



Créer les cépages de demain avec les outils d'aujourd'hui

Loic L. Le Cunff, Agota Fodor, Laurent Audegin, Agnes Doligez, Jean-Pierre Peros, Patrice P. This

► To cite this version:

Loic L. Le Cunff, Agota Fodor, Laurent Audegin, Agnes Doligez, Jean-Pierre Peros, et al.. Créer les cépages de demain avec les outils d'aujourd'hui. Innovations Agronomiques, 2015, 46, pp.63-74. 10.15454/1.4622670753648755E12 . hal-02635687

HAL Id: hal-02635687

<https://hal.inrae.fr/hal-02635687>

Submitted on 27 May 2020

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Créer les cépages de demain avec les outils d'aujourd'hui.

Le Cunff L.¹, Fodor A.¹, Audeguin L.¹, Doligez A.^{2,3}, Peros J.P.^{2,3}, This P.^{2,3}

¹ IFV, UMT Géno-Vigne® (IFV, INRA, Montpellier Supagro), 2 place Viala, 34060 Montpellier

² INRA, UMT Géno-Vigne® (IFV, INRA, Montpellier Supagro), 2 place Viala, 34060 Montpellier

³ INRA, UMR AGAP, 2 place Viala, 34060 Montpellier

Correspondance : loic.lecunff@vignevin.com

Résumé

La viticulture est aujourd'hui une des espèces les plus fortes utilisatrices de produits phytosanitaires en Europe. Par ailleurs, les évolutions du climat pourraient engendrer de profondes modifications et notamment dans la zone méditerranéenne. La création variétale peut-être une des solutions possibles pour répondre à ces défis. L'amélioration variétale est aujourd'hui assistée de plus en plus par des données moléculaires. Chez la vigne, les marqueurs utilisés actuellement sont liés à des caractères monogéniques comme la résistance au mildiou et à l'oïdium, l'apyrénie et le gout muscat. Les marqueurs sont connus et leur utilisation en routine est maintenant actée. Ce projet s'axe sur l'étude de méthodologies ou d'outils permettant i) la sélection de caractères complexes généralement contrôlés par plusieurs gènes ii) l'accélération du cycle de génération pour faciliter l'obtention de géniteurs élités. Ce projet permet la définition de plusieurs stratégies pour développer la création variétale chez la vigne.

Mots-clés: Sélection génomique, diversité, résistance, marqueurs moléculaires

Abstract: Develop new varieties in grapevine for tomorrow with present tools

Viticulture is one of the crops with the largest use of pesticides. Moreover viticulture is also impacted by climate change. Among the possible answers to these challenges, the development of new varieties is a possible way. Nowadays, the selection assisted by molecular markers (SAM) presents a great potential in plant breeding programs. In grapevine, SAM is only used to select simple traits (monogenic) as resistance to downy and powdery mildew, Muscat aroma or seedless. However many targeted traits as quality of berries show a complex inheritance (oligogenic or polygenic). In the project, we designed a strategy to speed up grapevine selection of complex traits including new methodologies and scientific knowledge.

Keywords: Genomic selection, diversity, resistance, molecular markers

Introduction

L'objectif principal de ce projet est d'acquérir les méthodologies et les connaissances nécessaires au développement de la création variétale chez la vigne avec comme but l'obtention de variétés de vigne aux propriétés organoleptiques intéressantes et résistantes au mildiou et à l'oïdium.

Afin d'atteindre ce but, nous avons identifié et donc décidé de travailler sur les thématiques suivantes:

- La prédiction phénotypique de caractères complexes (oligo ou polygéniques) via l'utilisation de marqueurs moléculaires et grâce aux méthodologies telles que la sélection génomique et la génétique d'association,
- L'étude d'un outil permettant de réduire le temps entre deux générations d'une vigne naine, appelé « Dwarf ».

1. Contexte actuel de la création variétale en vigne

La viticulture est aujourd'hui une des espèces les plus fortes utilisatrices de produits phytosanitaires en Europe. Par ailleurs, les évolutions du climat pourraient engendrer de profondes modifications et notamment dans la zone méditerranéenne. Les espèces apparentées du genre *Vitis* et du genre *Muscadinia*, interfertiles avec *Vitis vinifera* L. *sativa*, présentent des résistances monogéniques ou quantitatives aux stress biotiques et abiotiques, mais des caractères de qualité défavorables. Certaines de ces résistances ont déjà été utilisées au travers d'hybrides (HPD) mais du fait des caractères qualitatifs défavorables des espèces n'ont pas permis l'obtention de variétés de qualité. Les ressources moléculaires disponibles actuellement chez la vigne offrent dorénavant la possibilité de créer des variétés de vigne durablement résistantes tout en améliorant la qualité du raisin et donc du vin. L'UMT Géno-vigne® s'inscrit parfaitement dans ce cadre au travers de son activité de transfert de l'innovation vers la recherche appliquée. De plus, l'accès au domaine de Vassal (plus grande collection au monde de ressources génétique pour la vigne) et aux projets de génomique développés par l'INRA sur la vigne, sont des bases de connaissance et de savoir-faire incontestables pour développer ce projet.

Aujourd'hui, la viticulture française doit faire face à 3 grands défis:

- la réduction des intrants phytosanitaires (plan Ecophyto 2018),
- les changements climatiques,
- et de nouveaux concurrents sur le marché, notamment les pays du Nouveau Monde.

Même si l'organisation actuelle de la viticulture ne laisse que peu de place à la création variétale -- en effet, les AOC imposent un unique lieu de production mais aussi une liste de cépages pour produire un vin d'appellation -- et que le consommateur averti est sensibilisé aux cépages dans les vins, **la création variétale peut être une des solutions possibles pour répondre à ces défis**. Plusieurs interprofessions s'intéressent à cet axe de recherche (BNIC, CIVC, CIVB, CIVL, CIVP). Des projets de création variétale sont en cours avec des collaborations entre l'INRA (UMR SVQV et AGAP) et l'IFV. Ces collaborations sont renforcées par la signature d'un accord-cadre entre ces deux instituts et l'orientation scientifique de l'UMT Géno-Vigne® 2, dont le thème principal est la création variétale.

La vigne dispose de nombreux outils puissants permettant d'envisager la mise en place des dernières méthodologies d'amélioration :

- La vigne (*Vitis vinifera*) a des ressources génétiques très importantes liées à la diversité de l'espèce. Dans le cadre de ce projet, nous avons la possibilité de travailler avec la plus grande collection au monde de ressources génétiques disponibles de l'espèce *Vitis vinifera* et de ses espèces apparentées. Cette collection est localisée au domaine de Vassal et dépend de l'INRA de Montpellier.
- La séquence complète du génome (Jaillon *et al.* 2007 ; Velasco *et al.* 2007) qui facilite grandement la mise en place de méthodologies haut-débit d'identification de marqueurs. L'utilisation des nouvelles technologies de séquençage et d'approche du type « Genotyping by sequencing » (GBS) permettent notamment de découvrir de nombreux polymorphismes et de diminuer le prix du génotypage (Myles *et al.*, 2010 ; Davey *et al.*, 2011).

Cependant, l'amélioration de la vigne présente aussi des difficultés liées aux caractéristiques de l'espèce.

- Plusieurs études menées sur le déséquilibre de liaison (DL) chez la vigne ont détecté un niveau de DL très faible (Barnaud *et al.*, 2006 ; Lijavetzky *et al.*, 2007 ; This *et al.*, 2007 ; Myles *et al.*, 2010). Selon la littérature, le DL chute lorsque la taille du fragment est supérieure à 10Kb (le seuil de r^2 choisi est généralement 0,2).
- Les cycles de reproduction chez la vigne sont relativement longs (3 ans).

2. Description et place de ce projet dans le cadre de la création variétale vigne

L'amélioration variétale est aujourd'hui assistée de plus en plus par des données moléculaires. Chez la vigne, les marqueurs utilisés actuellement sont liés à des caractères monogéniques comme la résistance au mildiou et à l'oïdium, l'apyrénie et le gout muscat. Les marqueurs sont connus et leur utilisation en routine est maintenant actée. Ce projet s'intéresse à la sélection de caractères complexes généralement contrôlés par plusieurs gènes. Actuellement, deux méthodologies très proches sont utilisées chez d'autres plantes et animaux pour l'amélioration, la sélection génomique (GS) et la génétique d'association (GWAS). La finalité de ces deux approches est, dans un contexte de création variétale, de prédire la valeur génétique de l'individu qu'on appelle Genetic Estimated Breeding Values (GEBVs) (Schaeffer, 2006). Ces approches doivent être évaluées afin d'estimer leur intérêt chez la vigne. Même si les modèles statistiques utilisés dans ces deux méthodologies sont les mêmes, l'utilisation des résultats n'est pas la même. En effet, la méthodologie « GWAS » a pour but l'identification des marqueurs moléculaires (quelques) liés au déterminisme des caractères étudiés. L'utilisation d'un seuil (défini empiriquement) permettant un tri est implicite. Ces marqueurs peuvent ensuite être utilisés dans des schémas de Sélection Assistée par Marqueurs moléculaires (SAM). *A contrario*, la philosophie de la méthodologie « GS » est d'utiliser l'ensemble des marqueurs moléculaires disponibles et d'estimer un effet différent pour chacun, entre total ou nul, afin d'expliquer la variance génétique totale du caractère étudié.

Afin de réaliser ces prédictions, nous avons besoin de définir les paramètres de l'équation qui relie le phénotype au génotype. L'estimation de ces paramètres se fait sur une population dite d'entraînement, qui est génotypée et phénotypée. La diversité, le nombre d'individus et le nombre de marqueurs moléculaires disponibles sur cette population sont des critères essentiels. Il est aussi connu que la prédiction est plus précise quand la population d'entraînement et la population candidate sont génétiquement proches (Habier et al., 2010). Chez la vigne en absence de lignées élites et de populations de pré-breeding, nos populations d'entraînement seront composées i) de trois ensembles de cépages recouvrant la diversité de l'espèce *Vitis vinifera* ii) ou d'un croisement biparental entre les variétés Syrah et Grenache. Un des objectifs de ce projet est d'évaluer la composition de la population d'entraînement chez la vigne en fonction i) de la distance génétique qui la sépare de la population de validation, et ii) du caractère étudié (caractère corrélé avec la structure de la population d'entraînement).

Un autre levier de l'amélioration variétale afin notamment d'introgesser des caractères d'intérêt comme des gènes de résistance est le temps du cycle. Chez la vigne, le cycle est généralement de 3 ans. Afin d'accélérer les cycles de reproduction chez la vigne et de raccourcir ainsi le temps d'un programme de sélection, il est désormais possible d'utiliser un génotype particulier appelé « Dwarf » (nain). Ce génotype est le résultat d'une mutation spontanée (naturelle) apparue dans une des couches cellulaires du cépage Meunier le rendant insensible aux gibbérellines (Boss et al., 2002). Cette insensibilité a pour conséquence de réduire le temps de mise à fruit des génotypes porteurs de cette mutation. Lorsqu'une amélioration variétale nécessite plusieurs cycles de croisement, l'attente de mise à fruit des génotypes intermédiaires est longue chez la vigne (entre 2 à 3 ans minimum) alors que chez les génotypes « Dwarf » le passage du pépin au pépin est de 9 à 10 mois (Chaïb et al., 2010). Ces génotypes « Dwarf » sont donc de parfaits intermédiaires pour accélérer les programmes de sélection complexe chez la vigne. De plus, il est possible à chaque génération de revenir à un génotype « classique » puisque 50% des descendants d'un croisement « Dwarf X vigne génotype classique », sont morphologiquement normaux.

3. Résultats

La première partie de ce projet est l'évaluation de méthodologies de prédiction pour des caractères polygéniques via deux sous-actions :

- la réalisation de simulations afin d'estimer les paramètres les plus importants pour les modèles de prédiction dans un contexte de population d'entraînement à large base génétique (nombre d'individus dans la population d'entraînement, effet de l'éloignement génétique entre la population d'entraînement et la population à prédire, puissance de détection en fonction du nombre de marqueurs utilisés et de l'effet du QTL, impact de l'effet de la structure des populations étudiées sur le caractère ciblé, comparaison GWA, GS.).
- Des tests de sélection génomique sur des données acquises au champ. Pour cela, nous travaillons avec un échantillon de 279 individus représentatifs des 3 groupes de diversité identifiés chez *Vitis vinifera* L. (Bacilieri et al., 2013) qui sera notre population d'entraînement pour la GS. En revanche, les populations à prédire sont composées d'individus issus de différents croisements (détaillées dans la partie Matériel végétal).

La deuxième partie de ce projet comprend des tests d'utilisation du génotype « Dwarf » afin d'obtenir plus rapidement des géniteurs pluri-géniques homozygotes pour la résistance au mildiou et à l'oïdium.

3.1 Simulation

Pour réaliser l'étude de simulation, nous avons utilisé le logiciel quantiNemo (Neuenschwander et al., 2008), car il permet de simuler les génotypes et les phénotypes des individus. Ce logiciel de simulation est de type forward dans le temps, ce qui veut dire que les individus sont issus d'un scénario évolutif dont on doit préciser les détails.

3.1.1 Choix du scénario

La définition d'un scénario permettant de mimer les caractéristiques de l'espèce *Vitis vinifera* a été réalisée en concertation avec des spécialistes du domaine (Jean-Michel Boursiquot et Thierry Lacombe) mais aussi en s'aidant des informations présentes dans la littérature scientifique traitant du sujet. De nombreuses études sur la domestication de la vigne ont été réalisées (Levadoux, 1956 ; Aradhya et al., 2003 ; Grassi et al., 2003 ; Arroyo-García et al., 2006 ; This et al., 2006 ; Myles et al., 2011). Sur cette base nous avons construit **le scénario démographique le plus simple possible** illustrant les connaissances actuelles. Il aboutit à la construction des trois groupes de diversité de l'espèce *Vitis vinifera* L. *vinifera* (le compartiment cultivé de cette espèce) que nous avons observé par l'analyse de la collection de Vassal. Le premier groupe que nous appellerons « table de l'est – TE », correspond aux raisins de table d'aujourd'hui qui aurait pour origine le centre de domestication primaire, localisé dans le Caucase. Les 2 autres groupes regroupent les variétés de cuve, qui ont été sélectionnées pour donner du vin. Un de ces 2 groupes « cuve de l'est – CE » serait le résultat d'un événement de domestication secondaire. Suivant les migrations humaines, les variétés du centre de domestication primaire arrivent aux Balkans où elles se croisent avec les variétés sauvages locales en créant un deuxième centre de diversité. L'autre groupe « cuve de l'ouest – CO » serait le résultat d'une migration du groupe issu du centre de domestication secondaire, suivant les migrations humaines, de l'est vers l'ouest de l'Europe, suivie d'un croisement avec des individus sauvages de ces zones géographiques (Nord de l'Espagne, Sud de la France). Ces trois groupes sont très proches de ceux caractérisés par Negrul (1938) sur des observations phénotypiques: Orientales pour TE, Pontica pour CE et Occidentalis pour CO. Ces trois groupes peuvent être distingués par les marqueurs moléculaires (Figure 1).

Le scénario sélectionné, représenté sur la Figure 2, prend en compte les migrations entre pools et les sélections humaines. Nous avons aussi simulé des caractères quantitatifs dits simples (10 QTLs impliqués) ou complexes (100 QTLs impliqués), mais aussi liés à la structure (très variable dans une sous-population alors que quasi fixée dans une autre) ou pas. Pour le génotypage, nous avons simulé 100 000 marqueurs de type SNPs et 500 de type microsatellites (avec 10 allèles).

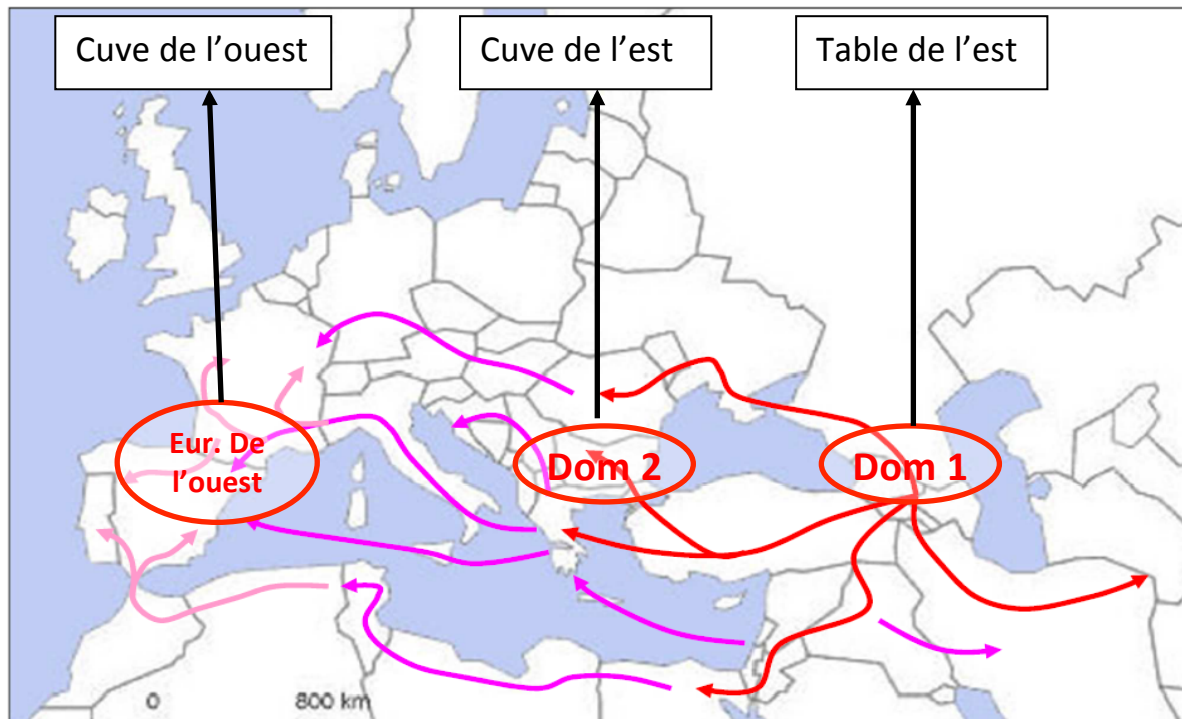


Figure 1. : Diffusion de la viticulture et de la vigne cultivée (d'après Labra et al. (2002)).

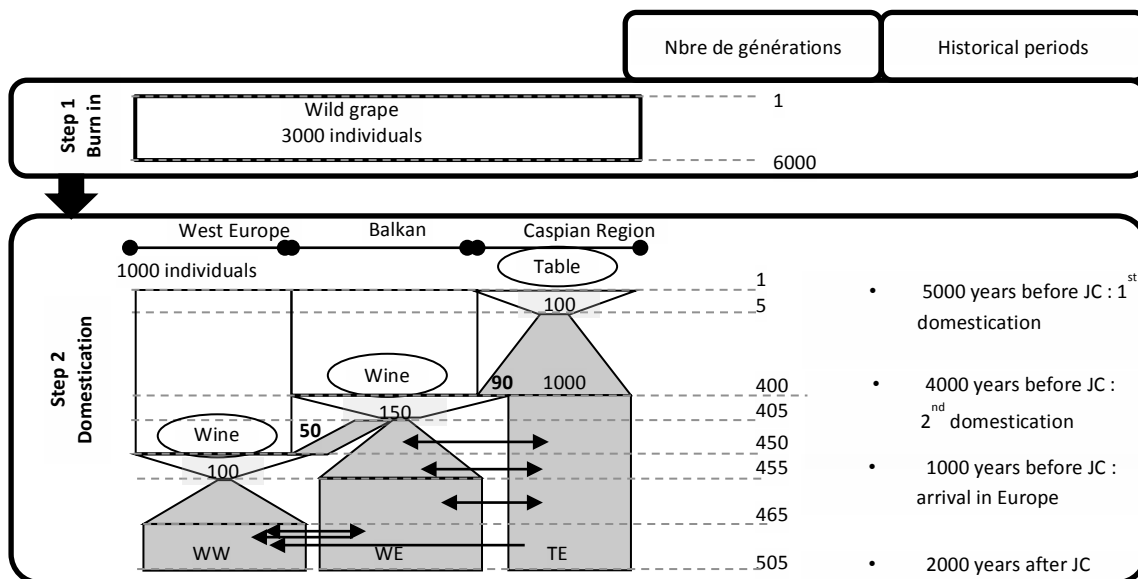


Figure 2 : Représentation schématique du scénario évolutif de l'espèce *Vitis vinifera* L.

Nous avons réalisé 10 fois le scénario choisi afin d'éviter un biais dans les analyses suivantes dû à la production d'un unique scénario. Toutes les analyses ont donc été réalisées avec les 10 scénarii et les résultats présentés sont la moyenne et l'écart-type des résultats obtenus. Afin de valider notre scénario, nous avons évalué des indices de diversité sur les populations issues de ce scénario et nous les avons comparés aux résultats obtenus par Laucou et collaborateurs (2010) sur la collection du domaine de Vassal qui est actuellement la collection la plus exhaustive pour l'espèce *Vitis vinifera* L. Nous avons aussi comparé les résultats avec des scénarii alternatifs en jouant sur le niveau de migration et de sélection entre populations.

Les résultats sont présentés dans le Tableau 1.

		Scenario retenu	Données publiées	Scenarii Alternatifs			
				Migration		Sélection	
				Sans	2 fois plus	Sans	2 fois plus
FST	WW-WE	0.04 (0.007)	0.05 ^a	0.34 (0.018)	0.01 (0.001)	0.01 (0.001)	0.05 (0.012)
	WW-TE	0.07 (0.012)	0.07 ^a	0.35 (0.001)	0.01 (0.003)	0.03 (0.003)	0.09 (0.015)
	WE-TE	0.04 (0.007)	0.05 ^a	0.45 (0.014)	0.01 (0.001)	0.03 (0.001)	0.04 (0.008)
Hétérozygotie		0.64 (0.026)	0.73 ^b	0.46 (0.011)	0.64 (0.019)	0.72 (0.005)	0.60 (0.012)

Tableau 1 : Comparaison d'indices de diversité calculés sur les populations issues de différents scénarii démographiques et des données publiées (Laucou et al. 2011)

Le scénario proposé permet d'obtenir des valeurs de FST (indice de distance génétique entre deux populations basée sur la fréquence des individus hétérozygotes au sein des populations considérées) entre populations et d'hétérozygotie, plus proche que les scénarii alternatifs. Ces résultats valident le scénario proposé et les données qui en sont issues

3.1.2 Résultats

Nous avons testé plusieurs combinaisons de populations d'entraînement et de populations de validation. Premièrement, les populations simulées ont servi de populations d'entraînement pour développer le modèle puis un sous-ensemble, tiré au hasard, a servi de population de validation (cross validation). Dans ce cas, les coefficients de régression entre les valeurs prédites par les modèles et celles du génotype sont toujours très proches de 1. C'est pourquoi nous avons décidé de simuler des populations issues du croisement entre deux individus tirés au hasard dans les populations simulées comme dans un schéma de sélection dans lequel le sélectionneur choisira deux géniteurs pour engendrer la population de sélection.

NB : pour faciliter la compréhension des codes utilisés, notons qu'il existe deux cas possibles i) les 2 individus géniteurs sont issus de la même sous population. Alors la descendance est notée d suivie du nom de la sous population ; par exemple dWW correspond à la descendance de deux individus issus de la population WW ; ii) en revanche si les deux individus géniteurs sont issus de sous populations différentes, la descendance est nommée Mixed. De plus, dans cette partie nous avons aussi cherché à diversifier les modèles de prédiction utilisés. Nous en avons utilisé 4 différents : i) le premier nommé cof n'utilise que les marqueurs issus de GWAS. Le second RR pour Ridge Regression et le troisième bLR (bayesian Linear Regression) correspondent à des modèles fréquemment utilisés en sélection génomique (ils diffèrent par les *a priori* possibles) et finalement un modèle développé dans cette étude nommé cofRR qui est un modèle mixte alliant les deux informations (celle de cof et celle de RR). Les résultats sont nettement différents que ceux obtenus lors de la cross-validation avec des valeurs plus faibles (Figure 4).

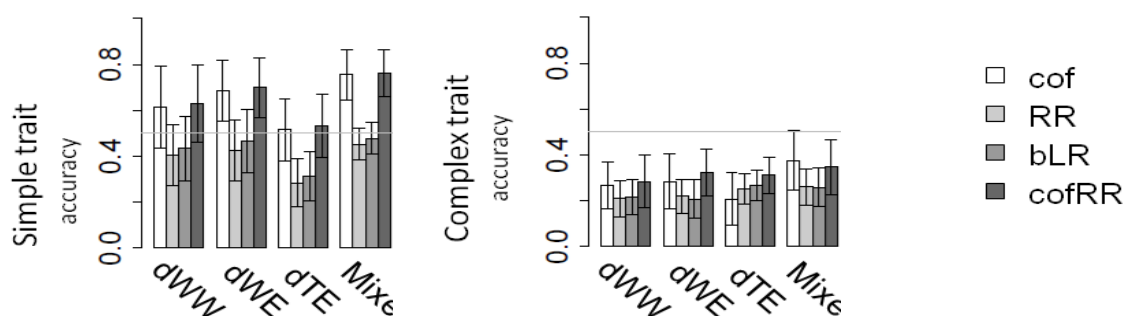


Figure 4 : Précision de la prédiction (accuracy) obtenue en utilisant des populations de validation non issues de la population d'entraînement.

La précision de prédiction change lorsque l'on cherche à prédire la valeur d'individus non présents dans la population d'entraînement. Ceci est important puisque l'on fait de la sélection, de création variétale, nous avons besoin de prédire la valeur de nouveaux génotypes.

Les modèles donnent des résultats similaires, mais l'architecture génétique du caractère est importante (simple ou complexe). Le modèle mixte cofRR est toujours le plus intéressant.

4. Données « Réelles »

Nous avons utilisé deux populations : i) un panel de diversité de 279 variétés couvrant la diversité de l'espèce *Vitis vinifera* L., ii) mais aussi une population issue d'un croisement biparental entre les variétés Syrah et Grenache. Nous avons cherché à identifier des marqueurs moléculaires liés à des caractères d'intérêt sur le panel de 279 variétés par des approches de GWAS. Mais nous avons aussi réalisé des prédictions avec 2 dispositifs. Le premier comprend comme population d'entraînement le panel de 279 variétés et comme population de validation des individus issus de la descendance de la Syrah et du Grenache.. Le second est comprend comme population d'entraînement et de validation une population biparentale (cross-validation) de 191 génotypes issus d'un croisement entre les variétés Syrah et Grenache. Pour réaliser nos objectifs nous avons dû génotyper avec la puce SNP Illumina définie dans le cadre du projet Européen GrapeReSeq. Cette puce de génotypage contient 18000 SNPs (Single Nucleotide Polymorphism). Nous avons choisi d'étudier plusieurs caractères complexes, aux héritabilités variables comme la qualité du raisin et la tolérance à la sécheresse. Une partie des données sont issues du projet ANR DL-Vitis. Ces caractères sont premièrement des cibles de la sélection mais également ceux pour lesquels la sélection génomique semble plus prometteuse que la SAM classique (Bernardo et Yu, 2007). Nous ne présenterons ici que les résultats obtenus sur la taille de la baie. Ce caractère est intéressant, son héritabilité est très élevée entre 0,9 et 0,95 (selon les études).

4.1 Détection de marqueurs moléculaires et prédiction en utilisant un panel de diversité

Nous avons phénotypé puis génotypé deux populations, la première correspondant au panel de 279 variétés de *Vitis vinifera* L. pour faire de la génétique d'association et identifier des marqueurs génétiques liés à la taille de la baie. La seconde population est composée de 24 génotypes pris au hasard dans la descendance issus du croisement entre les variétés Syrah et Grenache, il faut considérer ici comme une population de sélection. Cette dernière nous sert à évaluer notre capacité à prédire la taille de la baie d'un génotype nouveau au stade précoce (uniquement avec son ADN).

Après le phénotypage du panel de 279 variétés de *Vitis vinifera* L. et de la population candidate, nous avons mis en relation la taille de la baie (évalué par son poids en gramme) et l'origine du cépage. Il est clair (Figure 8) que la taille de la baie est différente en fonction de l'origine génétique de l'individu. Dans cet échantillon (279), nous retrouvons les trois origines génétiques de l'espèce *Vitis vinifera* L. (WW pour cuve de l'Europe de l'Ouest, WE pour cuve de l'Europe de l'Est et TE pour table de l'Europe de l'Est.).

Les résultats préliminaires indiquent que les gènes impliqués dans la variation de la taille de baie sont différents en fonction de l'origine génétique des individus (Figure 9). A titre d'exemple, dans la population WW neuf marqueurs ont été identifiés (points rouges) alors que ces neuf marqueurs n'ont pas d'impact sur la taille de la baie dans les 2 autres populations. Le même résultat est observé avec la population TE, dans laquelle trois marqueurs spécifiques sont impliqués.

Nous avons ensuite testé la prédiction (prédire la taille de la baie uniquement avec des marqueurs moléculaires issus d'une étape de sélection (GWAS) ou avec toute l'information GS) sur une population candidate (Syrah X Grenache). Le coefficient de corrélation entre les valeurs réelles (mesurées) et les valeurs prédites (déduites de l'information génétique) est différent selon la source des marqueurs

(population utilisée pour les identifier). Ce coefficient est le même (0,35) si les marqueurs utilisés sont ceux identifiés dans la population TE ou dans la population WW, et de 0,1 si l'on utilise la population WE. Étonnamment si l'on cumule l'information issue des trois populations, ce coefficient est faible (0,2).

Ce résultat confirme que ce sont des zones du génome différentes qui impactent la taille de la baie en fonction de l'origine génétique.

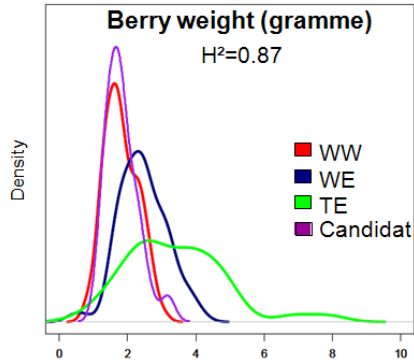


Figure 8 : Taille de la baie en fonction de la population, le panel de diversité (WW, WE et TE) et la descendance syrah Grenache (Candidat)

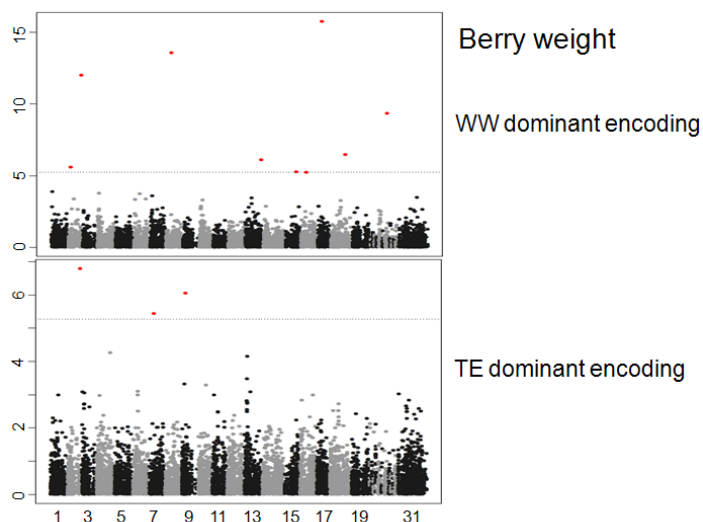


Figure 9: Manhattan plot (en abscisse les marqueurs sur les chromosomes et en ordonnée une valeur correspondant à $-\log$ p-value du marqueur. Plus la valeur est haute plus le marqueur est associé.

En rouge, les marqueurs qui passent le seuil de significativité : zones du génome liées avec la variation de la taille de la baie en fonction de l'origine génétique, les points rouges correspondent aux marqueurs liés.

4.2 Détection de marqueurs moléculaires et prédiction en utilisant une population biparentale

Nous avons aussi utilisé la population de 191 individus issus de la descendance Syrah X Grenache en cross validation (une partie comme population d'entraînement et l'autre comme population de validation). Les résultats présentés pour la taille de la baie montrent une meilleure précision de la prédiction pour les modèles issus de GS (RR et bLR) et le modèle mixte, que lorsque l'approche GWAS est utilisée seule (Figure 10). La précision est aussi meilleure que lors de l'utilisation de panel de diversité où elle plafonne à 0,3. De plus, ici le coût du génotypage est faible puisque nous avons utilisé 150 marqueurs microsatellites. L'utilisation de marqueurs préalablement identifiés comme liés au caractère (Approche QTL) donne des résultats inférieurs à l'utilisation de l'ensemble de l'information génétique (0,5 pour le modèle cof et 0,65 pour les autres modèles).

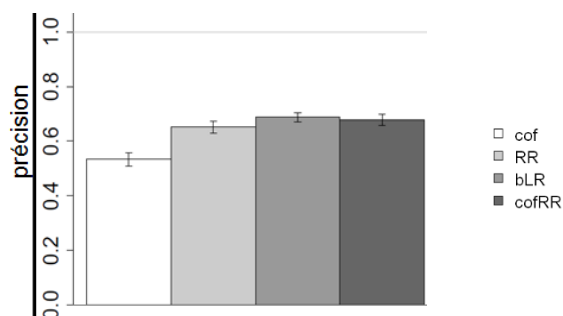


Figure 10 : Précision de la prédiction en cross validation dans la descendance de Syrah et Grenache

5. Utilisation du génotype « Dwarf »

Les croisements Dwarfs (2 HPD x *V. vinifera* et 2 *V. vinifera* X *V. vinifera*) ont été tout d'abord réalisés pour tester les différentes étapes nécessaires à l'utilisation de ce mutant dans les programmes d'amélioration variétale (notamment l'étape de sauvetage d'embryons obligatoire car ce mutant est insensible aux gibbérellines (Figure 11) (Bouquet A., com pers).

Les observations montrent un petit nombre de baies par grappe (~20 à 30), un grand nombre de pépins creux après pollinisation manuelle (~50%) et une faible reprise des embryons (~10%).

La lignée Dwarf ayant été utilisée comme femelle, nous allons tester des croisements inverses pour voir si cela augmente nos pourcentages d'individus acclimatés. Les résultats ne sont pas améliorés quel que soit le sens de croisement. Entre temps, l'INRA de Colmar (UMR SVQV) a développé une technique permettant la mise à floraison 1 an après l'obtention du pépin. Cette technique permettrait d'accélérer le cycle sans utiliser le génotype « Dwarf ».

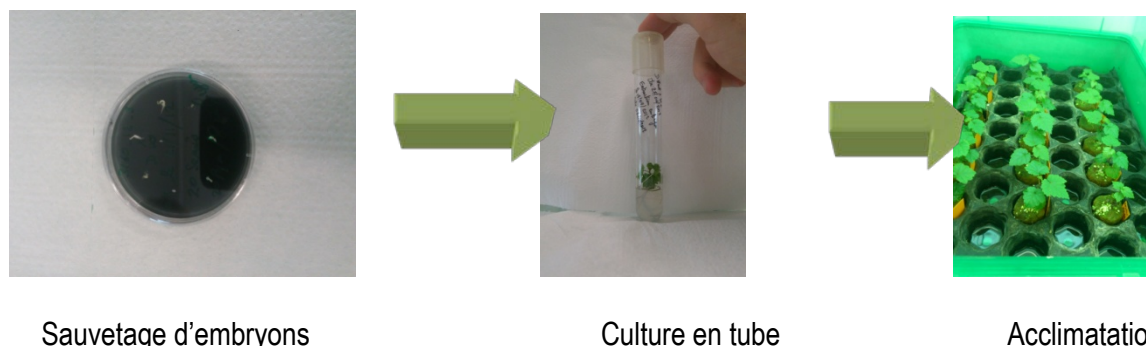


Figure 11 : Les différentes étapes du sauvetage d'embryons

De plus, afin d'obtenir des individus homozygotes pour les loci de résistance à l'oïdium et au mildiou, nous avons récolté les pépins de génotypes pyramidés pour ces résistances (run1-ren3 pour l'oïdium et rpv1-rpv3 pour le mildiou), obtenus dans le cadre de l'UMT Géno-Vigne® ce qui correspond à 3 loci puisque Run1 et Rpv1 sont sur le même locus. Environ 2000 pépins ont été semés parmi les plantes obtenues (très faible pourcentage de germination dû au phénomène de dépression de consanguinité). Nous avons actuellement en serre (première floraison en 2015) 7 génotypes homozygotes pour les 3 loci, 35 génotypes comprenant les 3 loci mais homozygotes pour 2 d'entre eux. Ces génotypes seront utilisés comme géniteurs ; ils permettront d'augmenter la fréquence de descendants pyramidés pour ces 4 gènes.

Finalement nous n'utiliserons pas le génotype Dwarf pour le pyramidage du fait notamment de l'obtention d'individus homozygotes pour les loci de résistance dans un fond génétique « non Dwarf ». Cependant le génotype Dwarf est toujours utilisé pour l'élaboration de populations d'étude qui peuvent

analysées en milieu contrôlé (de fait de sa petite taille) mais aussi pour des approches de sélection récurrente sur des caractères d'architecture génétique plus complexe.

6. Discussion et perspectives

Les résultats obtenus ici permettent de développer de nouvelles pistes de recherche et une réflexion générale sur « la modernisation de la création variétale chez la vigne ».

Même si le coût du génotypage est en constante baisse, alors que le phénotypage s'affine avec un coût qui de ce fait augmente, l'investissement réalisé est-il intéressant ? Afin d'apporter des éléments de réponse à cette question et suite aux résultats obtenus dans ce projet trois scénarii sont proposés. Nous tenterons de les présenter en présentant leurs avantages et leurs inconvénients.

6.1 Le développement d'une population d'entraînement universelle

Cette proposition est très séduisante ; elle permettrait de réaliser un seul investissement important rentable pour plusieurs années, voire des décennies et ce, quels que soient les idéotypes envisagés. Cet investissement comprendrait le génotypage avec 100 000 SNPs et le phénotypage multi-sites d'un grand nombre d'individus (1000). Le coût de génotypage des populations de validation pourrait être moindre grâce à l'utilisation d'algorithmes d'imputations. Au vu de la densité du marquage moléculaire envisagée, les modèles devraient être peu affectés par le nombre de générations. Cette proposition nécessite un investissement très important.

6.2 La prédiction en ségrégation

Nous avons aussi montré que la prédiction de phénotype sur une population en ségrégation donne des résultats très intéressants. Cette étude montre qu'avec un marquage simple (uniquement les cofacteurs identifiés) les prédictions étaient déjà satisfaisantes. Cependant, les modèles développés ne sont pas adaptés à d'autres croisements et le laps de temps entre le génotypage et le phénotypage de la population d'entraînement et l'utilisation des modèles sur une population de validation dépend du phénotype envisagé. Pour la vigne, le couplage avec la technologie développée par l'INRA de Colmar qui permet l'obtention d'une fructification dès la première année peut faire gagner du temps si l'on veut travailler sur des critères relatifs aux baies en serre. Cette stratégie, qui donne les meilleurs résultats, est aussi la moins coûteuse.

6.3 Set de géniteurs élités

Une proposition intermédiaire moins ambitieuse que la population « universelle » mais plus généralisable que « la prédiction en ségrégation » serait de sélectionner un set d'individus qui semble être les futurs géniteurs les plus importants pour répondre aux questions de la filière viticole. Ce set sera comme dans le cas de la population universelle, génotypé et phénotypé en multi-sites.

Conclusion

D'après les résultats obtenus et les outils et méthodologies actuellement disponibles (puce 18 KSNP), l'option d'une « population universelle » nécessite encore un effort afin de développer un outil de génotypage permettant d'atteindre 100 000 marqueurs moléculaires.

L'option « population en ségrégation » qui semble la plus intéressante au regard du rapport investissement / résultats est testée dans le cadre du projet EDGARR (lauréat 2014 de l'AAP CTPS « Semences et Plants »). Cette option est possible car la définition des idéotypes est très précise dans ce projet (des vignes résistance pour des vins rosés).

Cependant, la mise en place d'une population comprenant un set de géniteurs pertinents (*Vitis vinifera* et hybrides, voire espèces apparentées) semble être l'option la plus intéressante à moyen terme.

Références bibliographiques

- Aradhya M.K., Dangl G.S., Prins B.H., Boursiquot J.M., Walker M.A., Meredith C.P., Simon C.J., 2003. Genetic structure and differentiation in cultivated grape, *Vitis vinifera* L. *Genetics Research* 81, 179–192.
- Arroyo-García R., Ruiz-García L., Bolling L., Ocete R., López M.A., Arnold C., Ergul A., Söylemezoğlu G., Uzun H.I., Cabello F. et al., 2006. Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp. *sativa*) based on chloroplast DNA polymorphisms. *Molecular Ecology* 15, 3707–3714.
- Bacilieri R., Lacombe T., Di Vecchi-Staraz M., Laucou V., Le Cunff L., in prep. Genetic structure of the cultivated compartment of *Vitis vinifera* L. *BMC Plant Biol.*
- Barnaud A., Lacombe T., Doligez A., 2006. Linkage disequilibrium in cultivated grapevine, *Vitis vinifera* L. *Theoretical and Applied Genetics* 112, 708–716.
- Bernardo R., Yu J., 2007. Prospects for Genomewide Selection for Quantitative Traits in Maize. *Crop Science* 47, 1082.
- Boss P.K., Thomas M.R., 2002. Association of dwarfism and floral induction with a grape “green revolution” mutation. *Nature* 416, 847–850.
- 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. *Plant J.* 62, 1083–1092.
- Le Cunff L., Fournier-Level A., Laucou V., Vezzulli S., Lacombe T., Adam-Blondon A.F., Boursiquot J.M., This P., 2008. Construction of nested genetic core collections to optimize the exploitation of natural diversity in *Vitis vinifera* L. subsp. *sativa*. *BMC Plant Biology* 8, 31.
- Davey J.W., Hohenlohe P.A., Etter P.D., Boone J.Q., Catchen J.M., Blaxter M.L., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* 12, 499–510.
- 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, e19379.
- Grassi F., Labra M., Imazio S., Spada A., Sgorbati S., Scienza A., Sala F., 2003. Evidence of a secondary grapevine domestication centre detected by SSR analysis. *TAG Theoretical and Applied Genetics* 107, 1315–1320.
- 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, 5.
- Houel C., 2011. Caractérisation de la variation phénotypique de la taille de la baie chez la vigne *Vitis vinifera* L. et approches de génétique d’association et de recherche de traces de sélection pour ce caractère.
- Houel C., Martin-Magniette M.L., Nicolas S., Lacombe T., Le Cunff L., Franck D., Torregrosa L., Conéjéro G., Lalet S., This P. et al. (in prep.). Cell number in ovaries and cell extension after anthesis 1 determine the berry size in grapevine (*Vitis vinifera* L.). *Journal of Experimental Botany*.
- Kang H.M., Sul J.H., Service S.K., Zaitlen N.A., Kong S., Freimer N.B., Sabatti C., Eskin E., 2010. Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics* 42, 348–354.
- Kang H.M., Zaitlen N.A., Wade C.M., Kirby A., Heckerman D., Daly M.J., Eskin E., 2008. Efficient Control of Population Structure in Model Organism Association Mapping. *Genetics* 178, 1709–1723.
- Laucou, V., Lacombe, T., Dechesne, F., Siret, R., Bruno, J.P., Dessup, M., Dessup, T., Ortigosa, P., Parra, P., Roux, C., Santoni, S., Vares, D., Peros, 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.
- Levadoux L., 1956. Les populations sauvages et cultivées de *Vitis vinifera* L. *Annales De L'amélioration Des Plantes* 1, 59–118.

Lijavetzky D., Cabezas J., Ibáñez A., Rodríguez V., Martínez-Zapater J.M., 2007. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. BMC Genomics 8, 424.

Mangin B., Siberchicot A., 2011. Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. Heredity 108, 285–291.

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., et al., 2011. Genetic Structure and Domestication History of the Grape. PNAS 108, 3530–3535.

Myles S., Chia J.M., Hurwitz B., Simon C., Zhong G.Y., Buckler E., Ware D., 2010. Rapid Genomic Characterization of the Genus *Vitis*. PLoS ONE 5, e8219.

Negrul A., 1938. Evolution of cultivated forms of grapes. Comptes Rendus (Doklady) Academie Sciences USSR 18,.

Neuenschwander S., Hospital F., Guillaume F., Goudet J., 2008. quantiNemo: an individual-based program to simulate quantitative traits with explicit genetic architecture in a dynamic metapopulation. Bioinformatics 24, 1552–1553.

Pérez P., de los Campos G., Crossa J., Gianola D., 2010. Genomic-Enabled Prediction Based on Molecular Markers and Pedigree Using the Bayesian Linear Regression Package in R. The Plant Genome Journal 3, 106.

Schaeffer L.R., 2006. Strategy for applying genome-wide selection in dairy cattle. J. Anim. Breed. Genet. 123, 218–223.

Segura V., Vilhjálmsson B.J., Platt A., Korte A., Seren Ü., Long Q., Nordborg M., 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nature Genetics 44, 825–830.

This P., Lacombe T., Cadle-Davidson M., Owens C.L., 2007. Wine grape (*Vitis vinifera* L.) color associates with allelic variation in the domestication gene VvmybA1. Theoretical and Applied Genetics 114, 723–730.

This P., Lacombe T., Thomas M.R., 2006. Historical origins and genetic diversity of wine grapes. Trends in Genetics 22, 511–519.

Jaillon O. et al., 2007. French-Italian public consortium for grapevine genome characterization. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449, 443–467

Velasco R., Zharkikh A., Troggio M. et al., 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS One 12:e1326.

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 proanthocyanidins composition through QTL mapping. BMC Plant Biology 12, 30.

Marguerit E., Boury C., Manicki A., Donnart M., Butterlin G., Némorin A., Wiedemann-Merdinoglu S., Merdinoglu D., Ollat N., Decroocq S., 2009. Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theor. Appl. Genet. 118 (7), 1261–1278.

Cet article est publié sous la licence Creative Commons (CC BY-NC-ND 3.0)



<https://creativecommons.org/licenses/by-nc-nd/3.0/fr/>

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « Innovations Agronomiques », la date de sa publication, et son URL)