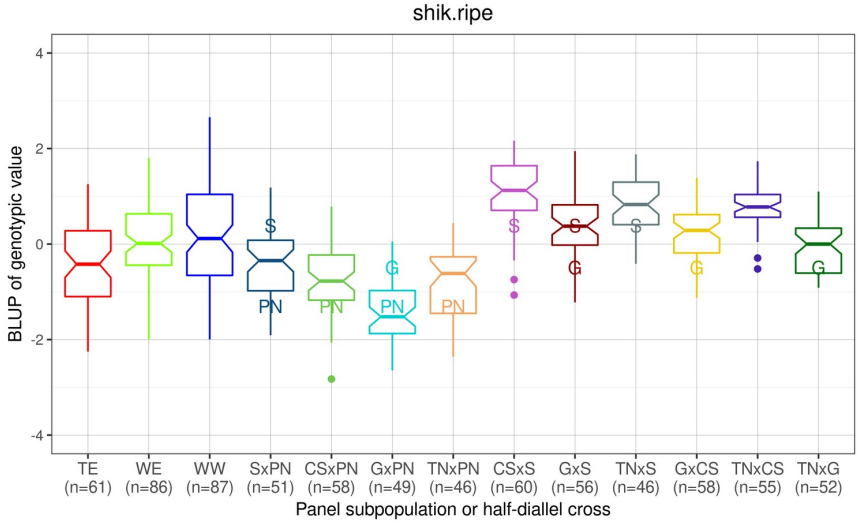
var.Cross.geno: $\sigma_C^2/\left(\sigma_C^2+\sigma_G^2\right)$ ratio, as defined in Methods section.

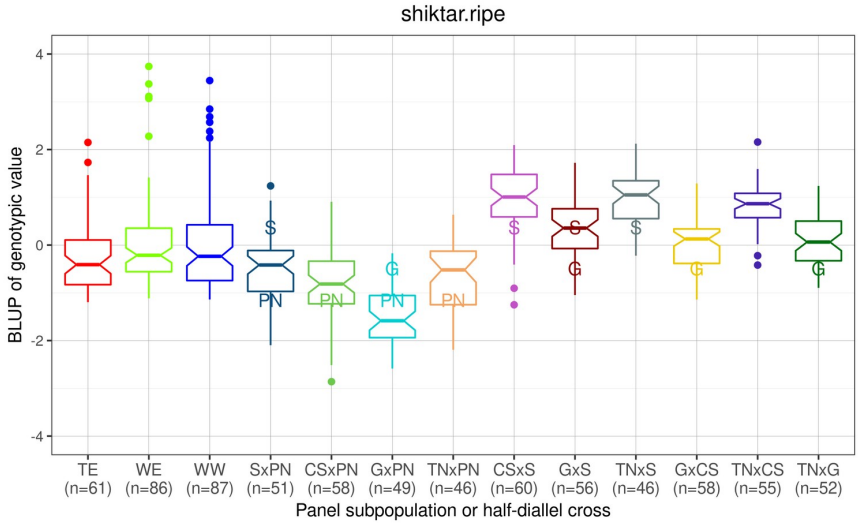
Figure S2 Per cross broad-sense heritability in the half-diallel.

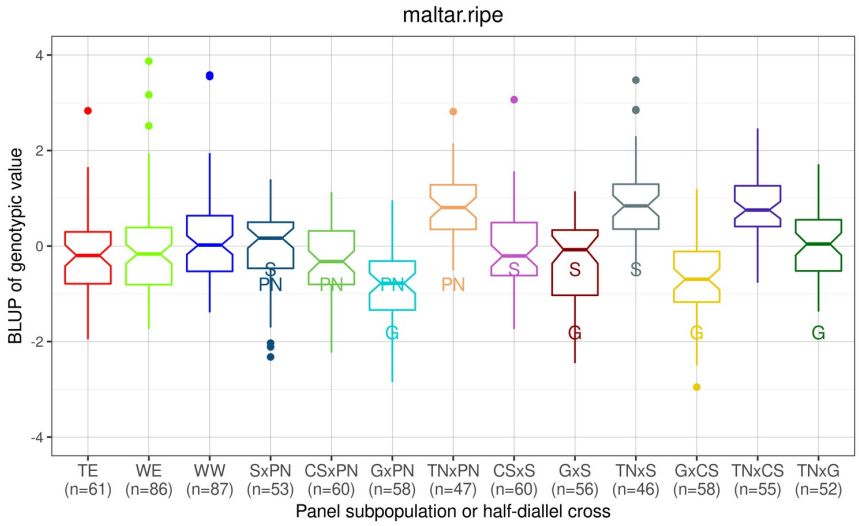
For crosses and traits, see abbreviation meaning in Methods section.

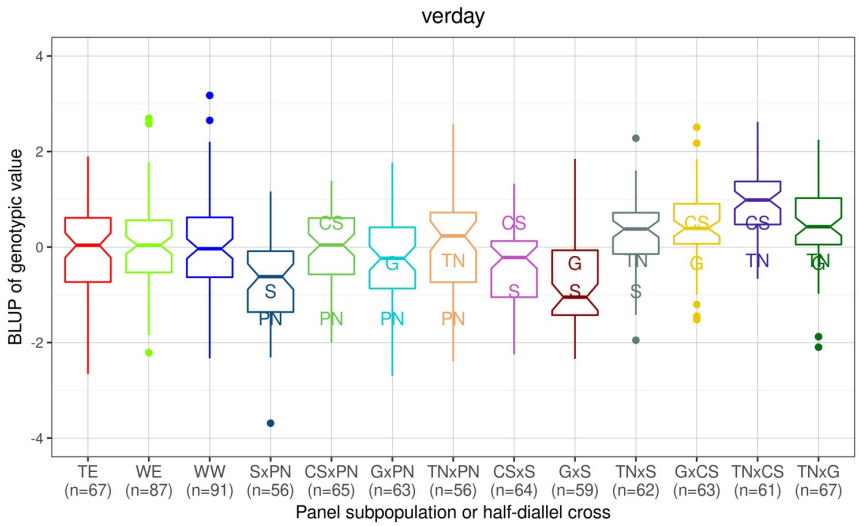


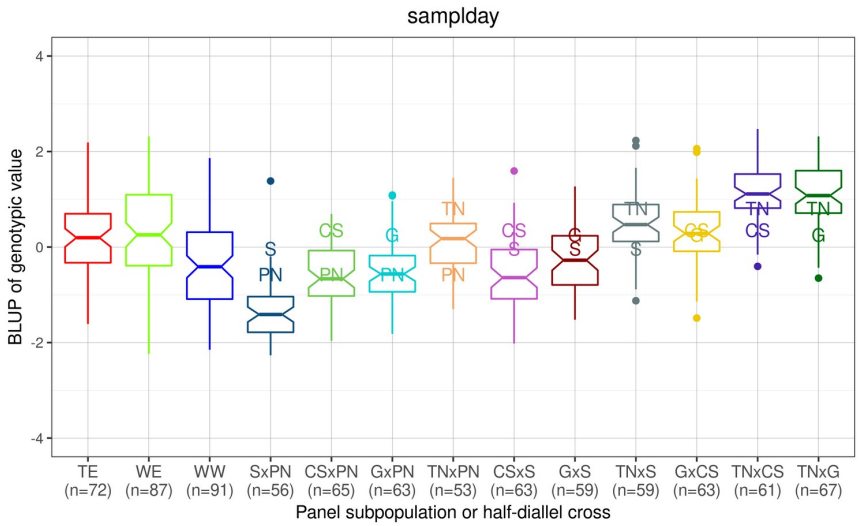

tar.ripe
BLUP of genotypic value
4
2
0
-2
-4
TE (n=61)
WE (n=86)
WW (n=87)
SxPN (n=53)
CSxPN (n=60)
GxPN (n=58)
TNxPN (n=47)
CSxS (n=60)
GxS (n=56)
TNxS (n=46)
GxCS (n=58)
TNxCS (n=55)
TNxG (n=52)
Panel subpopulation or half-diallel cross

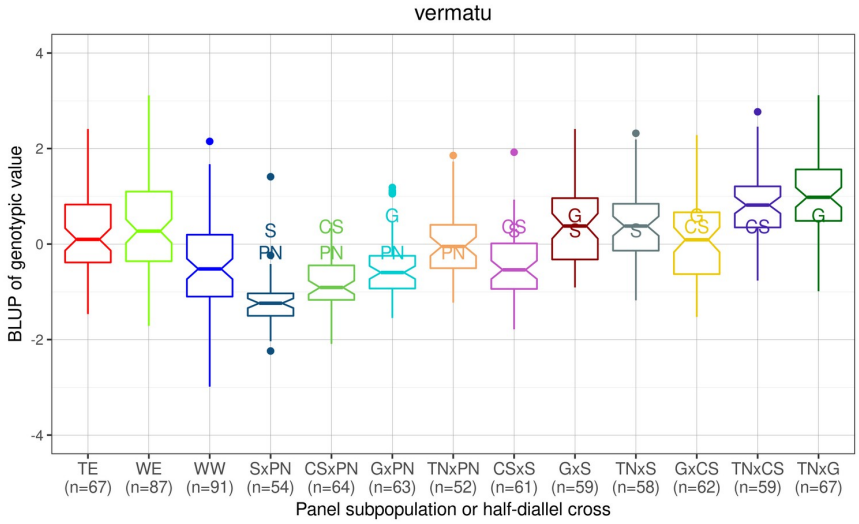

shik.ripe
BLUP of genotypic value
4
2
0
-2
-4
TE (n=61)
WE (n=86)
WW (n=87)
SxPN (n=51)
CSxPN (n=58)
GxPN (n=49)
TNxPN (n=46)
CSxS (n=60)
GxS (n=56)
TNxS (n=46)
GxCS (n=58)
TNxCS (n=55)
TNxG (n=52)
Panel subpopulation or half-diallel cross

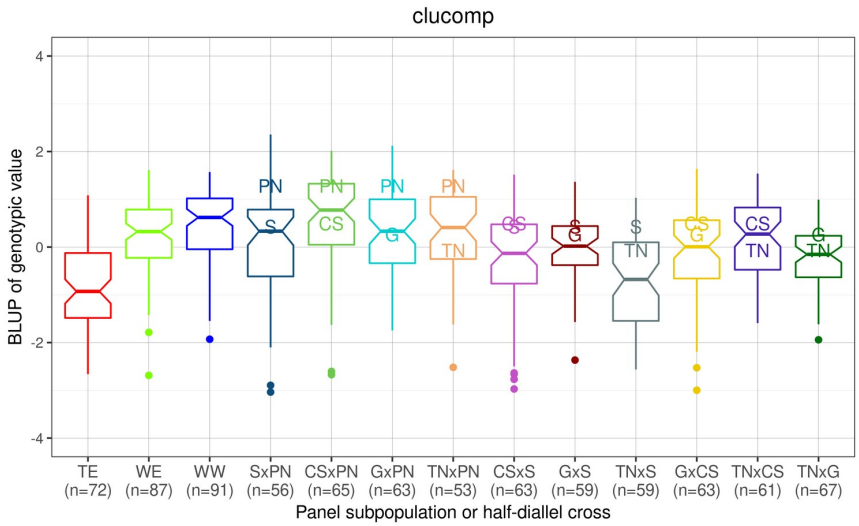

shiktar.ripe
BLUP of genotypic value
4
2
0
-2
-4
TE (n=61)
WE (n=86)
WW (n=87)
SxPN (n=51)
CSxPN (n=58)
GxPN (n=49)
TNxPN (n=46)
CSxS (n=60)
GxS (n=56)
TNxS (n=46)
GxCS (n=58)
TNxCS (n=55)
TNxG (n=52)
Panel subpopulation or half-diallel cross

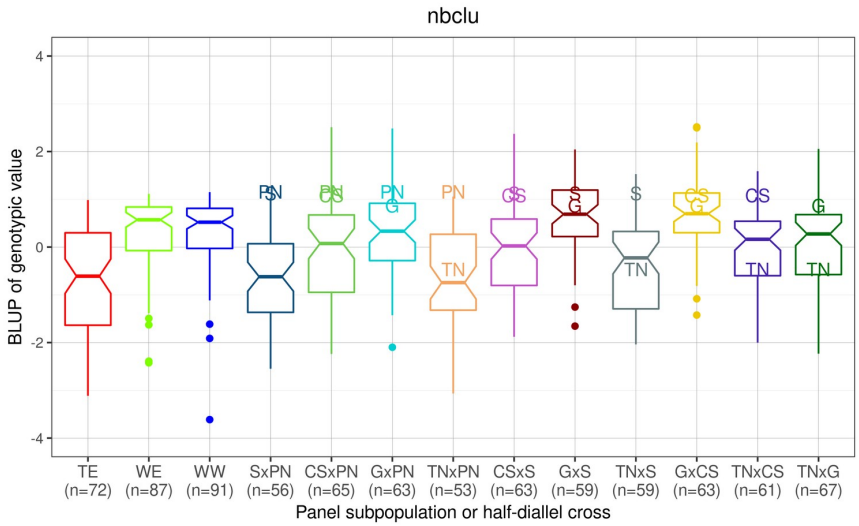

maltar.ripe
BLUP of genotypic value
4
2
0
-2
-4
TE (n=61)
WE (n=86)
WW (n=87)
SxPN (n=53)
CSxPN (n=60)
GxPN (n=58)
TNxPN (n=47)
CSxS (n=60)
GxS (n=56)
TNxS (n=46)
GxCS (n=58)
TNxCS (n=55)
TNxG (n=52)
Panel subpopulation or half-diallel cross

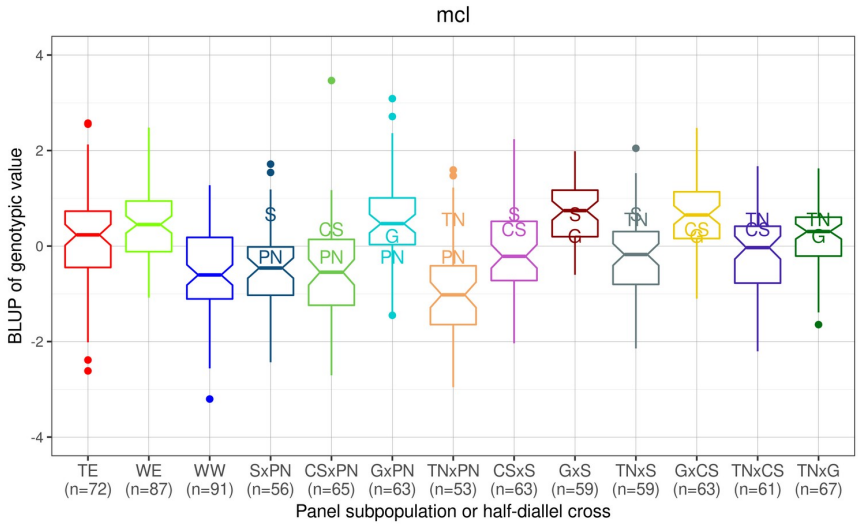

verday
BLUP of genotypic value
4
2
0
-2
-4
TE (n=67)
WE (n=87)
WW (n=91)
SxPN (n=56)
CSxPN (n=65)
GxPN (n=63)
TNxPN (n=56)
CSxS (n=64)
GxS (n=59)
TNxS (n=62)
GxCS (n=63)
TNxCS (n=61)
TNxG (n=67)
Panel subpopulation or half-diallel cross

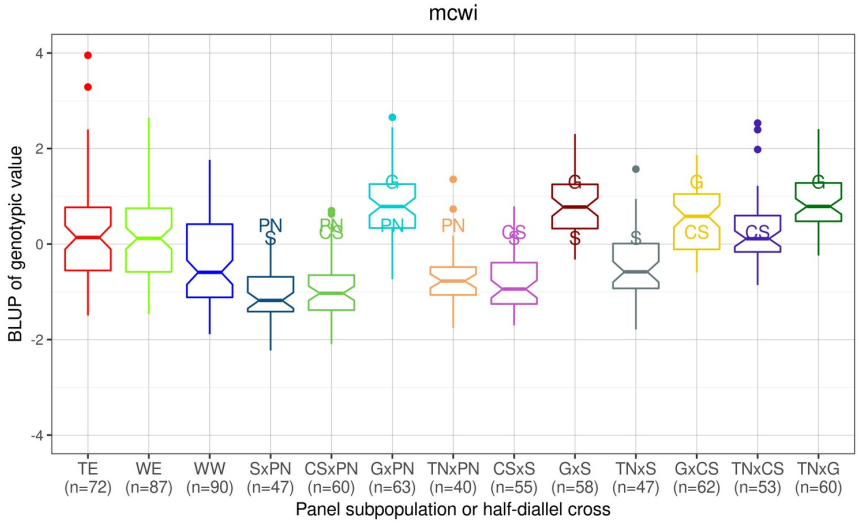

samplday
BLUP of genotypic value
4
2
0
-2
-4
TE (n=72)
WE (n=87)
WW (n=91)
SxPN (n=56)
CSxPN (n=65)
GxPN (n=63)
TNxPN (n=53)
CSxS (n=63)
GxS (n=59)
TNxS (n=59)
GxCS (n=63)
TNxCS (n=61)
TNxG (n=67)
Panel subpopulation or half-diallel cross

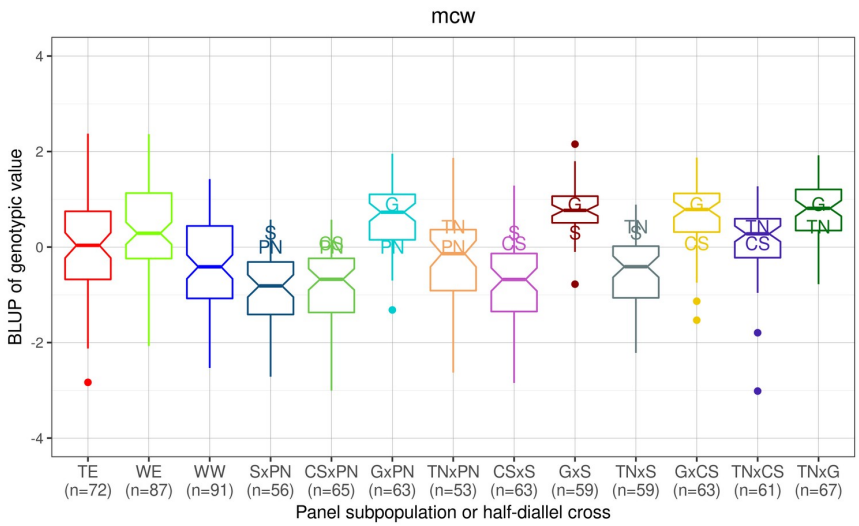

vermatu
BLUP of genotypic value
4
2
0
-2
-4
TE (n=67)
WE (n=87)
WW (n=91)
SxPN (n=54)
CSxPN (n=64)
GxPN (n=63)
TNxPN (n=52)
CSxS (n=61)
GxS (n=59)
TNxS (n=58)
GxCS (n=62)
TNxCS (n=59)
TNxG (n=67)
Panel subpopulation or half-diallel cross


clucomp
BLUP of genotypic value
4
2
0
-2
-4
TE (n=72)
WE (n=87)
WW (n=91)
SxPN (n=56)
CSxPN (n=65)
GxPN (n=63)
TNxPN (n=53)
CSxS (n=63)
GxS (n=59)
TNxS (n=59)
GxCS (n=63)
TNxCS (n=61)
TNxG (n=67)
Panel subpopulation or half-diallel cross


nbclu
BLUP of genotypic value
4
2
0
-2
-4
TE (n=72)
WE (n=87)
WW (n=91)
SxPN (n=56)
CSxPN (n=65)
GxPN (n=63)
TNxPN (n=53)
CSxS (n=63)
GxS (n=59)
TNxS (n=59)
GxCS (n=63)
TNxCS (n=61)
TNxG (n=67)
Panel subpopulation or half-diallel cross


mcl
BLUP of genotypic value
4
2
0
-2
-4
TE (n=72)
WE (n=87)
WW (n=91)
SxPN (n=56)
CSxPN (n=65)
GxPN (n=63)
TNxPN (n=53)
CSxS (n=63)
GxS (n=59)
TNxS (n=59)
GxCS (n=63)
TNxCS (n=61)
TNxG (n=67)
Panel subpopulation or half-diallel cross


mcwi
BLUP of genotypic value
4
2
0
-2
-4
TE (n=72)
WE (n=87)
WW (n=90)
SxPN (n=47)
CSxPN (n=60)
GxPN (n=63)
TNxPN (n=40)
CSxS (n=55)
GxS (n=58)
TNxS (n=47)
GxCS (n=62)
TNxCS (n=53)
TNxG (n=60)
Panel subpopulation or half-diallel cross


mcw
BLUP of genotypic value
4
2
0
-2
-4
TE (n=72)
WE (n=87)
WW (n=91)
SxPN (n=56)
CSxPN (n=65)
GxPN (n=63)
TNxPN (n=53)
CSxS (n=63)
GxS (n=59)
TNxS (n=59)
GxCS (n=63)
TNxCS (n=61)
TNxG (n=67)
Panel subpopulation or half-diallel cross

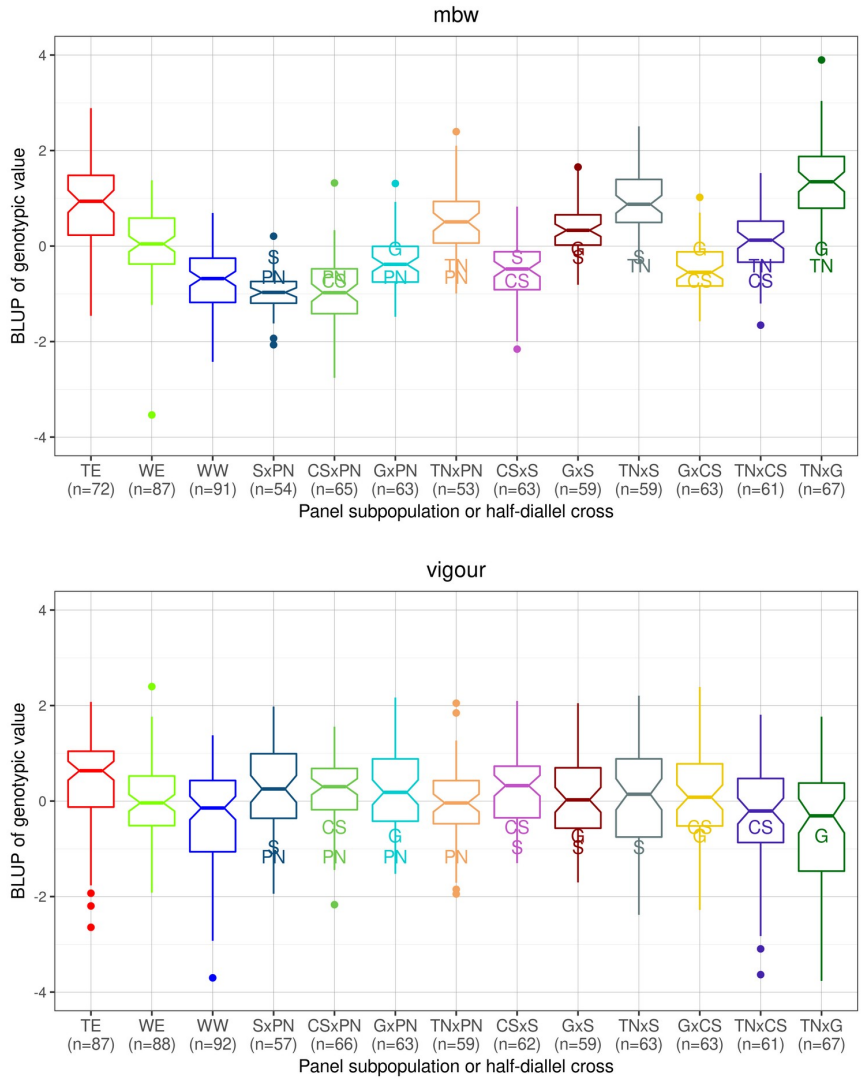


Figure S3 Distribution of genotypic value estimates (BLUPs) for 15 traits, in each diversity panel subpopulation and each half-diallel cross.

Subpopulation or cross size is indicated below its name. All BLUPs were centered and scaled separately for each subpopulation or cross. TE: Table East, WE: Wine East, WW: Wine West. For traits, see abbreviations meaning in Methods section. Values of the parents are indicated by their initials as defined in **Table S4**.

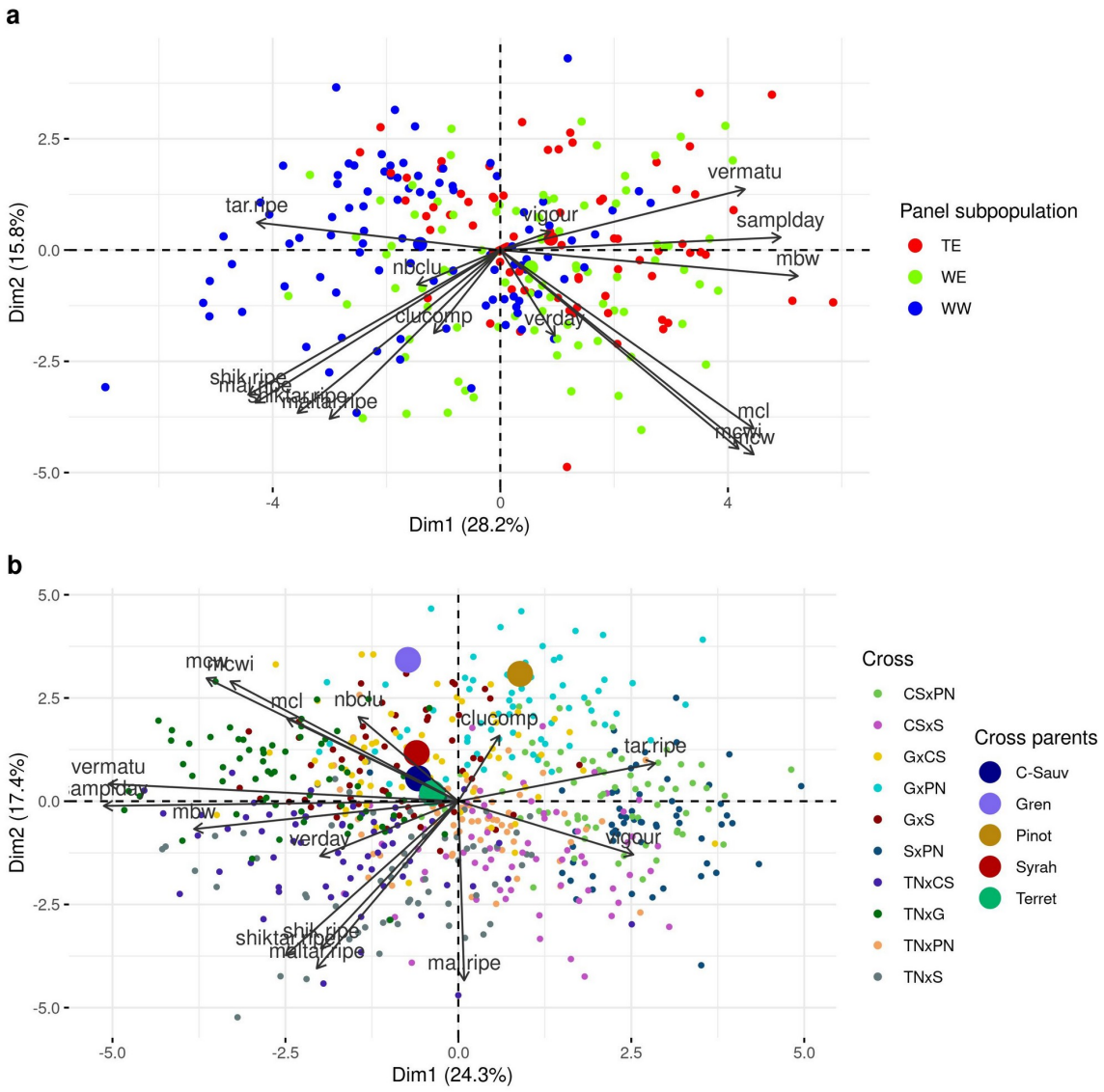


Figure S4 PCA applied to genotypic BLUPs for the 15 traits.

a: In the diversity panel population (TE: Table East, WE: Wine East, WW: Wine West); b: In the half-diallel population. For traits, see abbreviations meaning in Methods section.

Parent 1 genotype	Parent 2 genotype	Progeny genotypes (expected proportion in parenthesis)	Parental average genotype
0	**0**	**0** (1)	**0**
0	**1**	**0** (0.5) / **1** (0.5)	**0.5**
0	**2**	**1** (1)	**1**

1	**1**	**0** (0.25) / **1** (0.5) / **2** (0.25)	**1**
1	**2**	**1** (0.5) / **2** (0.5)	**1.5**
2	**2**	**2** (1)	**2**

Table S2 Computation of parental average genotypes.

For each cross and locus, parent genotypes (0, 1 or 2 for each parent) were used to derive the expected proportion of each genotype in the progeny under Mendelian segregation and from there, the parental average genotype. Probabilities are indicated in parentheses.

a:

Cross	S1a-RR	S1a-LASSO	S1b-RR	S1b-LASSO	S2-RR	S2-LASSO
CSxPN	**0.33**	0.03	**0.4**	0.38	0.22	**0.35**
CSxS	**0.68**	0.22	0.67	**0.74**	**0.72**	0.32
GxCS	**0.52**	0.43	0.42	**0.45**	**-0.08**	-0.23
GxPN	**0.6**	0.52	**0.56**	0.49	-0.45	**-0.3**
GxS	**0.77**	0.58	**0.71**	0.65	0	**0.11**
SxPN	**0.29**	0.06	0.37	**0.42**	**0.61**	0.35
TNxCS	**0.57**	0.15	**0.61**	0.59	0.41	**0.63**
TNxG	**0.72**	0.52	**0.85**	0.81	0.21	**0.61**
TNxPN	**0.38**	0.22	**0.67**	0.56	-0.02	**0.01**
TNxS	**0.78**	0.36	**0.78**	0.77	**0.25**	0.22

b:

Trait	S1a-RR	S1a-LASSO	S1b-RR	S1b-LASSO	S2-RR	S2-LASSO
mal.ripe	**0.9**	0.82	**0.91**	0.82	**0.73**	-0.01
tar.ripe	**0.95**	0.94	**0.95**	0.91	**0.7**	0.65
shik.ripe	1	1	0.98	**0.99**	0.19	**0.96**
shiktar.ripe	**1**	0.99	0.98	**1**	0.16	**0.56**
maltar.ripe	**0.98**	0.96	0.96	**0.97**	**0.72**	0.03
verday	0.93	**0.95**	**0.92**	0.9	0.66	**0.81**
samplday	0.99	0.99	0.98	0.98	**0.82**	0.59
vermatu	**0.97**	0.95	0.95	0.95	0.63	**0.72**
clucomp	**0.84**	0.55	**0.87**	0.75	0.26	**0.55**
nbclu	0.82	**0.91**	**0.95**	0.91	**-0.64**	-0.75
mcl	0.85	**0.93**	0.86	**0.92**	0.52	**0.76**
mcwi	0.94	0.94	**0.94**	0.92	0.77	**0.83**
mcw	**0.93**	0.79	0.96	0.96	0.74	**0.81**
mbw	**0.96**	0.94	0.96	**0.97**	**0.83**	0.63
vigour	**0.85**	0.67	**0.86**	0.77	-0.15	**-0.09**

Table S3 Predictive ability of cross mean.

Pearson's correlation between the observed cross mean and the one predicted based on parental average genotypes. Values are reported for scenarios 1a, 1b and 2, with RR and LASSO methods. The best PA value for each scenario is indicated in bold.

a: per-cross PA (correlation based on 15 observations); **b** per-trait PA (correlation based on 10 observations).

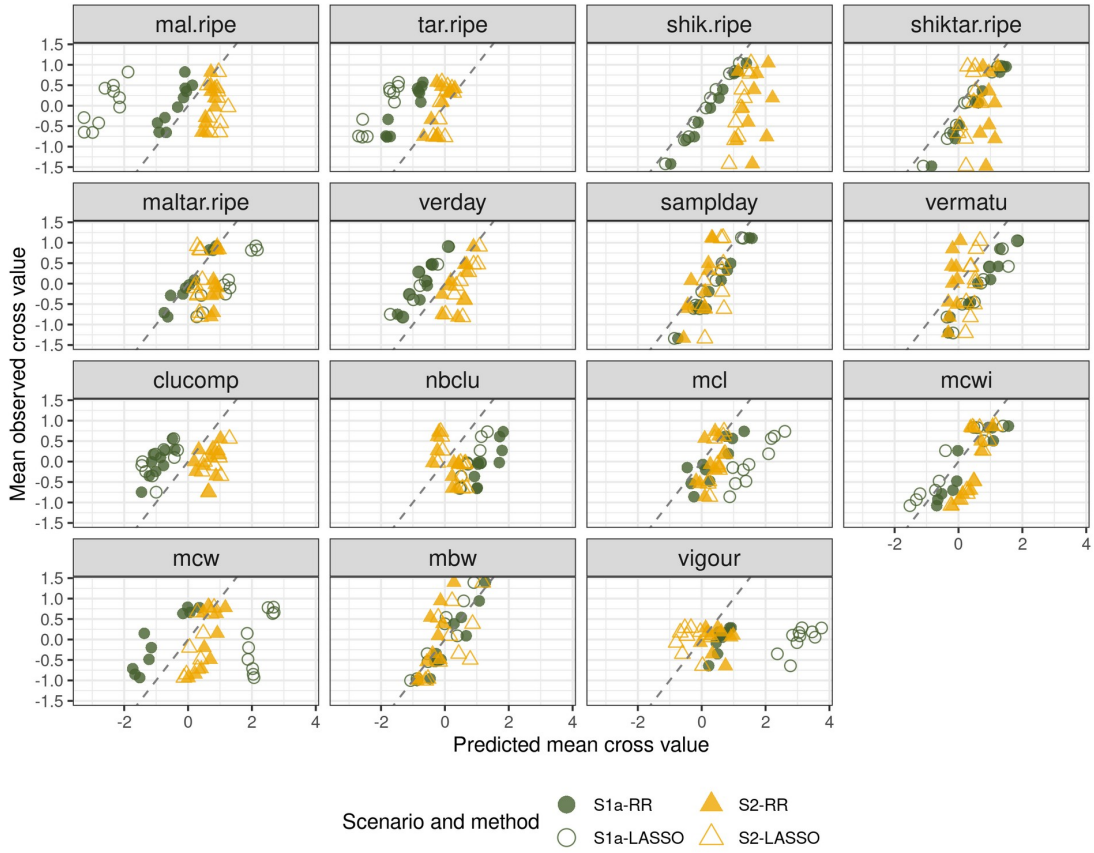
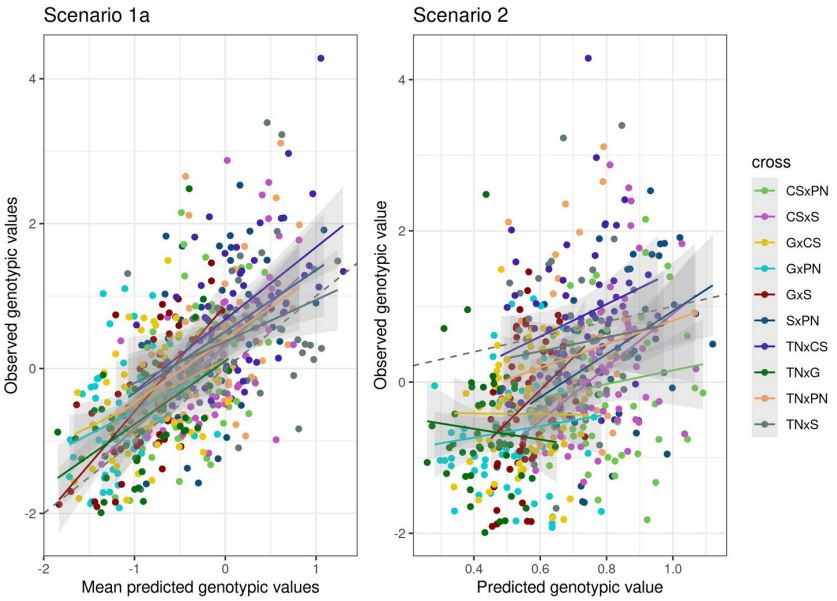


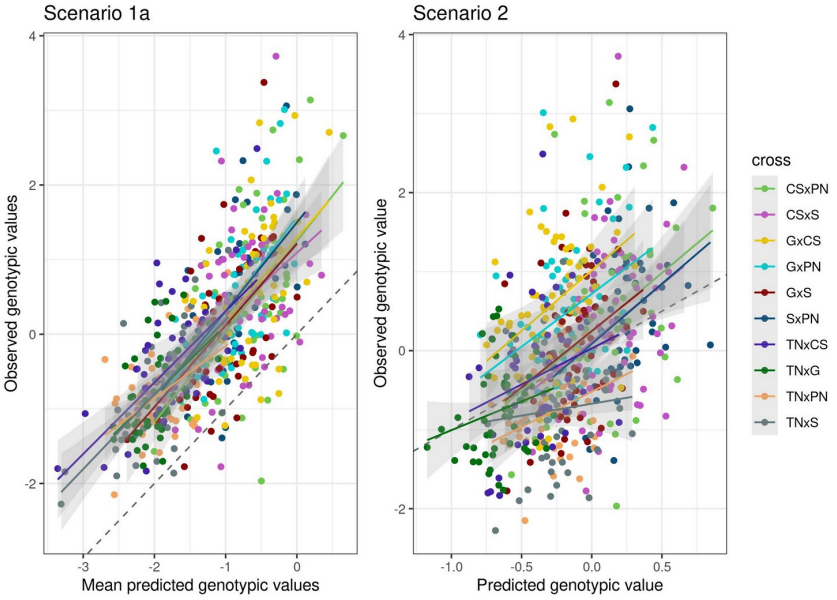
Figure S5 Observed vs predicted cross means for each trait in the half-diallel

Observed vs predicted (based on parental average genotypes) cross mean for each trait in the half-diallel, according to four prediction modalities: with allelic effects estimated in the half-diallel or in the diversity panel (*S1a and S2*, respectively) and with RR or the LASSO. The dashed line indicates the perfect fit (with slope=1 and intercept=0), points deviating from this line indicate bias.

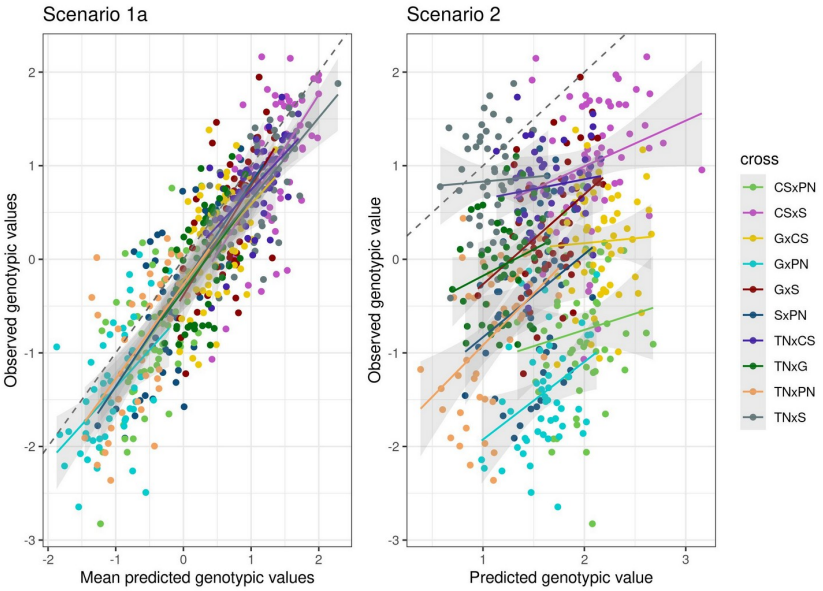
mal.ripe



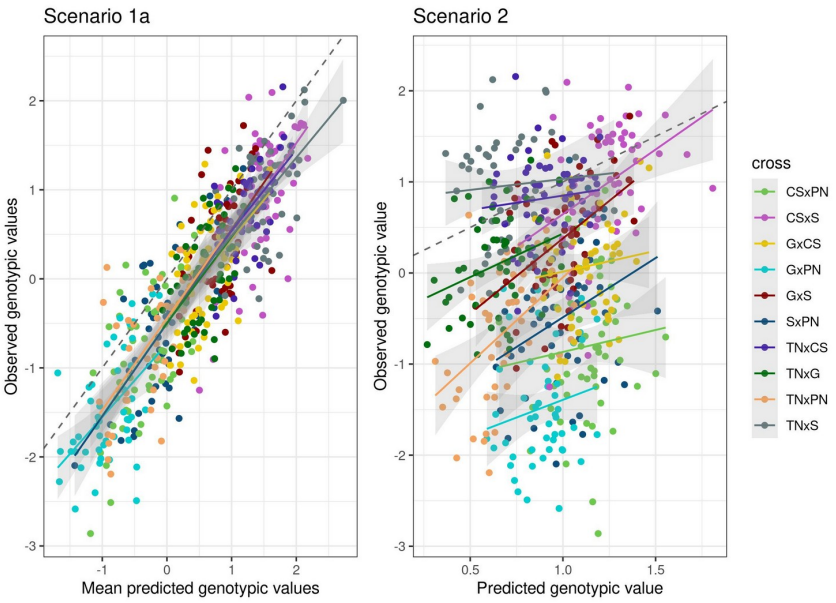
tar.ripe

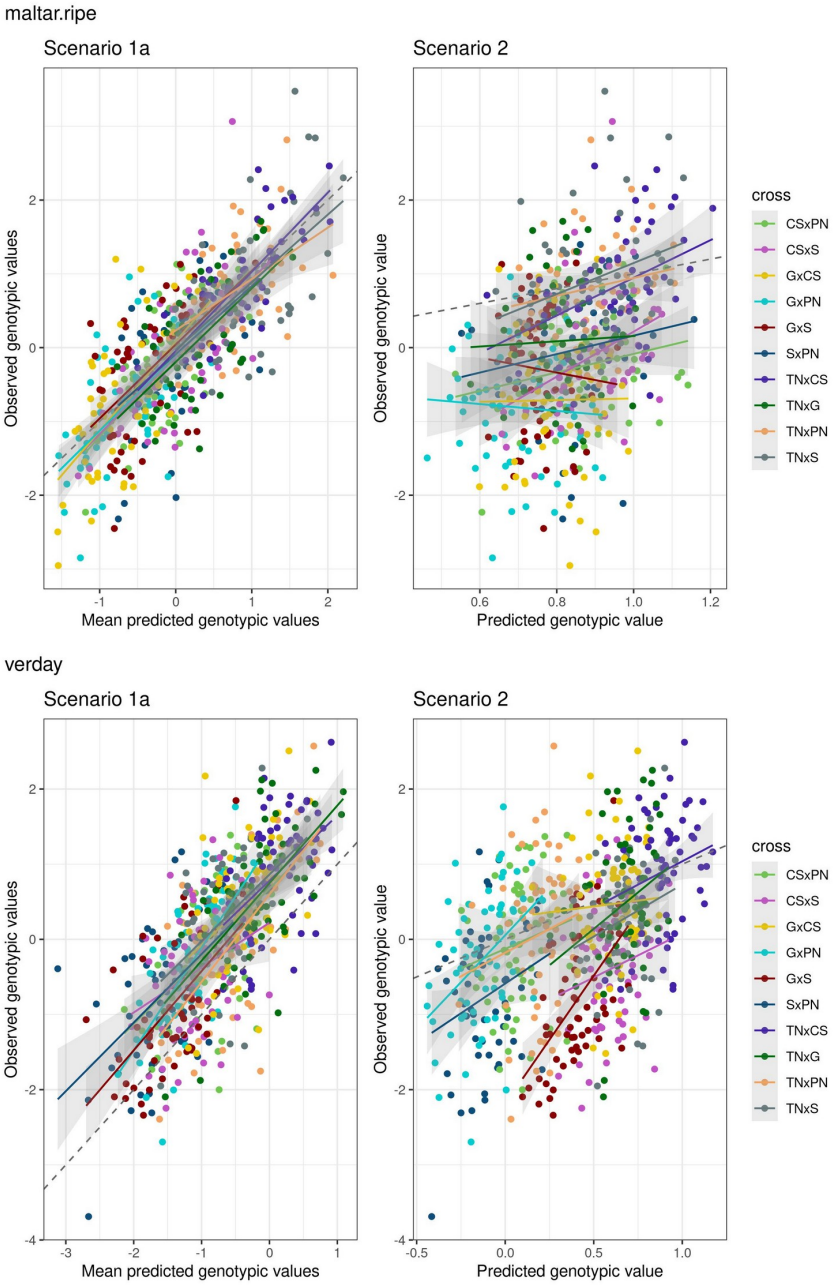


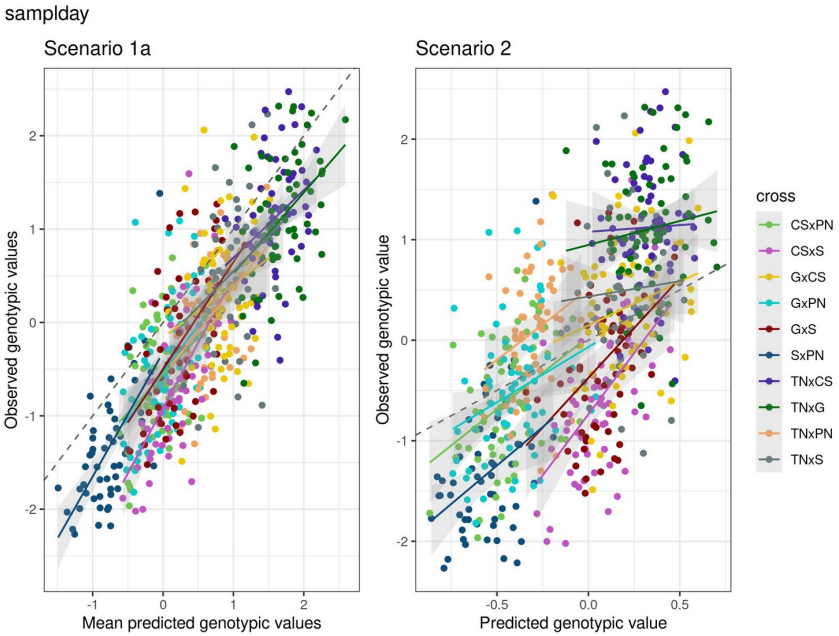
shik.ripe



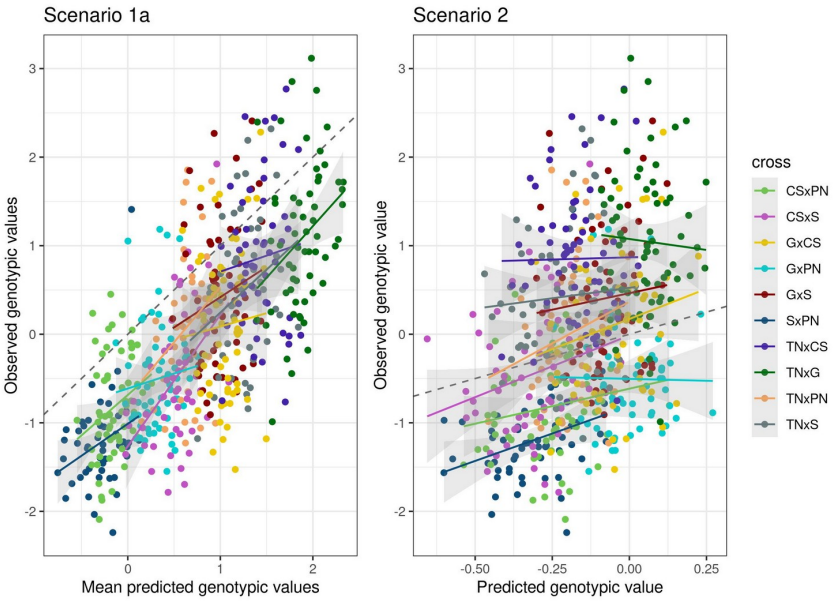
shiktar.ripe



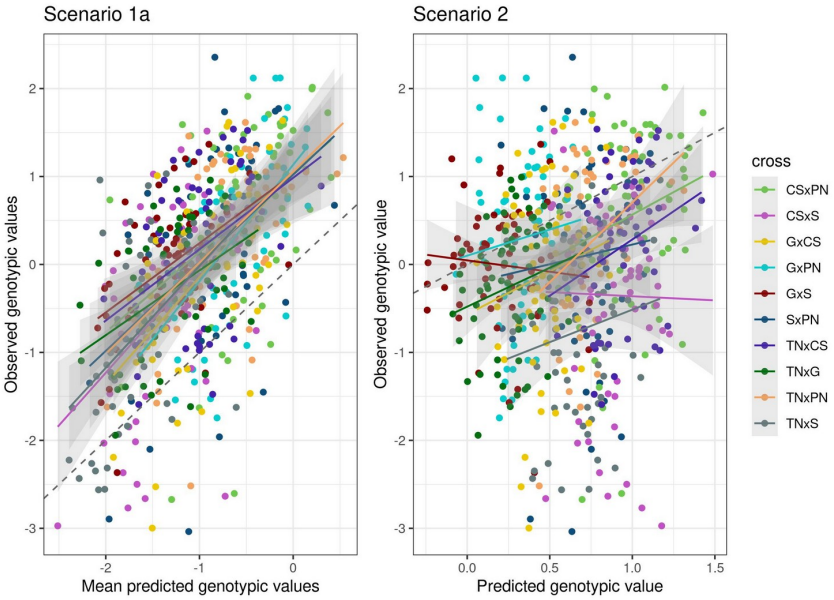

maltar.ripe
Scenario 1a
Scenario 2
Observed genotypic values
Observed genotypic value
Mean predicted genotypic values
Predicted genotypic value
cross
CSxPN
CSxS
GxCS
GxPN
GxS
SxPN
TNxCS
TNxG
TNxPN
TNxS
verday
Scenario 1a
Scenario 2
Observed genotypic values
Observed genotypic value
Mean predicted genotypic values
Predicted genotypic value
cross
CSxPN
CSxS
GxCS
GxPN
GxS
SxPN
TNxCS
TNxG
TNxPN
TNxS


samplday
Scenario 1a
Scenario 2
Observed genotypic values
Mean predicted genotypic values
Observed genotypic value
Predicted genotypic value
cross
CSxPN
CSxS
GxCS
GxPN
GxS
SxPN
TNxCS
TNxG
TNxPN
TNxS

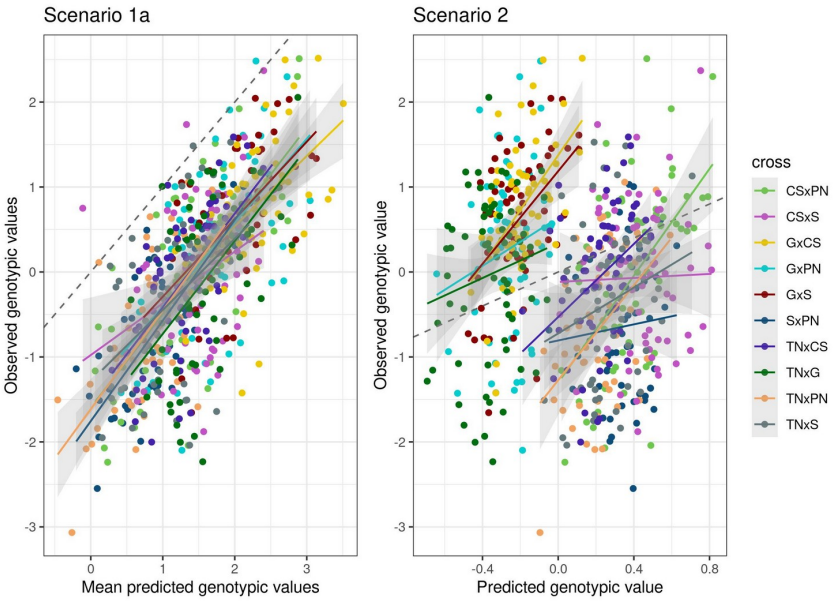
vermatu



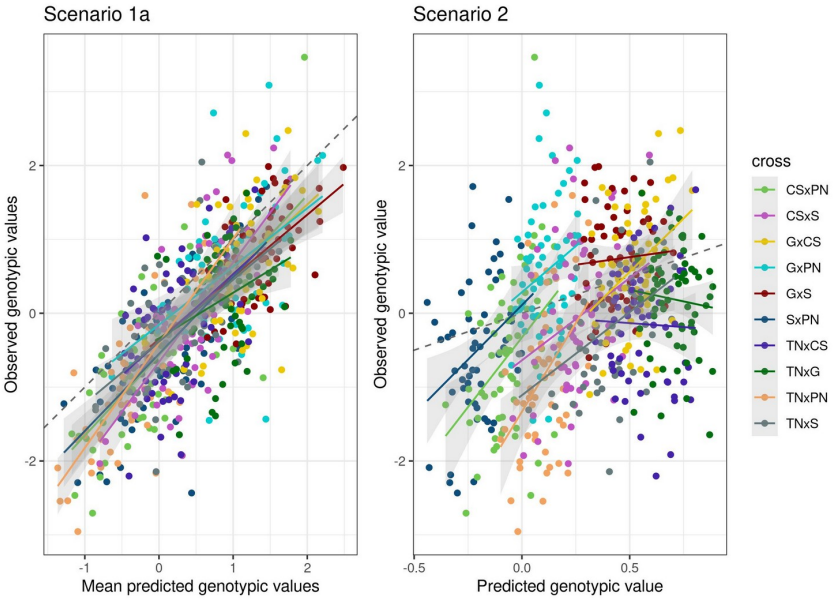
clucomp

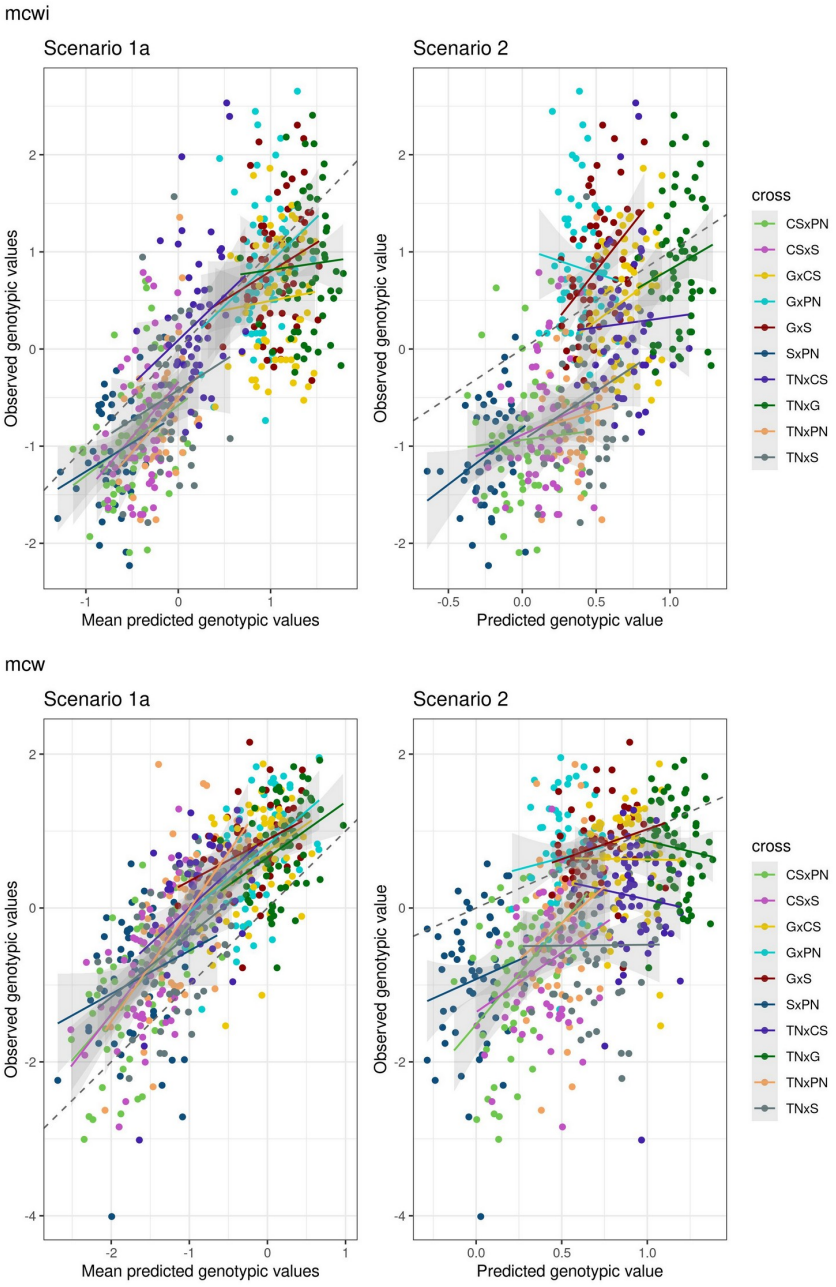


nbclu


Scenario 1a
Scenario 2
Observed genotypic values
Observed genotypic value
Mean predicted genotypic values
Predicted genotypic value
cross
CSxPN
CSxS
GxCS
GxPN
GxS
SxPN
TNxCS
TNxG
TNxPN
TNxS

mcl


Scenario 1a
Scenario 2
Observed genotypic values
Observed genotypic value
Mean predicted genotypic values
Predicted genotypic value
cross
CSxPN
CSxS
GxCS
GxPN
GxS
SxPN
TNxCS
TNxG
TNxPN
TNxS


mcwi
Scenario 1a
Scenario 2
Observed genotypic values
Observed genotypic value
Mean predicted genotypic values
Predicted genotypic value
cross
CSxPN
CSxS
GxCS
GxPN
GxS
SxPN
TNxCS
TNxG
TNxPN
TNxS
mcw
Scenario 1a
Scenario 2
Observed genotypic values
Observed genotypic value
Mean predicted genotypic values
Predicted genotypic value
cross
CSxPN
CSxS
GxCS
GxPN
GxS
SxPN
TNxCS
TNxG
TNxPN
TNxS

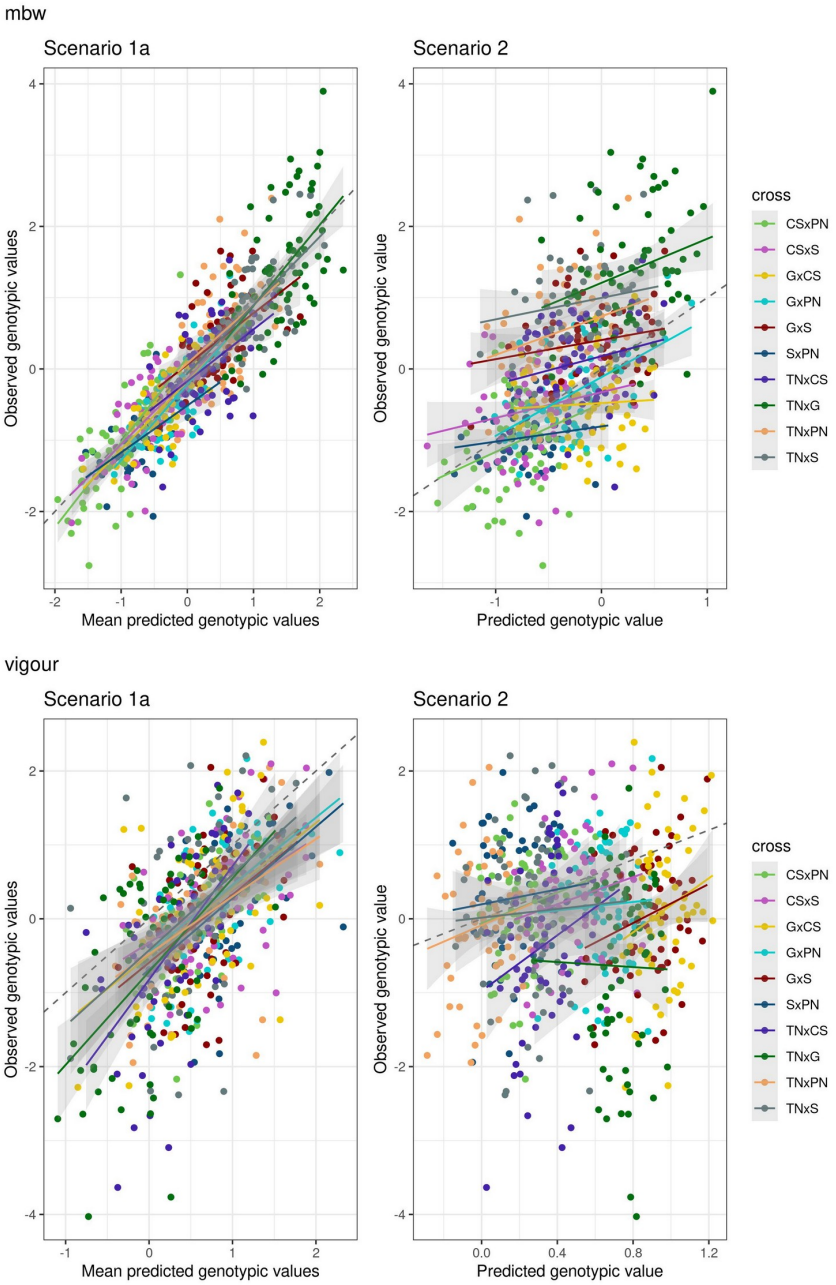


Figure S6 Observed vs predicted individual genotypic values for 15 traits.

Comparison between scenarios 1a (left) and 2 (right). Each point represents one offspring of one cross. A linear regression was fitted for each cross, with standard error displayed in grey. For scenario 1a, predicted genotypic values were averaged over 10 cross-validation repetitions. Identity (y=x) is displayed with a dashed line. Genotypic values were predicted using RR method.

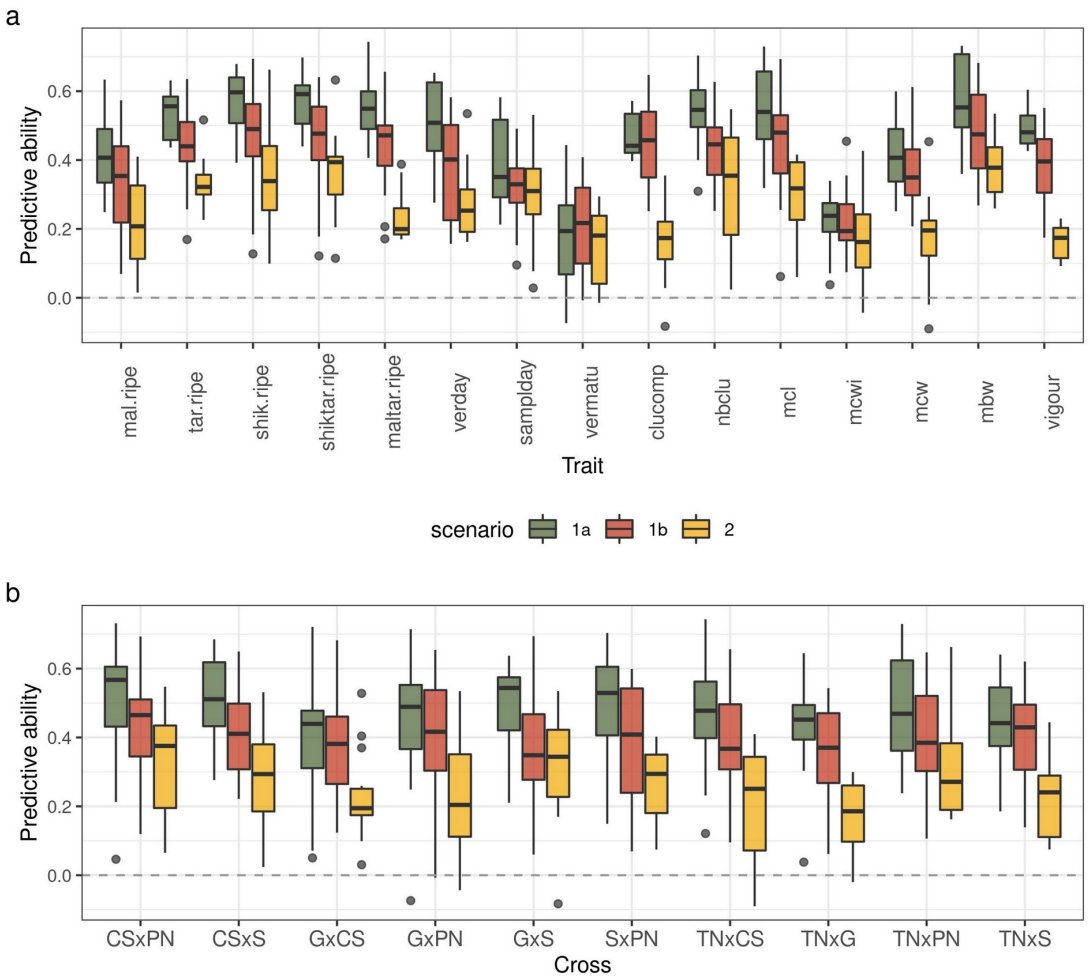


Figure S7 Distribution of predictive ability for Mendelian sampling genomic prediction.

Boxplots of PA values, calculated for all individuals within a cross, for scenarios 1a (green), 1b (red) and 2 (yellow), for the best method among RR and LASSO

a: distribution for each trait, **b**: distribution for each cross,

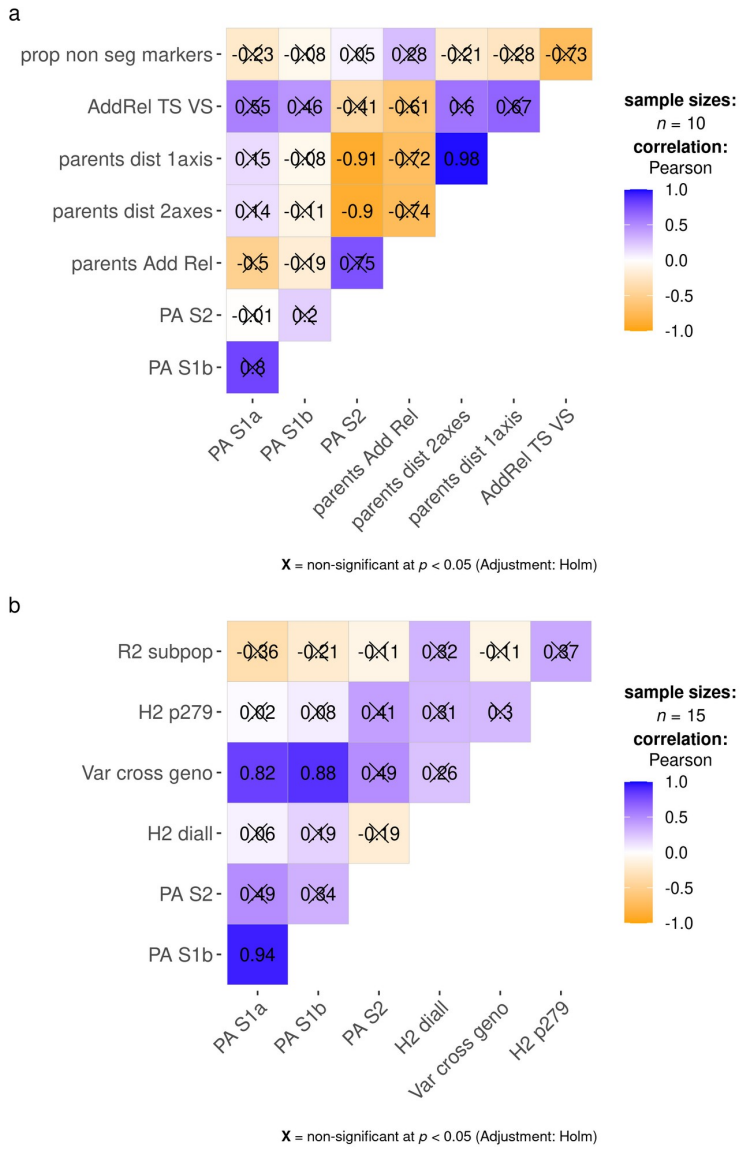


Figure S8 Correlation plot for PA of cross mean and potential explanatory variables.

PA S1a, PA S1b, PA S2 are PAs for the three scenarios 1a, 1b and 2, respectively.
a: per-cross PA and genetic variables. Prop non seg markers is the proportion of non-segregating markers in each half-diallel cross. AddRel TS VS is the mean additive relationship between TS and VS. Parents dist 1axis and parents dist 2axes are the pairwise distances between half-diallel parents on the PCA for the first axis or the first two axes, respectively. Parents Add Rel is the pairwise additive relationship between half-diallel parents.

b: per-trait PA and trait-related variables. H2 diall and H2 p279 are broad-sense heritability in the half-diallel and the diversity panel, respectively. Var cross geno is the proportion of genetic variance due to differences between crosses, as described in Methods.

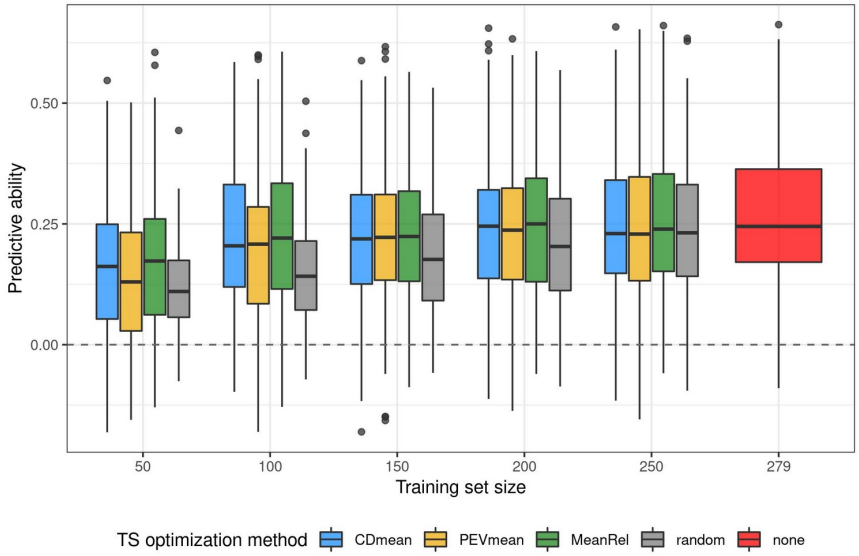


Figure S9 Distribution of predictive ability for Mendelian sampling genomic prediction, after training set optimization.

Boxplots of individual PA values over all traits and crosses in scenario 2, for different TS sizes and optimization methods, as described in Methods. “Random” method corresponds to random sampling of TS genotypes. “None" corresponds to the use of the whole diversity panel as TS. Optimization was performed for each cross of the half-diallel. The best predictive ability was kept between RR and LASSO for each trait and cross.

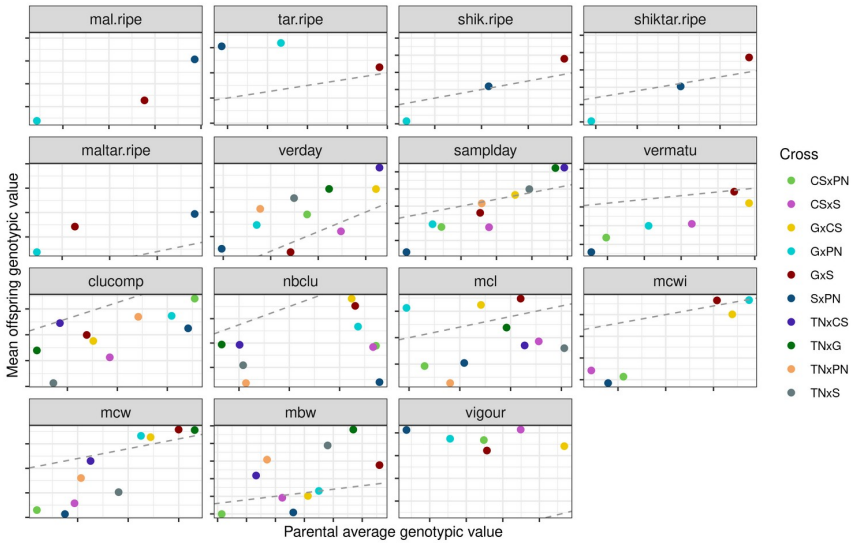


Figure S10 Mean offspring observed genotypic value vs parental average observed genotypic value in each half-diallel cross for 15 traits

The grey dashed line corresponds to identity (x=y).

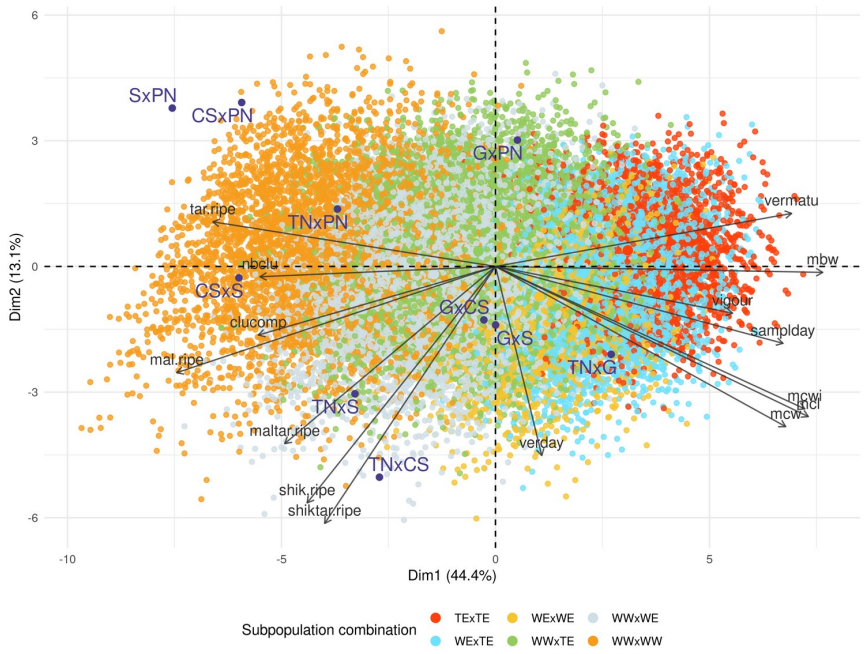


Figure S11 PCA of predicted cross mean genotypic values for all 38,781 possible simulated crosses between the 279 varieties of the diversity panel

Prediction was based on parental average genotypes and marker effects estimated with RR in the diversity panel. For each simulated cross, the dot color corresponds to the combination of panel subpopulations from which the parents of the cross originate. Values of the half-diallel crosses were projected.

Cross	Female	Male
SxPN	NA	PN: Pinot Noir
CSxPN	CS: Cabernet-Sauvignon	PN: Pinot Noir
GxPN	G: Grenache	PN: Pinot Noir
TNxPN	NA	PN: Pinot Noir
CSxS	CS: Cabernet-Sauvignon	S: Syrah
GxS	G: Grenache	S: Syrah
TNxS	TN: Terret Noir	S: Syrah
GxCS	G: Grenache	NA
TNxCS	TN: Terret Noir	NA
TNxG	TN: Terret Noir	NA

Table S4 Partial pedigree of half-diallel crosses used for marker imputation.

As the software **Fimpute3** does not handle hermaphroditism, we declared a partial pedigree which maximizes the number of crosses with both parents defined.

	Training set	Training set sizes (depending on trait)	Validation set	Validation set sizes (depending on trait)
Scenario 1a	Random sampling of 9/10th of the half-diallel	min: 477 max: 558	The remaining 1/10th of the half-diallel	min: 53 max: 62
Scenario 1b	Three bi-parental crosses with one common parent	min: 140 max: 193	The fourth cross with the same common parent	min: 40 max: 67
Scenario 2	The whole diversity panel	min: 234 max: 267	Each cross of the half-diallel	min: 531 max: 620

Table S5 Training and validation sets composition for each scenario used to assess genomic prediction

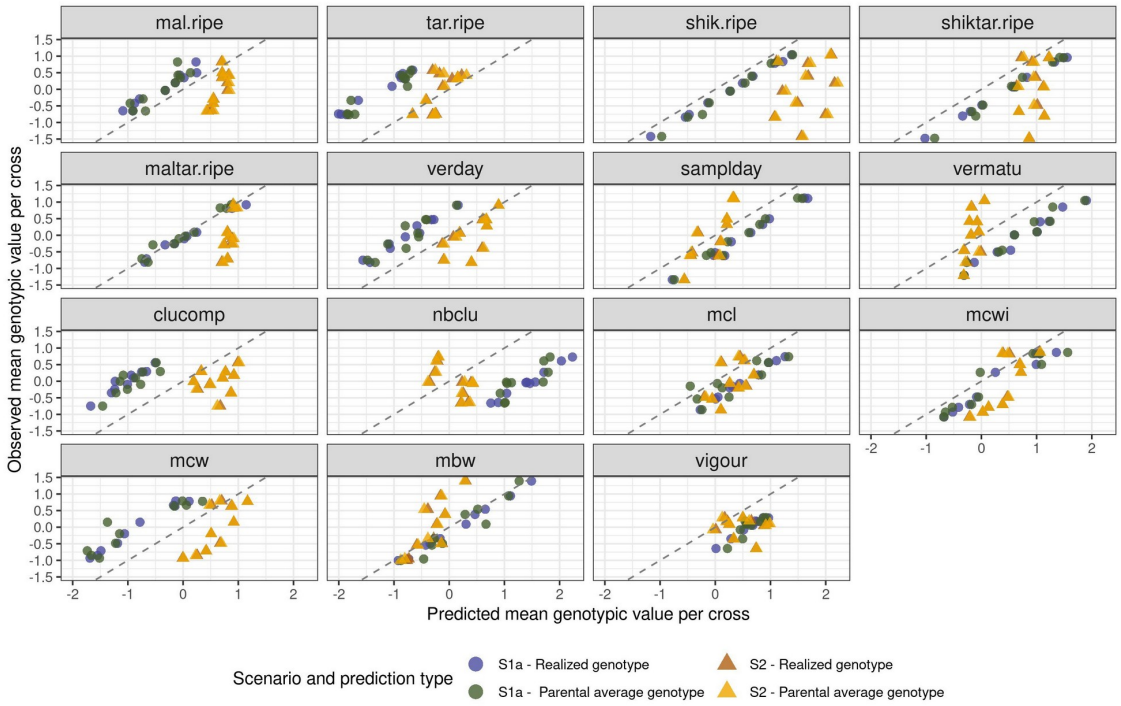


Figure S12 Observed vs predicted cross mean for 15 traits.

Predicted values per cross are displayed for realized genotypes or parental average genotype, and for scenarios 1a (circle) or 2 (triangle). Observed cross mean is the averaged genotypic value over all offspring within a cross. Identity (y=x) is displayed with a dashed line. Predicted values were obtained with RR.

Bibliography

VanRaden, P. M. Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* **91**, 4414–4423 (2008).

Appendix D

Appendix Chapter 4

D.1 Supplementary information for article III

Supplementary information

Interest of phenomic prediction as an alternative to genomic prediction in grapevine

Charlotte Brault, Juliette Lazerges, Agnès Doligez, Miguel Thomas, Martin Ecarnot, Yves Bertrand, Gilles Berger, Thierry Pons, Pierre François, Loïc Le Cunff, Patrice This, Vincent Segura

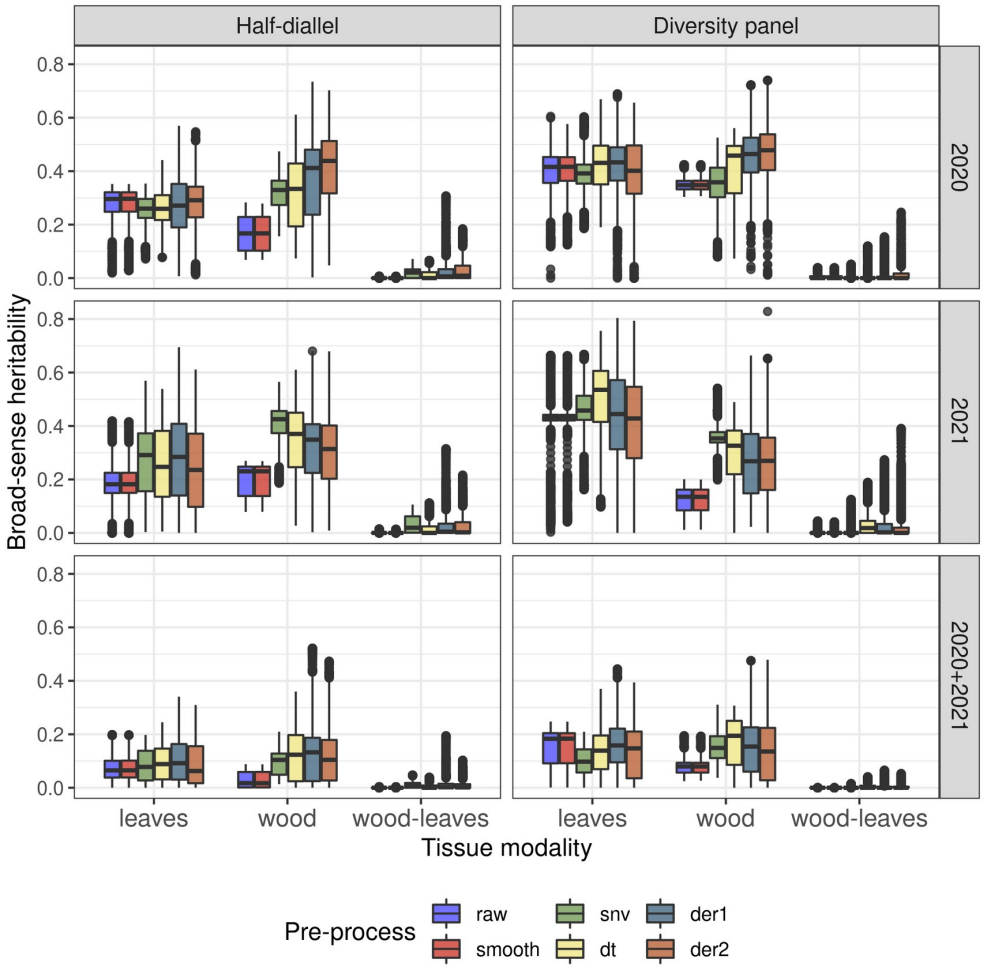


Figure S1 Heritability distribution across wavelengths for each pre-process

Comparison of broad-sense heritability between the six NIRS pre-processes, displayed for both populations, three year and three tissue modalities.

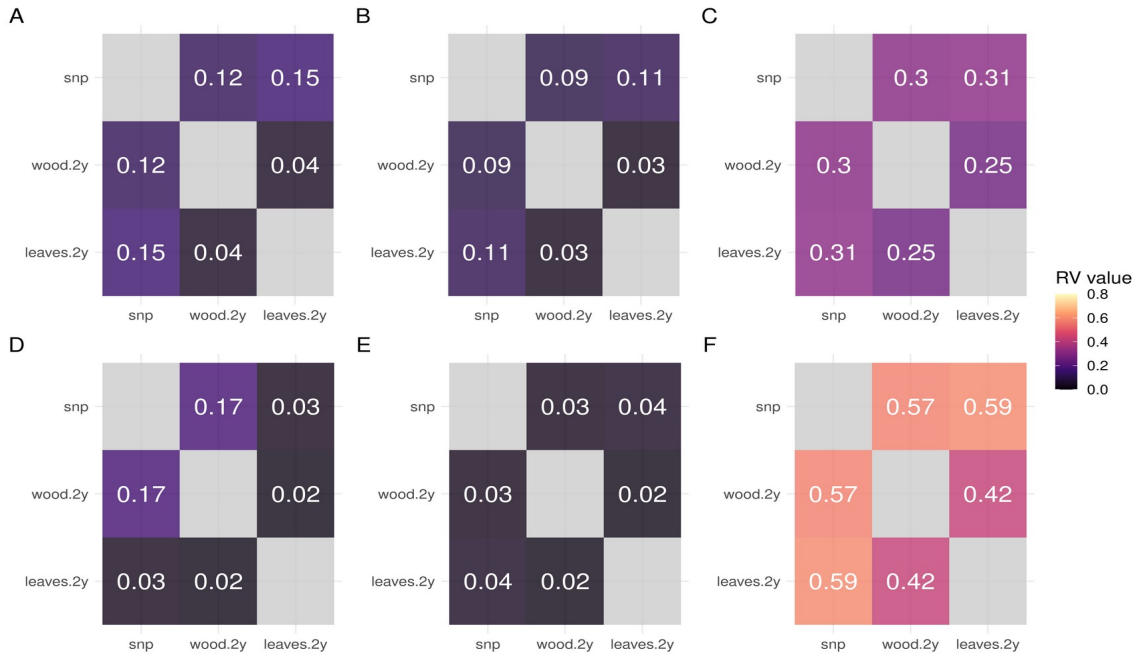


Figure S2 Co-inertia analysis between SNP and multi-year wood and leaves NIRS matrices

A, B, C: in the diversity panel

D, E, F: in the half-diallel

A, D: mixed model with only genotype effect, resulting NIRS matrix of genotype BLUPs

B, E: mixed model with genotype and *subpopulation* or *cross* effects, resulting NIRS matrix of genotype BLUPs only

C, F: mixed model with genotype and *subpopulation* or *cross* effects, resulting NIRS matrix of genotype + *subpopulation* or *cross* BLUPs

All mixed models were fitted after der1 pre-process.

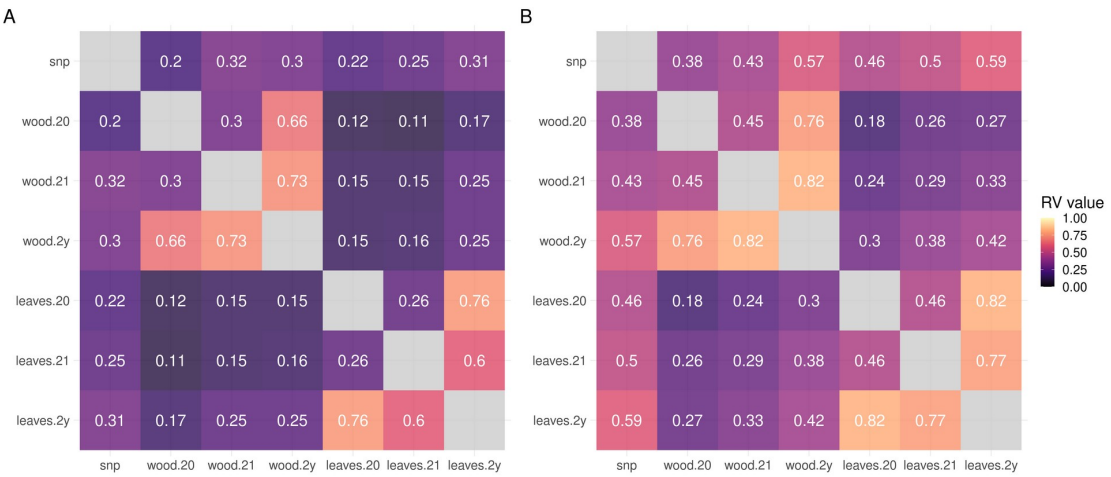


Figure S3 Co-inertia analysis between SNP, single and multi-year wood and leaves NIRS matrices

A: in the diversity panel

B: in the half-diallel

Suffixes “20”, “21”, “2y” correspond to 2020, 2021 and multi-year NIRS, respectively. All mixed models were fitted after der1 pre-process.

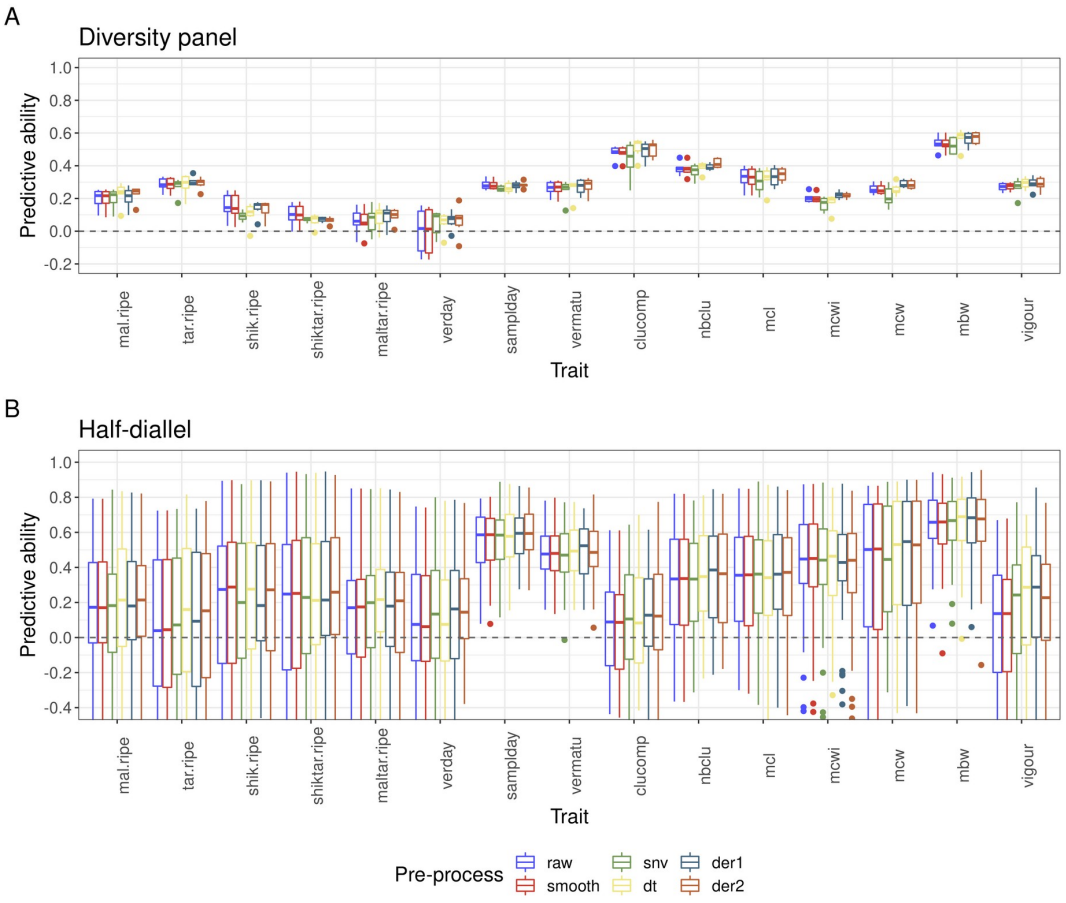


Figure S4 Distribution of phenomic prediction predictive ability per pre-process and per trait

A: in the diversity panel with rrBLUP method (implemented with glmnet); B: in the half-diallel with HBLUP method (implemented with lme4GS).

Distribution of predictive ability is displayed across two tissue x three year modalities, times ten crosses in the half-diallel.

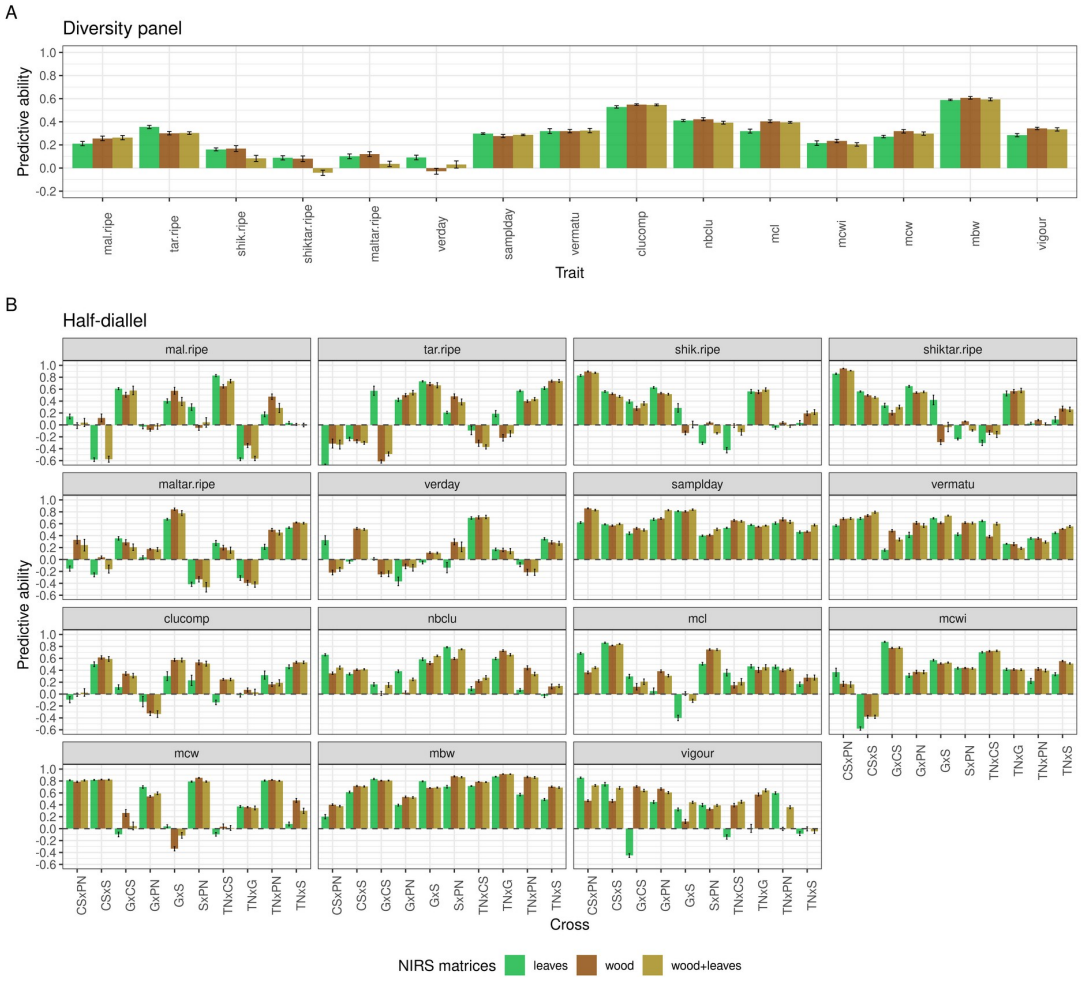


Figure S5 Predictive ability of phenomic prediction with a single vs both tissues

For “2 years” NIRS BLUPs derived after der1 pre-process. A: in the diversity panel, B: in the half-diallel. Predictive ability values are displayed per trait for both populations, and also per cross in the half-diallel. Prediction models were fitted with glmnet in the diversity panel (except for wood+leaves configuration) and with lme4GS in the half-diallel. Error bars correspond to 95% confidence intervals around the mean, calculated for the ten CV repetitions.

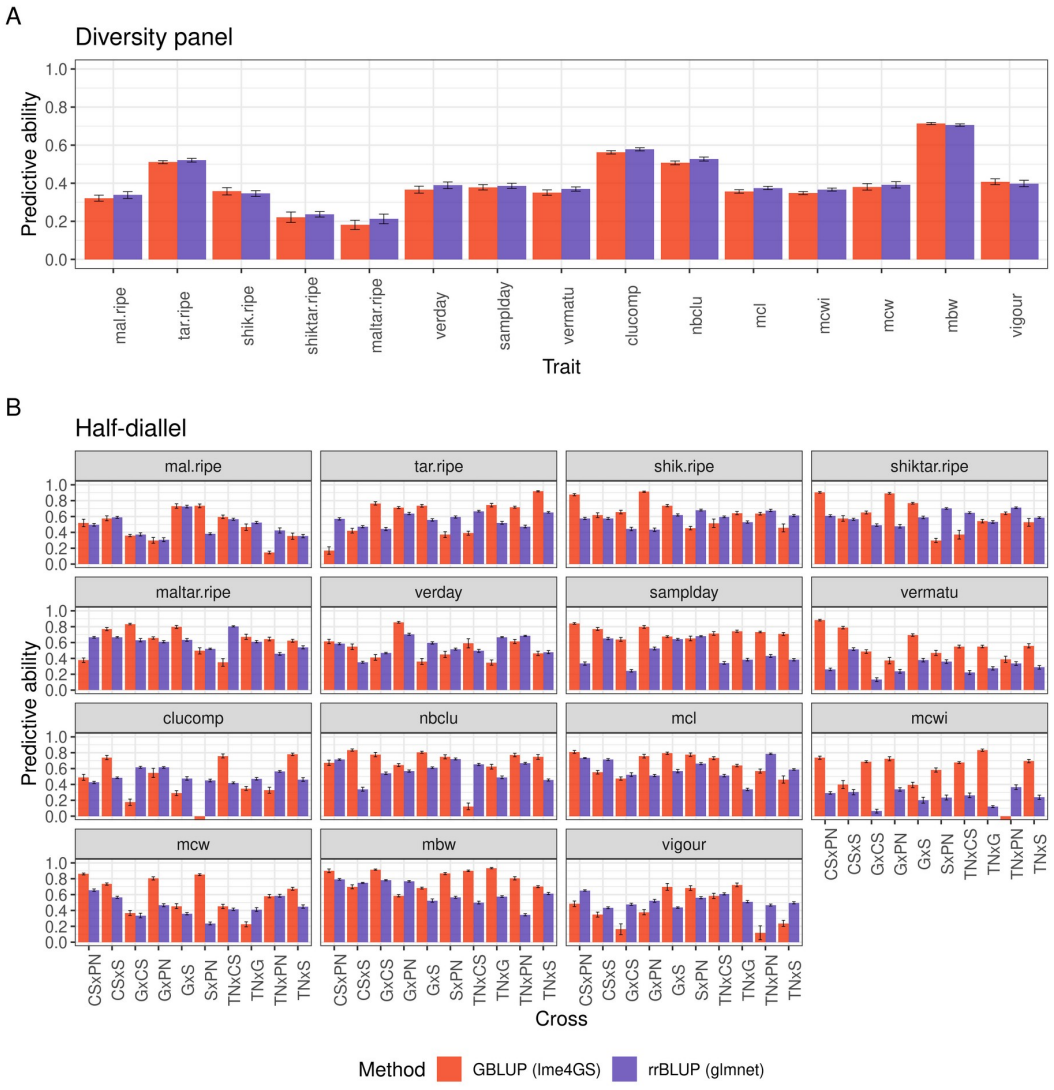


Figure S6 Comparison of methods for genomic prediction

A: in the diversity panel; B: in the half-diallel. Predictive ability values are displayed per trait for both populations, and also per cross in the half-diallel. Error bars correspond to 95% confidence intervals around the mean, calculated for the ten CV repetitions.

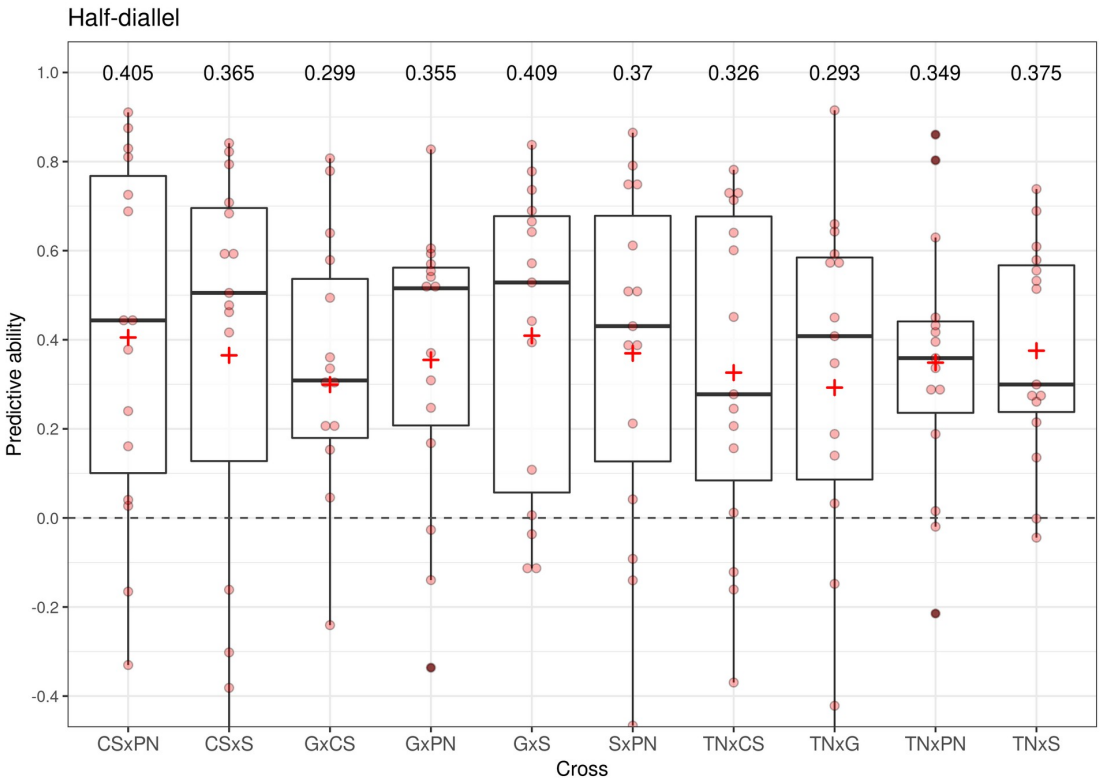


Figure S7 Distribution of phenomic prediction predictive ability over 15 traits in each half-diallel cross

For “2 years” NIRS BLUPs derived after der1 pre-process. Average PA per cross is displayed above each cross and with a red cross. Prediction models were fitted with lme4GS and included both wood and leaves NIRS relationship matrices.

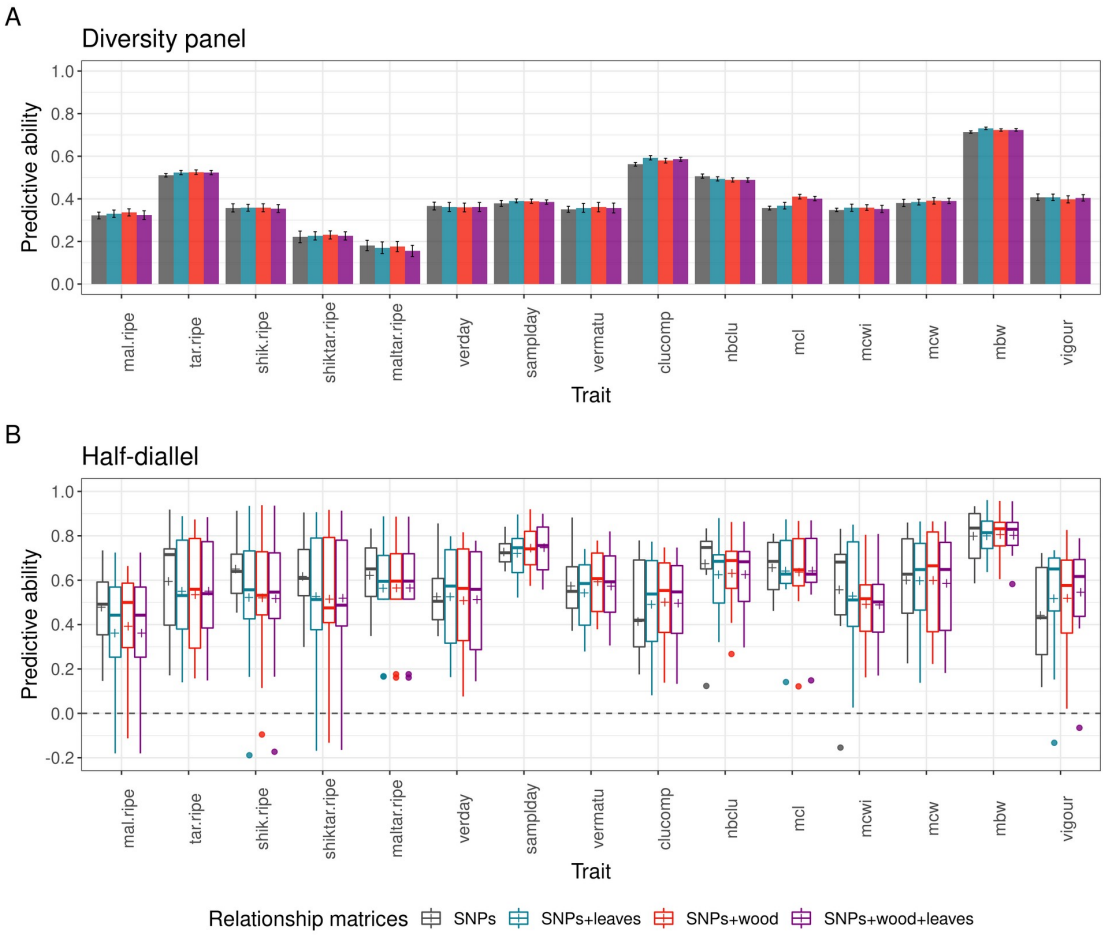


Figure S8 Predictive ability of combined vs genomic prediction models

A: in the diversity panel (error bars correspond to 95% confidence intervals around the mean, calculated for the ten CV repetitions); B: in the half-diallel (distribution over 10 crosses and 10 CV repetitions). Prediction models were fitted with lme4GS and included as relationship matrices: SNPs, SNPs and leaves NIRS, SNPs and wood NIRS or SNPs, wood and leaves NIRS. Wood and leaves NIRS BLUPs were derived from a mixed model including both years, after der1 pre-process

Abstract

Grapevine breeding needs to address two main issues over the next few years: reducing pesticide use and adaptating to climate change. If the selection of new resistant grapevine cultivars has been accelerated, breeding remains a long and costly process for a perennial species such as grapevine. That is why I tested and compared different methodologies for optimizing the breeding of new grapevine varieties. The first is genomic prediction (GP), which relies on the use of molecular markers to train a model to predict genetic values. The second is phenomic prediction (PP), which relies on the use of spectra measured on plant tissues, which is cheaper and more high-throughput than genotyping. I used GP and PP under different configurations to evaluate their interest in breeding programs. For that, I used three grapevine populations with contrasted relatedness, both genotyped and phenotyped. First, I compared univariate and multivariate GP models in a bi-parental population (N=188), on 14 traits related to drought. Multivariate methods did not perform better than univariate ones, and ranking between methods depended on the genetic architecture and heritability of the trait. Secondly, I tested across-population GP (a more applicable configuration for breeding) for 15 traits, by training the prediction model in a diversity panel (N=277) and using 10 bi-parental families of a half-diallel (N=622) as validation sets. For that, I first predicted the average genetic value of each family (for the first selection step of future crosses to be made) and then the genetic values of individuals within each family (for the second selection step of offspring within crosses once they have been realized). Prediction accuracy for these two steps appeared to be satisfying for the application of GP in breeding programs, compared to within-population prediction accuracy. Finally, I tested for the first time the application of PP in grapevine, by using near-infrared spectra measured on wood and leaves. Prediction accuracies were encouraging, despite a suboptimal framework because phenotypes and spectra were not measured in the same years. In conclusion, these results provide a good overview of methods and configurations that could be applied to optimize the selection of new grapevine varieties. For this optimization to be effective, it will be necessary to define the varietal ideotype and to have a reference population genotyped and phenotyped in several environments. My results allowed me to propose a new breeding scheme that comprises two steps, the selection of crosses to be made with GP and the selection of individuals within crosses with GP or PP. Such a scheme would allow to increase genetic gain while decreasing the cycle length from 15 to 9 years.

Résumé

La sélection génétique de la vigne doit répondre à deux enjeux majeurs dans les prochaines années : la réduction de l'utilisation de produits phytosanitaires et l'adaptation au changement climatique. Si la sélection de variété résistantes a été grandement accélérée, la sélection de caractères d'intérêt complexes reste un processus long et coûteux pour une plante pérenne comme la vigne. C'est pourquoi j'ai testé et comparé différentes méthodologies pour optimiser la sélection de nouvelles variétés de vigne. La première est la prédiction génomique (PG), qui repose sur les marqueurs moléculaires pour entraîner un modèle de prédiction des valeurs génotypiques. La seconde est la prédiction phénomique (PP), qui repose sur l'utilisation de spectres mesurés sur les tissus de la plante, plus rapides et moins chers à obtenir que le génotypage. J'ai utilisé la PG et la PP dans différentes configurations pour mesurer leur intérêt pratique dans les programmes d'amélioration. Pour cela, je me suis basée sur trois populations de variétés de vigne avec des apparentements génétiques contrastés, préalablement génotypées et phénotypées. J'ai d'abord comparé des modèles de PG univariés et multivariés dans une population bi-parentale (N = 188), sur 14 caractères relatifs au stress hydrique. Les précisions de prédiction étaient bonnes pour les caractères héritables, les méthodes multivariées n'étaient pas meilleures que les méthodes univariées, et le classement entre les méthodes dépendait de l'architecture génétique du caractère. Dans un second temps, j'ai testé la PG inter-population (configuration plus applicable en sélection) pour 15 caractères, en entraînant le modèle de prédiction dans un panel de diversité (N = 277) pour prédire dans 10 familles bi-parentales connectées en demi-diallèle (N = 622). Pour cela, j'ai prédit d'abord la moyenne des familles (pour la sélection des futurs croisements à réaliser), puis la valeur des individus au sein de chaque famille (pour la sélection d'individus au sein des croisements un fois qu'ils ont été réalisés). Pour ces deux étapes, les précisions de prédiction en inter-population restaient satisfaisantes (par rapport à l'intra-population) pour l'application de la PG dans les programmes d'amélioration. Finalement, j'ai testé la PP pour la première fois chez la vigne, en utilisant des spectres dans le proche infrarouge obtenus sur bois et sur feuilles. Les précisions de prédiction en PP se sont révélées encourageantes, malgré un cadre suboptimal car les spectres et les phénotypes n'ont pas été mesurés les mêmes années. En conclusion, ces résultats donnent un bon aperçu de ce qui pourrait être appliqué pour optimiser la sélection de nouvelles variétés de vigne. Pour que cette optimisation soit effective, il faudra cependant bien définir l'idéotype variétal et avoir à disposition une population de référence génotypée et phénotypée dans plusieurs environnements. Le nouveau schéma de sélection que je propose sur la base de mes résultats comprend deux étapes : la sélection des croisements à réaliser par PG et la sélection des individus au sein des croisements par PG ou PP. Un tel schéma pourrait permettre d'augmenter le gain génétique par cycle et de diminuer la durée du cycle de 15 à 9 ans.