



HAL
open science

Identification de marqueurs associés à une résistance stable au mildiou de la pomme de terre par génétique d'association

Nadège Le Garrec

► **To cite this version:**

Nadège Le Garrec. Identification de marqueurs associés à une résistance stable au mildiou de la pomme de terre par génétique d'association. Sciences du Vivant [q-bio]. 2019. dumas-02369540

HAL Id: dumas-02369540

<https://dumas.ccsd.cnrs.fr/dumas-02369540>

Submitted on 19 Nov 2019

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.

CFR Angers CFR Rennes

<p>Année universitaire : 2018-2019 Master Biologie, Agrosociétés Parcours Amélioration, Production, Valorisation du végétal Option : Génétique, Génomique et Amélioration des Plantes</p>	<p>Rapport de stage</p> <p><input type="checkbox"/> d'Ingénieur de l'Institut Supérieur des Sciences agronomiques, agroalimentaires, horticoles et du paysage</p> <p><input checked="" type="checkbox"/> de Master de l'Institut Supérieur des Sciences agronomiques, agroalimentaires, horticoles et du paysage</p> <p><input type="checkbox"/> d'un autre établissement (étudiant arrivé en M2)</p>
---	--

Identification de marqueurs associés à une résistance stable au mildiou de la pomme de terre par génétique d'association

Par : Nadège LE GARREC

Soutenu à Rennes

le 26/06/2019

Devant le jury composé de :

Président : Mélanie JUBAULT

Maître de stage : Florence ESNAULT

Enseignant référent : Mélanie JUBAULT

Rapporteur : Régine DELOURME

Examineur : Maria MANZANARES-DAULEUX

Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celles d'AGROCAMPUS OUEST et l'université de Rennes 1

Ce document est soumis aux conditions d'utilisation

« Paternité-Pas d'Utilisation Commerciale-Pas de Modification 4.0 France »

disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr>



Remerciements

Je tiens à remercier l'UMR IGEPP de Ploudaniel de m'avoir accueillie pour ce stage.

Je tiens à remercier ma maitre stage, Florence Esnault, pour son accueil, pour la confiance qu'elle m'a accordée ainsi que pour le partage de ses connaissances ce qui m'a permis de réaliser les objectifs de ce stage. Je tiens aussi à la remercier pour le temps qu'elle m'a consacré et l'aide qu'elle m'a apporté pour la rédaction de ce rapport.

Je tiens à remercier toute l'équipe de l'UMR IGEPP pour son accueil. Et plus particulièrement les personnes qui m'ont fait découvrir leur travail sur le terrain ce qui m'a permis de découvrir le domaine de l'amélioration de la pomme de terre. Je tiens particulièrement à remercier Sylvie Marhadour pour sa disponibilité et les conseils qu'elle m'a donné.

Je tiens également à remercier Jérôme Morice pour le temps qu'il a consacré pour réaliser la représentation graphique du circos.

Table des matières

Introduction	1
La pomme de terre.....	1
Le mildiou et les moyens de lutte.....	2
L'analyse d'association chez la pomme de terre.....	3
Matériel et méthodes	4
Matériel végétal.....	4
Données phénotypiques.....	5
Données de génotypage.....	5
Structuration de la population	6
Analyse d'association.....	6
Analyse des marqueurs significatifs.....	7
Analyse du déséquilibre de liaison des régions génomiques identifiées.....	7
Résultats	8
Données phénotypiques et génotypiques	8
Structuration de la population	8
Analyse d'association.....	8
Déséquilibre de liaison	12
Discussion	13
Conclusion et perspectives	16
Bibliographie.....	17
Annexes	

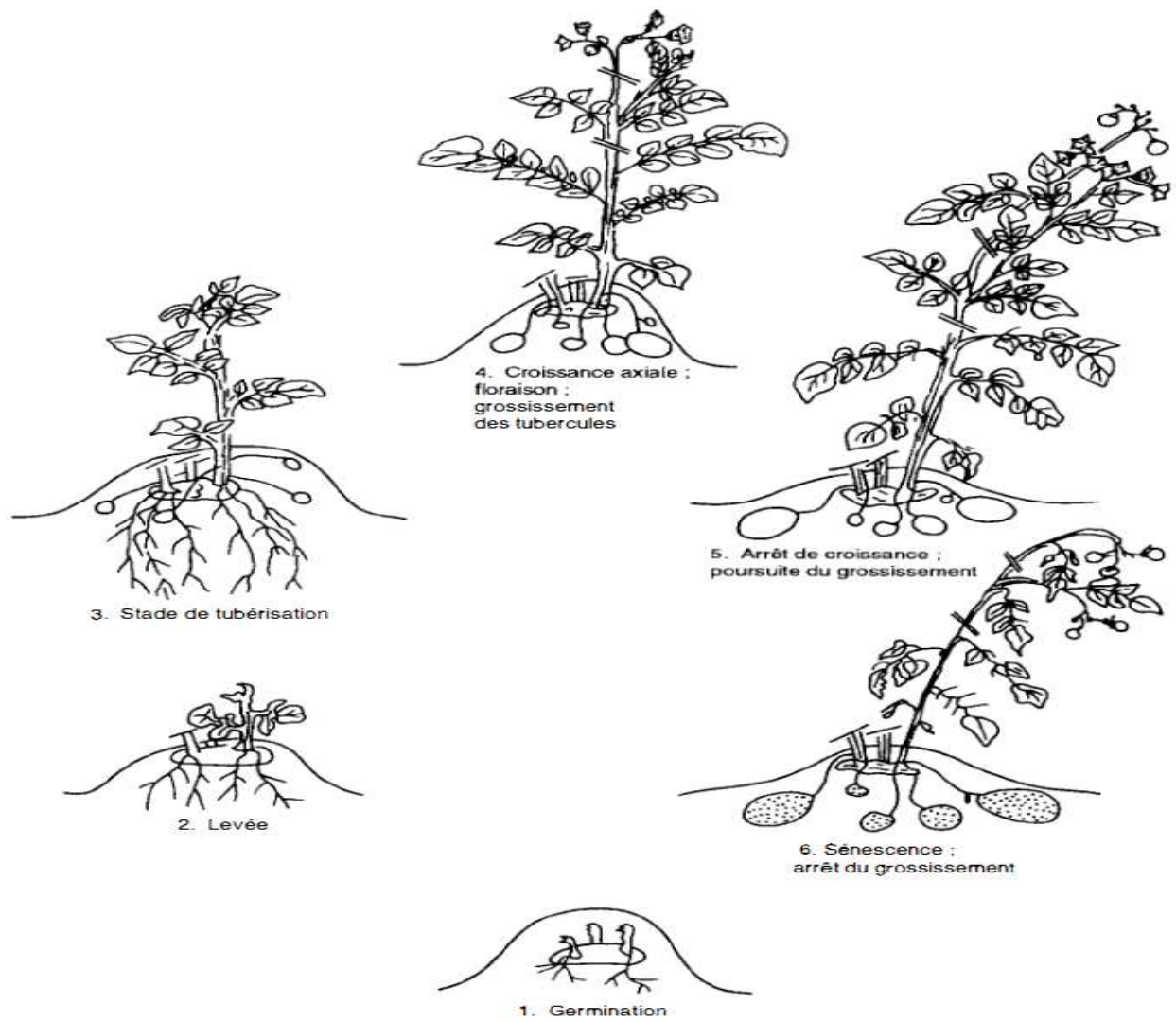


Figure 1: Schéma du cycle végétatif de la pomme de terre (Rousselle, Robert, Crosnier 1996).

Introduction

La pomme de terre

La pomme de terre cultivée (*Solanum tuberosum* L.) est la quatrième espèce d'importance agronomique au niveau mondial. En France cette culture représentait 194 055 ha en 2017 (Agreste-statistiques). Cette espèce a été domestiquée il y a environ 9 000 ans en Amérique du Sud dans la région du lac Titicaca. Le centre de variabilité maximum se situe dans les Andes (Rousselle, Robert, Crosnier 1996). La pomme de terre a été introduite en Europe à la fin du XVI^e siècle d'abord dans les îles Canaries, puis en Espagne et dans les Iles Britanniques.

L'origine de la pomme de terre cultivée n'est pas encore totalement résolue. Les espèces sauvages et cultivées de la pomme de terre font partie du genre *Solanum*, section *Petota*. *Solanum tuberosum* L. comprend des landraces cultivées en Amérique du Sud et réparties en deux groupes qui sont *Chilotanum* et *Andigenum*, et les variétés dites modernes cultivées hors Amérique du Sud. L'origine de *S. tuberosum* groupe *Andigenum* serait un événement de domestication unique à partir d'espèces sauvages du complexe *S. brevicaulis*. Cet événement aurait eu lieu dans le sud du Pérou. Ensuite les formes triploïdes et tétraploïdes du groupe *Andigenum* auraient probablement diverses origines. Elles ont pu se former suite à des événements de polyploïdisation avec des gamètes non réduits et/ou des hybridations interspécifiques (Spooner et al. 2014). Par la suite la différenciation du groupe *Andigenum* implique des introgressions à partir d'espèces sauvages qui ne font pas partie du complexe *S. brevicaulis* ce qui a permis d'apporter aux espèces cultivées une capacité d'adaptation à des conditions écologiques plus diverses (Spooner et al. 2014). Ainsi *S. tuberosum* groupe *Chilotanum* s'est différencié à partir du groupe *Andigenum* et est adapté aux jours longs. Les dernières études montrent que les variétés cultivées dites modernes sont plus proches génétiquement du groupe *Chilotanum*.

Chez la pomme de terre l'organe d'intérêt agronomique est le tubercule. Le tubercule sert aussi à la multiplication végétative. Le cycle végétatif de la pomme de terre se déroule en six grands stades (Fig.1). A la fin de ce cycle les tubercules produits ont besoin d'une période de repos végétatif (entre 17 et 40 semaines) pour que les bourgeons puissent se développer en germes (Rousselle, Robert, Crosnier 1996).

La pomme de terre est une espèce autotétraploïde ($2n=4x=48$), fortement hétérozygote dont le génome a été séquencé en 2011 à partir d'un clone monoploïde doublé dérivé d'une accession de *S. tuberosum* groupe *Phureja*. La plupart des caractères sont à déterminisme polygénique et ont une ségrégation de type tétrasomique. Lors de la méiose, plusieurs types de bivalents et même des multivalents peuvent se former entre les quatre chromosomes homologues (Schmitz Carley et al. 2017). Sa nature autotétraploïde et l'hérédité tétrasomique rendent les analyses génétiques compliquées (Rousselle, Robert, Crosnier 1996).

Les critères de sélection sont nombreux pour répondre aux besoins du marché. La production de pomme de terre se divise en trois filières : pommes de terre de consommation, pommes de terre féculières (destinées aux industries : papetière, agro-alimentaire et pharmaceutique) et plants. La filière pomme de terre de consommation représente la plus grande partie de la production, 75% (source UNPT : Union Nationale Des Producteurs de Pomme de Terre). La sélection doit donc répondre aux besoins de ces différentes filières. Des critères de sélection comme l'adaptation au lieu de culture, la productivité, la qualité des tubercules liée aux utilisations ainsi que la résistance aux ravageurs et aux pathogènes sont travaillés (Rousselle, Robert, Crosnier 1996)... Nous allons nous intéresser plus particulièrement dans ce rapport à

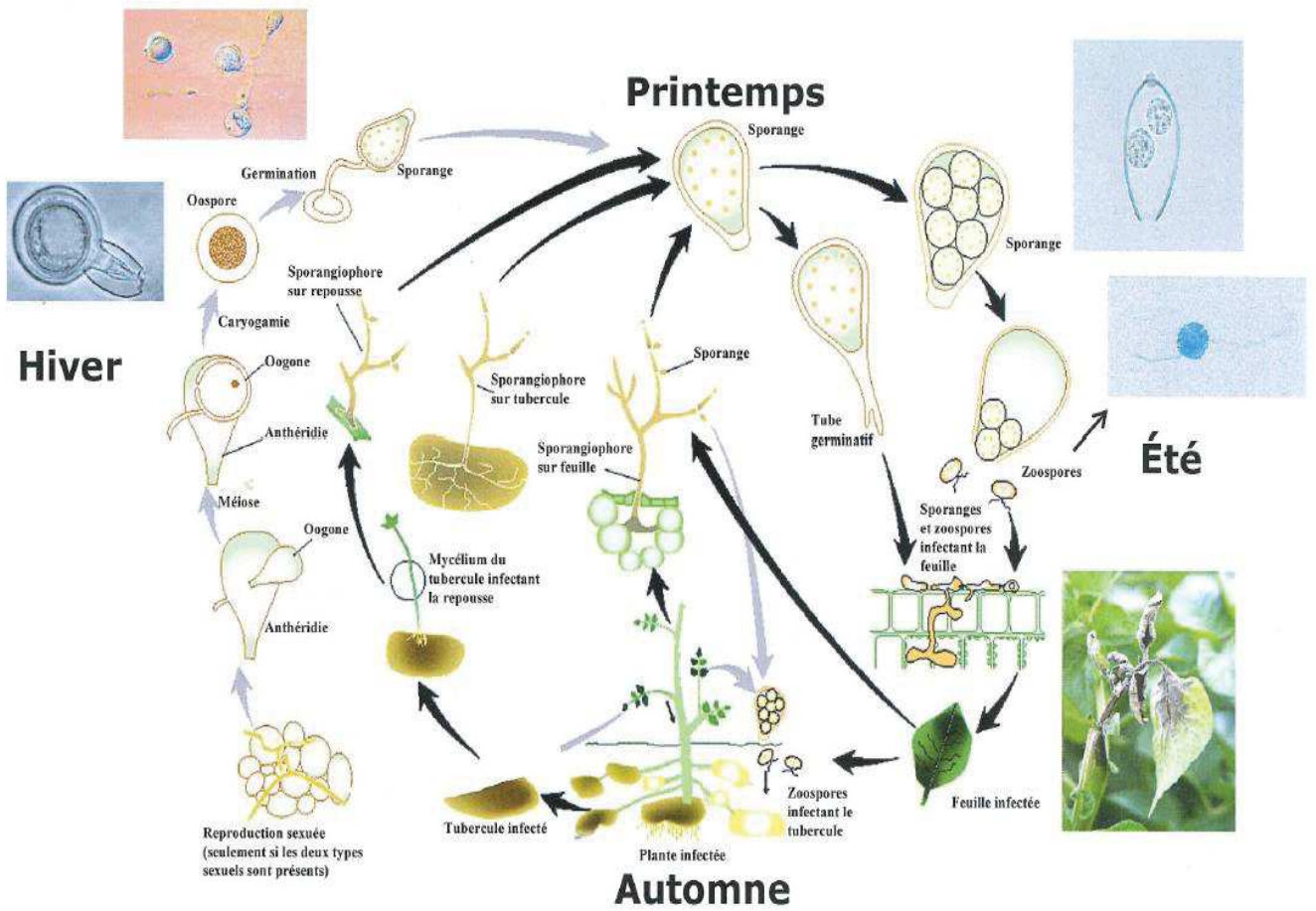


Figure 2: Cycle de développement de *P. infestans* en fonction des saisons. Les deux types de reproduction, sexué et asexué, y sont présentés. (Source : <http://www.eucablight.org>) Photos : R. Corbière (UMR IGEPP, INRA Bretagne Normandie).

la résistance à l'agent pathogène responsable du mildiou. Dans un contexte de changement climatique, le développement de variétés adaptées à une agriculture durable est essentiel.

Le mildiou et les moyens de lutte

Le mildiou est une des maladies qui causent le plus de perte de rendement sur la culture de pomme de terre. Cette maladie est causée par l'oomycète *Phytophthora infestans*. Elle peut provoquer de grands dégâts sur les cultures. L'oomycète peut se développer sur différents organes de la plante : feuilles, tiges et tubercules. Sur les parties aériennes les symptômes se présentent sous forme de nécroses. Le développement de la maladie (Fig.2) peut ainsi aboutir à une destruction complète des parties aériennes des plantes. Les tubercules contaminés lors de la phase de grossissement (Fig.1) eux présentent des pourritures sèches qui se développent de l'épiderme vers l'intérieur du tubercule (Rousselle, Robert, Crosnier 1996). Les tubercules sont contaminés à partir des sporocystes et/ou zoospores produits sur le feuillage qui passent dans le sol par lessivage.

L'inoculation primaire peut venir de différentes origines. L'oomycète peut se conserver pendant l'hiver sur les tubercules restés dans le sol ou dans des tas de déchets au bord des parcelles, ce qui représente la principale source d'inoculum primaire. Il peut aussi venir des tubercules utilisés comme plants. La propagation de la maladie est favorisée par des conditions météorologiques humides et fraîches (Rousselle, Robert, Crosnier 1996). Un champ entier peut être contaminé en quelques jours (Fry 2008).

Ce pathogène présente deux types sexuels nommés A1 et A2. La présence des deux types sexuels peut conduire à une reproduction sexuée (Fig.2). Cependant le type de reproduction le plus observé est la reproduction asexuée (Fig.2). Il a été observé que les populations évoluaient rapidement avec l'apparition de nouvelles lignées clonales et la disparition d'autres. Les populations peuvent connaître de nombreux changements au cours d'une année dûs à leur capacité de dispersion et d'adaptation (Mariette et al. 2016). En effet le mildiou est un agent pathogène à forte capacité d'évolution, capable de s'adapter aux facteurs limitants son développement. Il peut ainsi développer des résistances aux fongicides et contourner des gènes de résistance des variétés (Colon et al. 1995). La structure génétique des populations françaises montre que ces populations sont essentiellement clonales (Corbière et al. 2016). Toutefois il existe de la variabilité au sein d'une même lignée clonale notamment pour la vitesse de sporulation et le taux de croissance. Ce qui permet à ces variants de s'adapter aux variétés présentes dans un lieu (Jeffrey, Jinks, Grindle 1962). En France il existe actuellement quatre lignées clonales dominantes qui sont 13_A2, 6_A1, 1_A1 et 37_A2.

Il existe différents moyens de lutte pour faire face à ce pathogène. La lutte culturale et sanitaire consiste à réduire les sources d'inoculum primaire en plantant des plants sains et en éliminant les tas de déchets qui bordent les champs suite à la récolte. Le moyen de lutte principal est la lutte chimique. La culture de cette espèce présente un fort IFT (Indice de Fréquence de Traitement), 18.2 en 2014 (Agreste) dont 14.4 correspond à l'IFT fongicide. L'application de fongicide est principalement faite pour lutter contre le mildiou. Ces produits phytosanitaires présentent des inconvénients. La grande plasticité des populations de *P. infestans* et l'utilisation en grande quantité de ces produits ont conduit au développement de résistance à certains fongicides (Fry 2008). La lutte chimique représente aussi un coût élevé pour les producteurs. Les fongicides ont aussi un impact environnemental avec notamment la pollution des eaux et sur la santé des utilisateurs (Haverkort et al. 2008). Ces inconvénients et la demande de diminution de l'utilisation des produits phytosanitaires dans le cadre du plan Ecophyto II,

montrent que la lutte chimique ne peut pas être un outil de lutte durable et unique contre le mildiou.

La lutte génétique a aussi été développée dans les programmes de sélection pour faire face au mildiou. Il existe deux grands types de résistance au mildiou, l'une spécifique aux souches et l'autre non spécifique. La résistance spécifique dépend de la présence de gènes de résistance dits R. C'est une résistance à déterminisme génétique simple. Dans cette situation la plante ne présente pas ou très peu de symptômes de la maladie. La résistance non spécifique est associée à une résistance quantitative, c'est le résultat de l'action de plusieurs gènes (Lindqvist-Kreuze et al. 2014). Jusqu'à présent 35 gènes R ont été identifiés chez différentes espèces sauvages apparentées de la pomme de terre dont 11 chez *S. demissum*. Les premiers gènes R découverts dans le début du XXème siècle ont été introduits dans les variétés commerciales. Mais la culture à grande échelle de variétés avec ces gènes R a conduit à un contournement de ce type de résistance par le pathogène (Rousselle, Robert, Crosnier 1996). Une des solutions proposées a été de créer des variétés avec plusieurs gènes R. Cependant cette solution n'a pas permis d'empêcher les contournements de ces gènes R (Fry 2008). La résistance génétique la plus durable dans le temps est la résistance quantitative (Colon et al. 1995; Van Der Plank 1971). La stabilité de la résistance quantitative peut être évaluée dans le temps et dans l'espace. Son évaluation nécessite la mise en place d'expérimentations sur différents sites sur plusieurs années avec les mêmes génotypes. Il a été montré que certains génotypes conservaient un niveau de résistance élevée au fil des années. La stabilité dans l'espace a été peu étudiée et seulement dans des zones géographiques limitées (Forbes et al. 2005). L'identification de sources de résistance durable est un enjeu majeur pour lutter contre le mildiou.

De nombreux travaux d'identification de résistances polygéniques ont été réalisés. Une méta-analyse des QTL identifiés dans la bibliographie a permis de mettre en évidence la présence de 24 méta-QTL répartis sur tous les chromosomes (Danan, Veyrieras, Lefebvre 2011). Toutefois, la majorité de ces QTL ont des contributions à la variation du caractère inférieures à 20%. D'autre part, il a été observé que la résistance quantitative dans certains cas est liée à une maturité tardive des plantes. Comme les sélectionneurs cherchent plutôt à créer des variétés avec une maturité précoce, c'est-à-dire un temps de production de tubercule plus court (Danan, Veyrieras, Lefebvre 2011), il faut donc identifier des QTL de résistance au mildiou qui ne soient pas associés à une maturité tardive.

L'analyse d'association chez la pomme de terre

Le caractère polygénique des résistances quantitatives les rend plus difficiles à sélectionner, en particulier chez une espèce autotétraploïde. Il est alors intéressant d'identifier des marqueurs génétiques associés à ce caractère qui permettront d'aider à sélectionner des individus présentant une résistance quantitative. L'introduction de résistances polygéniques dans les variétés actuelles pourrait permettre d'obtenir une résistance plus stable et plus durable contre le mildiou de la pomme de terre (Danan, Veyrieras, Lefebvre 2011). La génétique d'association génome complet est une méthode d'analyse de collections de ressources génétiques qui permet d'identifier des marqueurs moléculaires en déséquilibre de liaison avec le gène impliqué dans le caractère d'intérêt. (Stich et al. 2013). Afin de réduire le risque de fausses associations dans ce type d'analyse, il est important de prendre en compte la structuration de la population.

La pomme de terre est une espèce bien adaptée à la génétique d'association car le déséquilibre de liaison y est étendu (Sharma et al. 2018). Ainsi, différentes études ont permis d'identifier des marqueurs pour des caractères de qualité du tubercule (Douches et al. 2014; Sharma et al. 2018) ou de résistance au mildiou chez la pomme de terre (Lindqvist-Kreuze et al. 2014). Les

difficultés de l'analyse d'association sur une telle espèce sont dues à sa nature autotétraploïde et à son génome fortement hétérozygote.

Les SNP (Single Nucleotide Polymorphism) sont des marqueurs appropriés pour l'analyse d'association. (Xu et al. 2011). La ségrégation d'un locus biallélique chez des individus autotétraploïdes conduit à trois classes génotypiques hétérozygotes (Aaaa, AAaa et AAAa) et deux classes homozygotes (aaaa et AAAA) (Spooner et al. 2014). Les analyses d'association sur les espèces autotétraploïdes ont d'abord été faites en utilisant des logiciels adaptés aux diploïdes. Dans ces analyses, les trois classes hétérozygotes pour un marqueur biallélique étaient regroupées pour ne prendre en compte qu'une seule classe d'hétérozygote comme pour une espèce diploïde. Maintenant de nouveaux logiciels sont disponibles permettant d'évaluer le dosage allélique et ainsi de distinguer les trois classes d'hétérozygotes (Voorrips, Gort, Vosman 2011; Schmitz Carley et al. 2017). De plus un logiciel de génétique d'association adapté aux autotétraploïdes a été développé qui permet de prendre en compte les cinq classes génotypiques et de proposer différents modèles d'action des gènes (Rosyara et al. 2016).

Il existe une puce appelée « SolCAP » (Solanaceae Coordinated Agricultural Project) qui permet le génotypage de 8 303 SNP bialléliques répartis tout au long du génome de la pomme de terre. Elle permet une bonne couverture de tout le génome avec une haute densité de marqueurs. Cette puce comprend des SNP identifiés à partir de données de transcriptome de variétés de pomme de terre d'Amérique du Nord (Hamilton et al. 2011). Il a été montré que cette puce permettait aussi d'analyser des génotypes européens (Stich et al. 2013) et elle a déjà été utilisée sur le matériel de l'INRA de Ploudaniel (Esnault et al. 2012; Marhadour et al. 2012).

Le sujet de mon stage est intégré au projet PoTStar 'Stabilité des résistances au mildiou de la pomme de terre et identification de marqueurs associés' financé dans le cadre des actions de recherche du plan Ecophyto. Les objectifs de ce projet sont 1. d'identifier des sources de résistances stables au mildiou de la pomme de terre au sein de la collection du Centre de Ressources Biologiques BrACySol de l'INRA de Ploudaniel, 2. d'identifier les facteurs génétiques et de développer des marqueurs associés à cette résistance stable utilisables en sélection assistée par marqueurs, et 3. d'analyser la réponse des populations de mildiou face à ces résistances. Mon stage porte sur le deuxième volet de ce projet. La stratégie utilisée est une méthode de génétique d'association. L'acquisition des données de génotypage et de phénotypage a été faite avant le stage sur 258 géniteurs améliorés issus des programmes de l'INRA. Les données de phénotypage illustrant la progression du mildiou sur les génotypes au champ sont issues de deux années d'expérimentation réalisées sur cinq sites situés dans des zones géographiquement différentes où les conditions météorologiques et les populations de mildiou sont différentes. Ainsi c'est une résistance stable dans l'espace qui est étudiée ici. Les données de génotypage ont été obtenues en utilisant la puce SolCAP de 8 303 SNPs. Dans le cadre de ce stage ces données sont exploitées pour réaliser l'analyse de structuration génétique de la collection étudiée et l'analyse d'association permettant d'identifier les marqueurs SNP associés à la résistance au mildiou.

Matériel et méthodes

Matériel végétal

Le matériel végétal utilisé est constitué de 287 génotypes au total. Ce matériel comprend 258 géniteurs améliorés maintenus au sein du centre de ressources biologiques BrACySol (Brassica

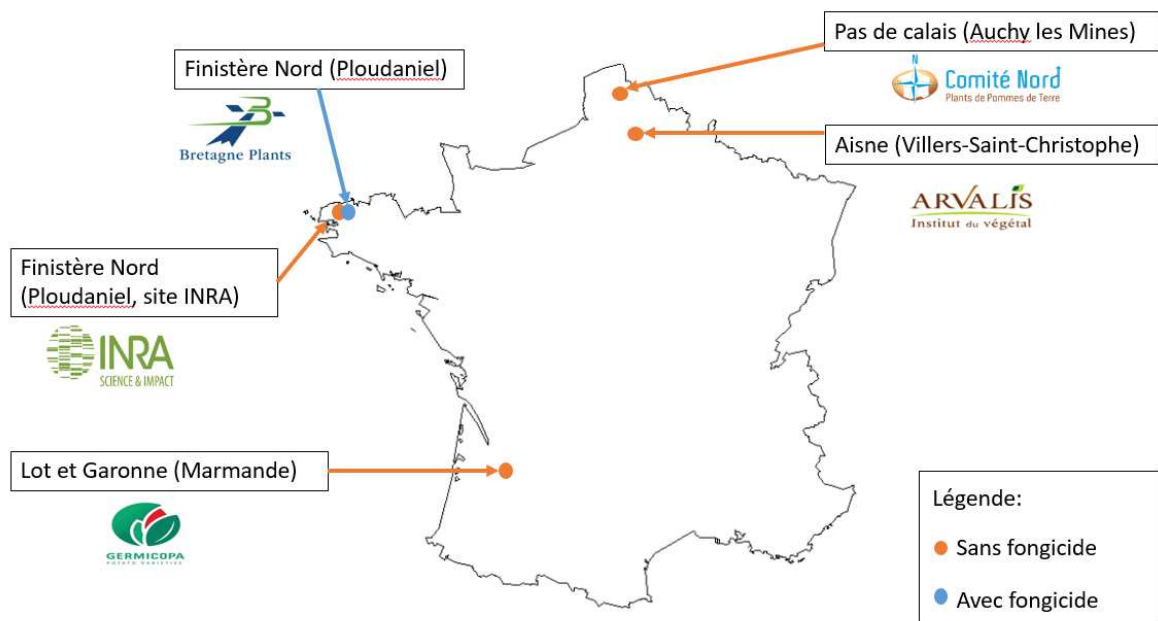


Figure 3: Carte des sites de phénotypage avec les noms des partenaires du projet qui ont effectué les expérimentations. Les points orange montrent les sites où il n'y a pas eu d'application de fongicide. Le point bleu montre le site où une couverture fongicide partielle a été appliquée.

Allium Cynara Solanum) de l'INRA. Les géniteurs améliorés sont issus de neuf programmes de recherche. Ces géniteurs ont été sélectionnés pour la résistance à différents pathogènes : le mildiou (117), le nématode à kyste *Globodera pallida* (51), les nématodes à galles *Meloidogyne* (15), la double résistance mildiou et *G. pallida* (8), des bactéries de type *Pectobacterium* (32), la fusariose (9), des maladies à virus (8). Certains géniteurs ont été sélectionnés pour des caractères de qualité du tubercule dont le sucrage à basse température (10) et la diminution du noircissement à la cuisson (8). Les génotypes restants se divisent en deux catégories : des variétés témoins qui ont des comportements connus vis-à-vis du mildiou avec différents niveaux et types de résistance (18), les hôtes différentiels qui permettent d'avoir des informations sur la virulence des souches de mildiou (11).

Données phénotypiques

Les données de phénotypage ont été obtenues à partir d'essais au champ. Les essais ont été réalisés sur quatre sites qui présentent des conditions climatiques différentes : Finistère Nord, l'Aisne, Pas de Calais et Lot-et-Garonne (Fig.3). Dans le Finistère Nord il y a un 5eme essai dans lequel quelques traitements fongicides ont été appliqués afin d'évaluer les génotypes en condition de pression d'inoculation plus faible. Le dispositif expérimental comporte deux répétitions qui sont en blocs randomisés et des rangs d'infecteurs avec la variété sensible Bintje pour favoriser le développement de l'épidémie. Les essais ont été répétés sur deux ans (2017 et 2018).

Le protocole initial prévoyait que les essais soient en conditions d'inoculation naturelle. Toutefois des inoculations artificielles ont dû être faites sur certains sites en 2017 du fait de conditions climatiques trop sèches et donc non favorables au développement de la maladie.

Les notations de progression de la maladie ont été faites selon le protocole de Pellé et al. (2005). La première notation se fait lors de l'observation de la première tâche foliaire. Une notation de pourcentage de destruction du feuillage est ensuite réalisée toutes les semaines. Des courbes de progression de la maladie ont été tracées à partir de ces notations en fonction du temps. Différents paramètres ont été calculés : l'aire sous la courbe de progression de la maladie (AUDPC : Area Under Disease Progress Curve), delta t qui correspond à la mesure du retard de l'apparition de l'épidémie par rapport à un témoin sensible (Bintje) et delta a qui est la mesure de la différence de vitesse d'évolution de l'épidémie sur chacun des génotypes par rapport au témoin sensible (Andrivon, Pellé, Ellissèche 2006). Les valeurs d'AUDPC ont été normalisées selon la méthode de Fry (1978) et donnent les relative AUDPC (rAUDPC).

Des histogrammes représentant la distribution des rAUDPC selon les sites ont été réalisés pour chaque année.

Données de génotypage

Les données de génotypage ont été obtenues au préalable à partir de la puce SolCAP 8K (8303 SNPs) avec la technique de séquençage Illumina. Il y a des données de génotypage pour les 258 géniteurs améliorés et 16 variétés témoins (les deux variétés témoins sensibles n'ont pas été génotypées). Une variété témoin a été répétée trois fois.

La classe génotypique des individus pour chaque SNP a été attribuée avec deux méthodes. La première méthode est celle de Hirsch et al. 2013 qui permet de définir les cinq classes génotypiques en fonction du niveau de fluorescence de chacun des deux allèles. Les bornes des différentes classes génotypiques ont déjà été définies par ces auteurs pour 5032 SNP. Une

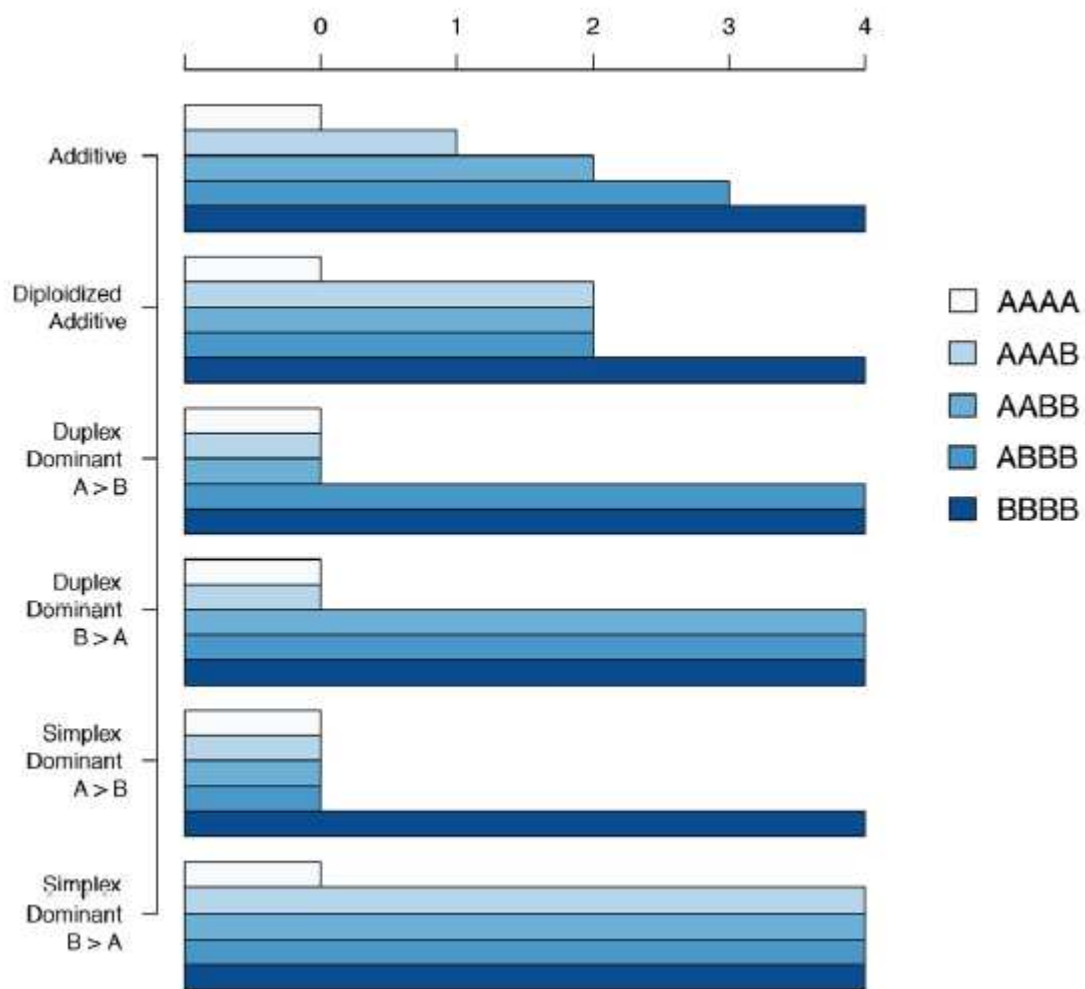


Figure 4: Représentation graphique des différents modèles testés dans le package GWASpoly pour expliquer l'effet des SNP sur le caractère étudié (Rosyara et al. 2016).

deuxième méthode a été utilisée afin d'obtenir le maximum de données de génotypage. C'est un package du logiciel R : FitTetra qui a été développé par Voorrips, Gort, Vosman (2011). Ce logiciel recherche le modèle qui permet d'expliquer au mieux la répartition des individus dans les cinq classes génotypiques à partir du rapport d'intensité des deux allèles (Voorrips, Gort, Vosman 2011). Pour chaque SNP les résultats de l'analyse qui génèrent le moins de données manquantes ont été conservés.

Ensuite les données obtenues sur les SNP ont été filtrées : concordance entre les deux méthodes d'analyse, cartographie à un seul locus (Felcher et al. 2012), moins de 20% de données manquantes et concordance des résultats pour le témoin répété. Ainsi les données pour 5274 SNP ont été exploitées pour les analyses d'association. Les positions physiques des SNP de la puce SolCAP sur le génome de référence ont été obtenues d'après Sharma et al. (2013).

La fréquence de l'allèle mineur a été calculée pour chaque SNP et un histogramme représentant la distribution de la fréquence de cet allèle mineur a été réalisé.

Structuration de la population

Une analyse de structuration génétique de la collection des 258 géniteurs améliorés avec les 16 variétés résistantes a été réalisée en utilisant les données de génotypage d'une sélection de dix SNP répartis le long de chaque chromosome.

La méthode de type bayésienne implémentée dans le logiciel Structure 2.3.4® (Pritchard, Stephens, Donnelly 2000) a été utilisée. Les paramètres choisis sont un modèle avec admixture et les fréquences alléliques corrélées entre les populations. Le nombre de burn-in est de 50 000 suivi de 100 000 répétitions de MCMC (Markov Chain Monte Carlo). Le nombre de sous-populations testées va de 1 à 10. Pour chaque test de K sous-populations, cinq répétitions ont été effectuées. Le nombre de sous-populations, K, a été déterminé à l'aide du graphique du delta K (probabilité de la moyenne de la variation de K) en fonction du nombre de sous-populations testées (Evanno, Regnaut, Goudet 2005). Ce graphique a été obtenu avec Structure Harvester (Earl, vonHoldt 2012).

Analyse d'association

Le trait étudié est le rAUDPC. Pour chaque site une valeur unique de rAUDPC par génotype a été calculée à partir des résultats d'une analyse de variance ayant pour modèle :

$$Y_i = \mu + G + R + e$$

Y_i : valeur de rAUDPC de l'individu i

μ : moyenne

G : effet du génotype

R : effet de la répétition

e : résidus

Les analyses de variance ont été réalisées avec le logiciel R®.

Des BLUPs (Best linear unbiased predictors) ont été calculés avec le package lme4, ce qui permet d'avoir un estimateur par génotype. Le modèle à effets aléatoires utilisé est :

$$Y_i = \mu + 1|G + 1|S + 1|G*S + 1|R\%in\%S + e$$

Y_i : valeur de rAUDPC de l'individu i

μ : moyenne

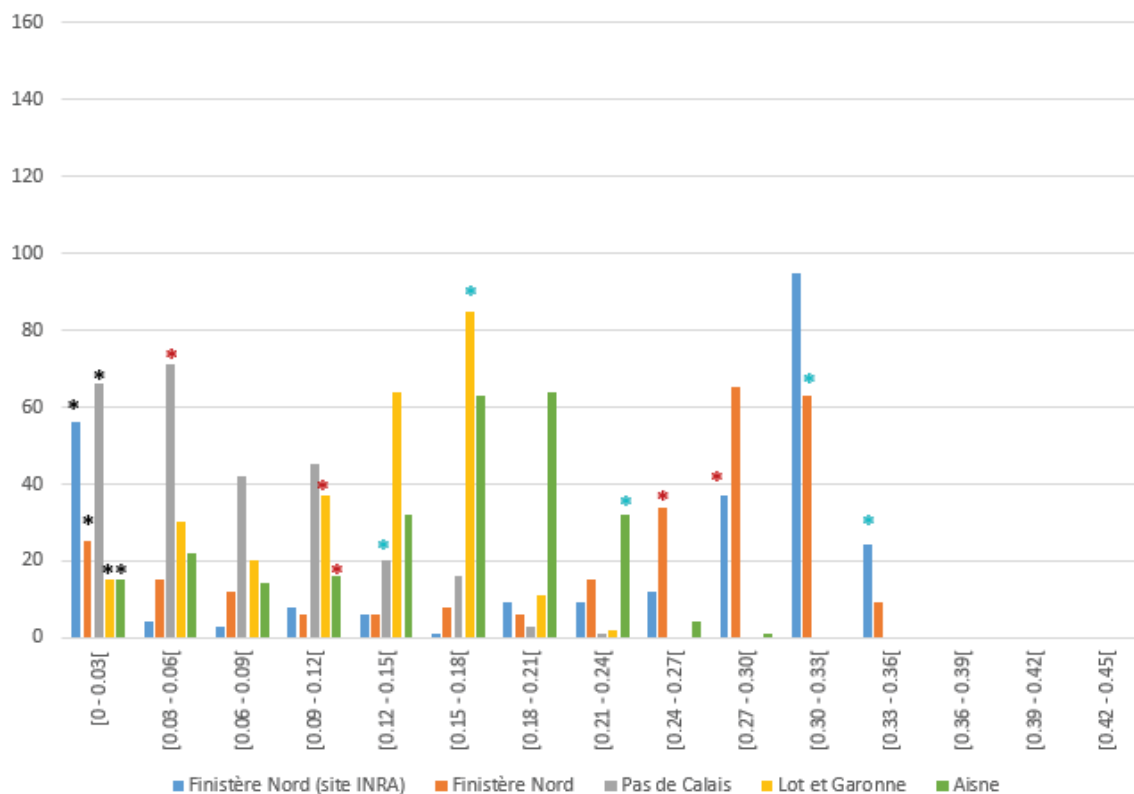


Figure 5: Distribution des valeurs des rAUDPC par site en 2017. Certains témoins sont représentés par des points de couleur : en noir, une variété résistante (Sarpo Mira) ; une variété de résistance intermédiaire en rouge (Arka) ; en bleu, une variété sensible (Gwenne).

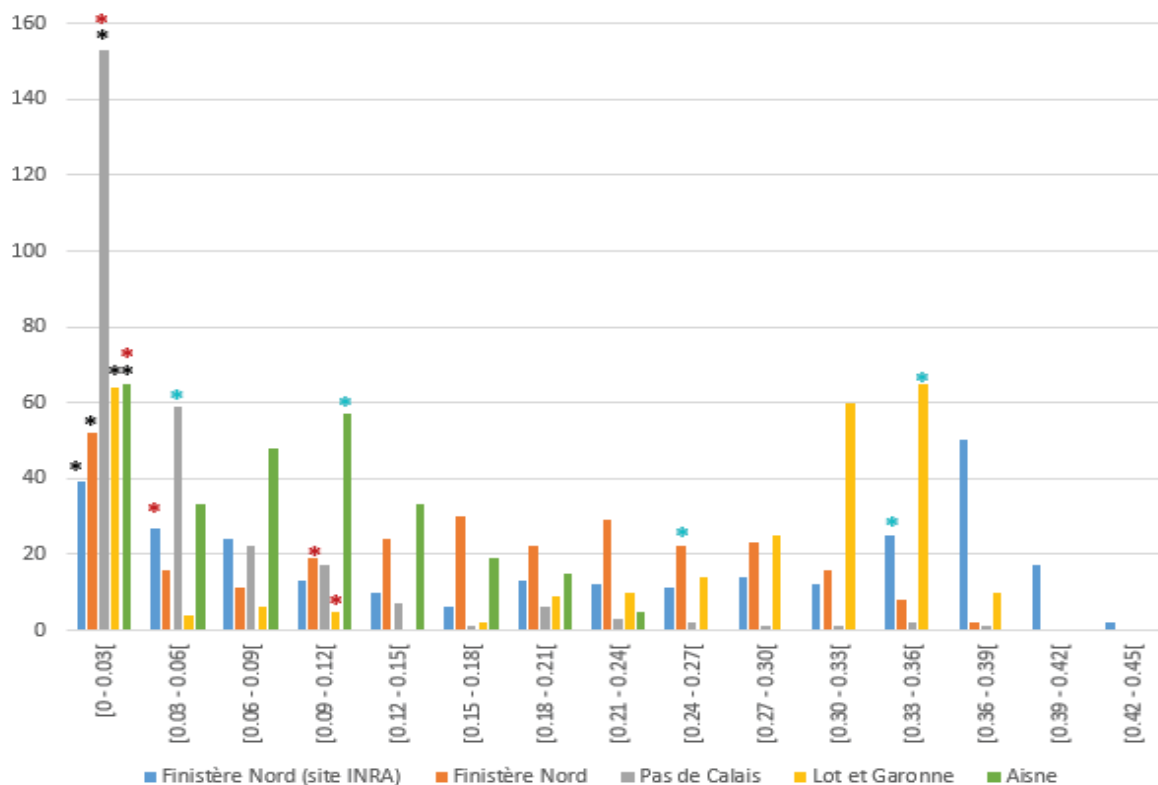


Figure 6: Distribution des valeurs des rAUDPC par site en 2018. Certains témoins sont représentés par des points de couleur: en noir, une variété résistante (Sarpo Mira) ; une variété de résistance intermédiaire en rouge (Arka) ; en bleu, une variété sensible (Gwenne).

G : effet du génotype
S : effet du site
G*S : effet de l'interaction entre le génotype et le site
R%in%S : effet de la répétition niché dans l'effet site
e : résidus
Ainsi un estimateur a été calculé par génotype par année.

Puis des analyses d'association site par site pour chaque année ont été réalisées d'une part et une par année avec les BLUPs d'autre part à l'aide du package GWASpoly du logiciel R (version 3.4.4) (Rosyara et al. 2016) adapté aux tétraploïdes. Un modèle linéaire mixte a été utilisé en prenant en compte l'apparentement par le calcul d'une matrice de parenté (realized relationship matrix) et la structuration de la population obtenue avec l'analyse Structure. Nous avons également utilisé les paramètres par défaut qui considèrent que la fréquence de l'allèle mineur doit être supérieure à 0.05 et que la fréquence maximale d'une classe génotypique ne doit pas être supérieure à 0.95. Ainsi, l'analyse a pris en compte 4 504 SNP. Dans cette analyse différents modèles d'action des SNP détectés sont testés en plus du modèle général (Fig.4). Dans le modèle général, l'effet de chaque classe génotypique est arbitraire. Dans le modèle additif, l'effet du SNP est proportionnel au dosage de l'allèle mineur. Le modèle simplex dominant considère que les trois classes hétérozygotes sont équivalentes à une des classes homozygotes. Il y a deux classes homozygotes, il y a donc deux types de modèles simplex dominant. Le dernier modèle est le modèle duplex dominant. Il peut être de deux types : la classe hétérozygote (AABB) a un effet similaire soit avec les classes AAAB et AAAA, soit avec les classes ABBB et BBBB (Rosyara et al. 2016). Une correction de type FDR (False Discovery rate) a été appliquée comme correction de tests multiples pour éviter la détection de faux positifs. Le FDR correspond à la proportion de résultats de tests faussement positifs parmi tous les tests significatifs (Balding 2006). Le seuil de significativité fixé était de 0.05.

Analyse des marqueurs significatifs

Une analyse plus approfondie des SNP les plus significatifs (identifiés sur plusieurs sites et/ou au cours des 2 années,) a été réalisée. Ainsi, pour ces SNP, l'effet de chaque classe génotypique a été étudié site par site par analyse de variance selon le modèle suivant :

$$Y_i = \mu + M + e$$

Y_i : valeur de rAUDPC de l'individu i
 μ : moyenne
 M : effet du marqueur
 e : résidus

Un test de comparaison de moyenne par paire, test de Tukey, a été réalisé sur les résultats de l'analyse de variance avec le package R Agricolae.

Analyse du déséquilibre de liaison des régions génomiques identifiées

Afin d'exploiter les données de dosage allélique et ainsi prendre en compte les 3 classes d'hétérozygotes de manière indépendante, le déséquilibre de liaison entre paires de SNP d'un même chromosome identifiés dans l'analyse d'association a été évalué à partir du coefficient de corrélation de Pearson. Le package Hmisc a été utilisé. Ensuite les p-values obtenues avec les tests de Pearson ont été corrigés pour tests multiples par le calcul de q-value avec le package qvalue de R (Storey, Tibshirani 2003). Le seuil de significativité retenu pour le DL est de 0.01.

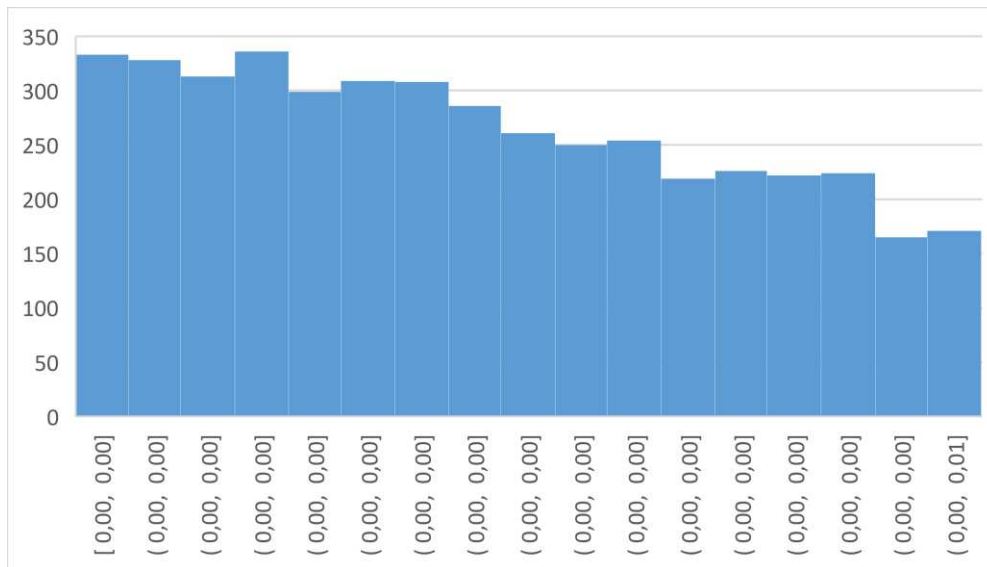


Figure 7: Distribution de la fréquence allélique de l'allèle mineur ($MAF > 0.05$) des 4504 SNP dans la collection.

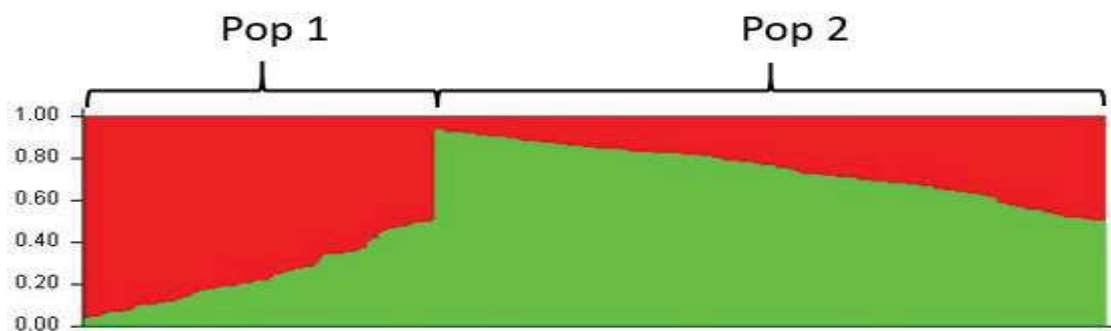


Figure 8 : Graphique représentant la structuration de la population des 273 génotypes de pomme de terre pour une structuration de la population avec $K=2$. Chaque génotype est représenté par une barre. La proportion de couleur dans chaque barre montre l'appartenance de chaque individu à chacune des deux sous-populations.

Des graphiques ont été construits à l'aide du package corrplot.

Résultats

Données phénotypiques et génotypiques

Les figures 5 et 6 représentent la distribution des rAUDPC selon les sites en 2017 et 2018 respectivement. Une valeur de rAUDPC élevée reflète un génotype sensible et inversement. La variation de ce caractère est plus étendue sur les deux sites du Finistère que sur les autres sites, et en 2018 par rapport à 2017. Ces courbes présentent une distribution plus ou moins bimodale selon les sites, résultant de la ségrégation des deux types de résistance au sein de cette collection. La résistance spécifique avec des rAUDPC très faibles, proches ou égales à 0 est illustrée par le premier pic de la distribution. Le deuxième pic de distribution correspond à la ségrégation de la résistance quantitative avec les individus sensibles les plus à droite. La distribution des rAUDPC sur le site du Pas-de-Calais en 2018 est décalée vers des valeurs faibles indiquant que cet essai est peu discriminant. Le positionnement des variétés témoins dans cet essai le montre également. Ainsi cet essai n'a pas été pris en compte pour les analyses d'association.

La distribution de la fréquence de l'allèle mineur (MAF) des 4504 SNP ayant une MAF supérieure à 0.05 est illustrée figure 7. L'histogramme montre que cette collection de géniteurs améliorés présente un grand nombre de SNP avec des valeurs élevées de fréquence allélique de l'allèle mineur (> 0.1). Ce qui est souhaité pour mettre en œuvre une stratégie de génétique d'association.

Structuration de la population

Les résultats de l'analyse avec le logiciel Structure® (Fig. 8) montrent que la structuration la plus probable est un nombre de sous populations égal à deux, avec 94 individus dans la sous-population 1 et 179 individus dans la sous-population 2. Un individu est assigné à une sous-population lorsqu'il a au moins 50% d'appartenance à cette population.

L'analyse des programmes d'origine des géniteurs (Fig. 9) montre que la sous-population 1 comprend majoritairement (52%) des géniteurs issus du programme de résistance à *G. pallida* et la sous-population 2 majoritairement (59%) des géniteurs issus du programme de résistance au mildiou.

Analyse d'association

Analyse 1

Une première analyse d'association site par site a permis de détecter 78 SNP à partir des données 2017 et 12 SNP à partir des données 2018. Six de ces SNP sont détectés sur les deux ans. Le nombre de SNP associés de façon significative à la résistance est de 5 SNP sur le site du Pas-de-Calais, 6 SNP dans le Finistère Nord, 17 SNP dans le Finistère Nord (site INRA), 26 SNP dans le Lot et Garonne et 43 SNP sur le site de l'Aisne.

Le but étant d'identifier des résistances stables au mildiou nous avons sélectionné les SNP qui ont été détectés au moins sur deux sites et/ou sur les deux ans. Ensuite une analyse de variance

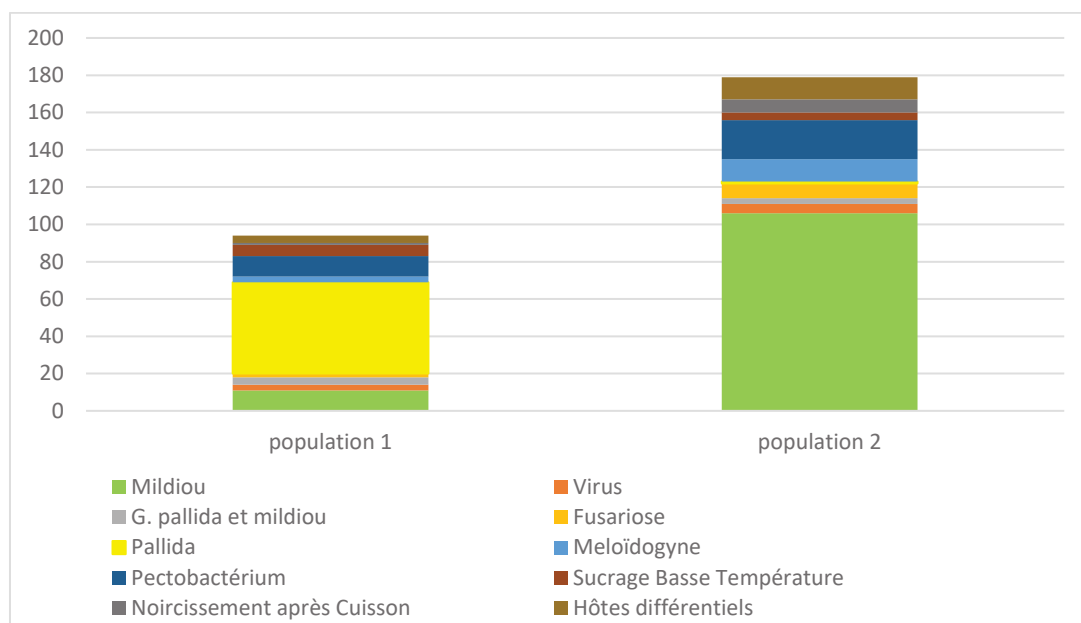


Figure 9 : Répartition des génotypes dans les deux sous-populations en fonction de leur programme d'origine.

Tableau 1: Description des SNP détectés le plus souvent dans l'analyse d'association 1. Les colonnes de la gauche vers la droite sont : les noms des SNP, le numéro de chromosome, la position physique, les deux allèles du SNP, le modèle avec lequel le SNP a été détecté dans l'analyse, le nombre de fois où le SNP a été détecté en fonction de l'année, le pourcentage d'explication de la variabilité expliqué par ce SNP et le pourcentage d'explication de la variabilité avec une analyse de variance des données de BLUP.

SNP	chr	Position (Mpb)	allèle	modèle	nombre de détection		R ² (analyse de variance par SNP)	R ² BLUP
					2017	2018		
solcap_snp_c2_27763	9	61.5	A/T	additif	1	1	22-23%	
solcap_snp_c1_8560	9	61.1	G/A	simplex dominant	4	4	24-50%	23-43%
solcap_snp_c2_56418	9	60.2	T/G	additif	2	2	19-27%	23%
solcap_snp_c2_23049	5	4.8	A/G	additif	2		10-18%	
solcap_snp_c2_18945	8	43.4	G/A	duplex dominant	2		9-15%	
solcap_snp_c2_41467	8	35.9	C/T	duplex dominant	1	1	7%	
solcap_snp_c2_36060	4	59.1	C/T	duplex dominant	2		6%	
solcap_snp_c1_10762	4	58.9	A/C	duplex dominant	1	1	5-8%	

a été réalisée pour tester l'effet de ces SNP sur la variation du caractère de résistance. Les SNP ne présentant pas d'effet significatif dans cette deuxième analyse ont été éliminés. Ainsi huit SNP ont été retenus (Tableau 1 et Fig.10) : trois sur le chromosome 9, deux sur le chromosome 8, deux sur le chromosome 4 et un sur le chromosome 5. Les SNP détectés le plus souvent sont présents sur le chromosome 9 (Tableau 1). Le site du Pas-de-Calais présente une particularité car c'est le seul site où aucun des SNP du chromosome 9 ne présente d'effet significatif (Fig. 10). Les analyses faites avec les BLUPs nous permettent de détecter les SNP qui ressortent sur le plus de sites : solcap_snp_c1_8560 et solcap_snp_c2_56418.

Le SNP détecté le plus souvent (solcap_snp_c1_8560) est un SNP du chromosome 9 (tableau 1 et Fig. 11). Il est détecté sur quatre des cinq sites et sur les deux ans. Ce premier marqueur a été mis en évidence dans un modèle de type simplex dominant (Fig. 12 a) sur tous les sites où il a été détecté (Annexe II). Ce qui signifie que la présence d'une seule copie de l'allèle favorable (allèle A) suffit pour avoir un phénotype moins sensible au mildiou. Toutefois dans la collection étudiée, il n'y a pas de génotypes homozygotes pour l'allèle favorable et il n'y a qu'un seul individu avec trois copies de cet allèle, classe génotypique qui n'est donc pas bien représentée ici. L'analyse de variance montre que ce SNP permet d'expliquer entre 24 et 50% de la variabilité des rAUDPC selon les sites. Les valeurs moyennes des rAUDPC par classe génotypique sont présentées dans le tableau 2. Les classes les plus favorables (avec des rAUDPC plus faibles) sont les classes AABB et ABBB. La fréquence allélique de l'allèle A de ce SNP est de 0.09 dans la population. Cette fréquence est de 0.05 dans la sous-population 1 et de 0.11 dans la sous-population 2.

Le deuxième SNP est le solcap_snp_c2_56418. Celui-ci est ressorti sur trois sites (Finistère Nord, Finistère Nord (INRA) et Lot-et-Garonne) et au cours des deux ans pour un site (Finistère Nord). Il a été mis en évidence avec un modèle de type additif (Fig.12 b et Annexe I : a). Le niveau de résistance augmente avec le nombre d'allèles B présents. L'analyse de variance montre que ce SNP permet d'expliquer entre 19 et 27% de la variabilité des rAUDPC selon les sites (Tableau 1). La fréquence allélique de l'allèle B de ce SNP est de 0.15 dans la collection. Elle est de 0.12 dans la sous-population 1 et de 0.17 dans la sous-population 2.

Le troisième SNP du chromosome 9, solcap_snp_C2_27763, associé de façon significative à la résistance est détecté en 2017 sur le site du Finistère Nord et en 2018 sur le site du Lot-et-Garonne. Il a été détecté avec un modèle de type additif (Annexe I : b et Annexe III : a) et permet d'expliquer entre 22 et 23% de la variabilité phénotypique sur les sites où il est détecté (Tableau 1). Les classes génotypiques les plus favorables sont les classes AAAB, AABB et ABBB. Ce SNP est cartographié dans un gène de sucrose synthase (PGSC0003DMG400006672). La fréquence allélique de l'allèle B de ce SNP est de 0.21 dans la collection. La fréquence dans les sous-populations 1 et 2 sont de 0.15 et 0.24 respectivement.

Sur le chromosome 5, le SNP solcap_snp_c2_23049 est détecté sur les sites du Pas-de-Calais et de Lot-et-Garonne en 2017. Il a été détecté avec le modèle additif (Tableau 1 et Annexe I : c). Ce SNP explique entre 10 et 18% de la variabilité phénotypique. Les classes génotypiques les plus favorables sont ABBB et BBBB (Annexe III : b). La position physique de ce SNP sur le génome de référence montre qu'il se situe dans un gène qui code pour une protéine de résistance au mildiou (PGSC0003DMG400018441). La fréquence allélique de l'allèle B de ce SNP est de 0.61 dans la collection. La fréquence dans les sous-populations 1 et 2 sont de 0.54 et 0.64 respectivement.

Sur le chromosome 8, deux SNP sont détectés comme associés de façon significative à la résistance (Tableau 1). Il y a le SNP solcap_snp_c2_18945 qui est détecté en 2017 (Annexe I : d) sur les sites de l'Aisne et du Finistère Nord (INRA). Il explique entre 9 et 15% de la

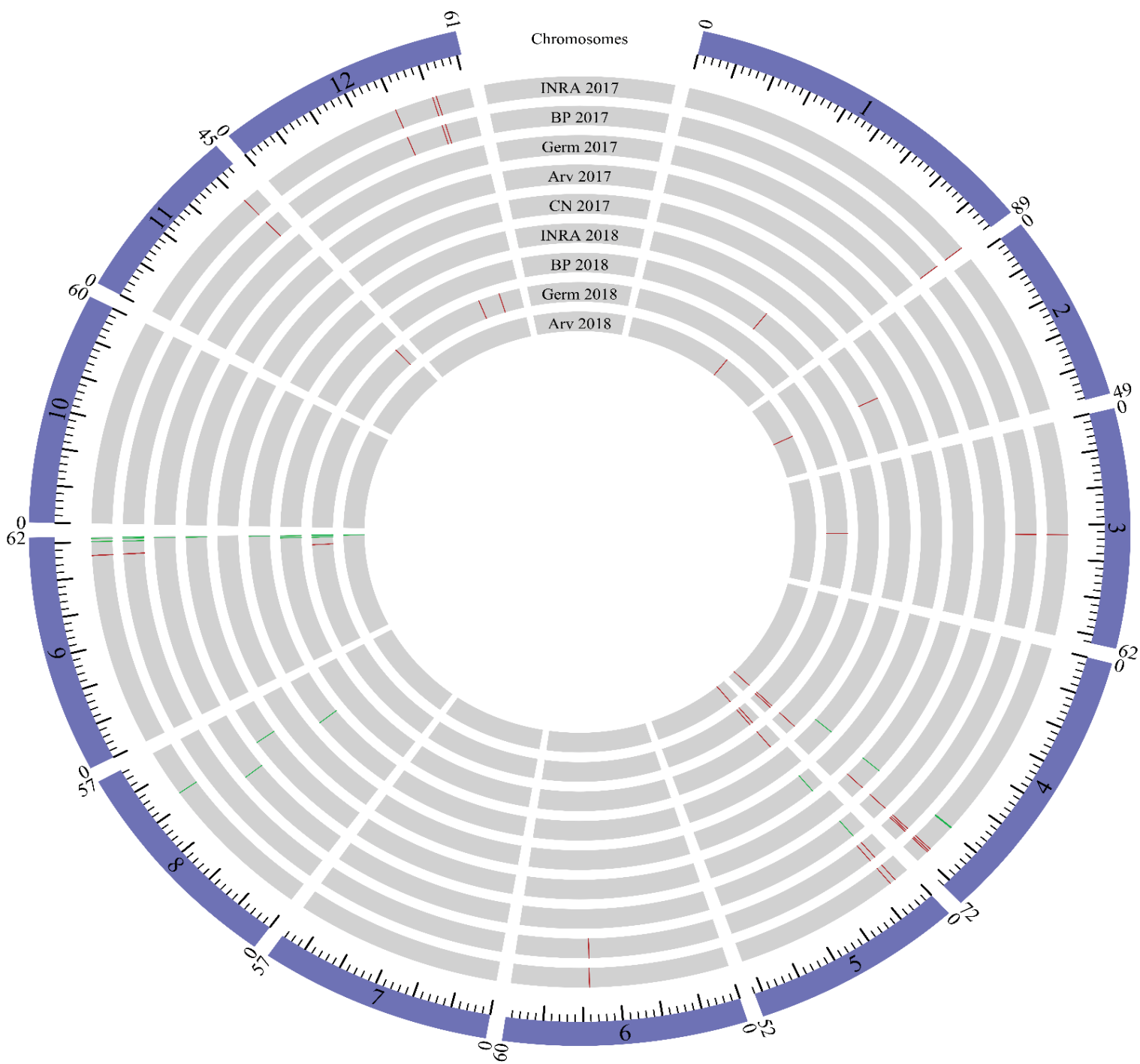


Figure 10 : Représentation de la position des SNPs sur les 12 chromosomes du génome. Le cercle le plus externe représente les chromosomes. Les SNPs sont représentés par des barres de couleur : en vert ceux de la 1^{ère} analyse et en rouge ceux de la 2^{ème}. Les cercles gris représentent les sites et les différentes années. (INRA : Finistère Nord (INRA), BP : Finistère Nord, Germ : Lot-et-Garonne, Arv : Aisne et CN : Pas-de-Calais)

variabilité. La classe la plus résistante est ABBB (Annexe III : c). Le deuxième SNP détecté sur le chromosome 8 est le solcap_snp_c2_41467 (Annexe I : e). Il est détecté en 2017 sur le site du Lot-et-Garonne et en 2018 sur le site du Finistère Nord (INRA). Le pourcentage d'explication pour ce SNP est de 7% (Tableau 1). Les classes les plus favorables pour ce SNP sont : AABB et ABBB (Annexe III : d).

Sur le chromosome 4, deux SNP sont détectés (Tableau 1). Les analyses de variance par site pour ces SNP montrent les pourcentages d'explication les plus faibles (Tableau 1). Le SNP solcap_snp_c2_36060 a un pourcentage d'explication de 6%. Il n'a été détecté qu'en 2017 sur deux sites qui sont : l'Aisne et le Finistère Nord (INRA) (Annexe I : f). Les classes les plus favorables sont : AABB, ABBB et BBBB (Annexe III : e). La fréquence de l'allèle B est de 0.39 dans la population et de 0.28 et 0.45 dans la sous-population 1 et 2 respectivement. Ce SNP est présent dans un gène qui code pour un précurseur de l'extensine (PGSC0003DMG400000783). Le deuxième SNP est le solcap_snp_c1_10762 (Annexe I : e). Il a été détecté sur le site du Finistère Nord sur les deux ans. La classe AAAA est la classe la plus favorable (Annexe III : f). La fréquence de l'allèle A dans la population est de 0.52 et 0.49 et 0.54 dans la sous-population 1 et 2 respectivement. Sur le génome de référence ce SNP se trouve dans un gène qui code pour une protéine sensible à la photopériode (PGSC0003DMG400000791).

Cette première analyse d'association montre qu'il y a un QTL à effet fort sur le chromosome 9. Nous nous sommes alors demandé si ce QTL à effet fort pouvait masquer d'autres QTL à effets plus faibles situés dans d'autres régions du génome. Les géniteurs améliorés qui présentent les classes génotypiques favorables pour les deux SNP à effets les plus forts : solcap_snp_c1_8560 (classe : ABBB) et solcap_snp_c2_56418 (classes : AABB, ABBB et BBBB) ont été identifiés afin de faire une analyse sans ces génotypes. Ce qui représente 93 individus dont 22 qui cumulent les deux combinaisons alléliques favorables. Parmi ces génotypes 69% sont issus des programmes de recherche de résistance au mildiou. Les autres génotypes proviennent de six autres programmes de recherche (*Pectobacterium*, *G. pallida*, fusariose, noircissement après cuisson, *Meloïdogyne* et Virus) dont 10% du programme *Pectobacterium*.

Analyse 2

Cette analyse a détecté plus de SNP, au total 284 SNP différents Il y a 96 SNPs détectés avec les données 2017 et 209 avec les données de 2018. Le nombre de SNP associés à la résistance sont de 188 dans l'Aisne, 29 dans le Finistère Nord, 26 dans le Finistère Nord (INRA), 0 dans le Pas-de-Calais et 89 en Lot-et-Garonne.

Les critères de sélection utilisés lors de l'analyse 1 ont aussi été appliqués aux résultats de cette analyse. Suite à cette sélection 17 SNPs différents ont été retenus, cartographiés sur 8 des 12 chromosomes (Fig.10). Les résultats présentés dans le tableau 3 montrent que tous les SNPs ont été détectés avec un modèle de type général sauf deux. Deux types d'analyses de variance ont été faites sur ces SNPs : une qui ne prend en compte que le sous-ensemble de génotypes et une deuxième qui prend en compte l'ensemble des génotypes.

Le SNP détecté le plus souvent dans cette analyse est sur le chromosome 4, c'est le SNP : solcap_snp_c2_32099 (Tableau 3). Il est détecté sur quatre sites, trois (Finistère Nord, Lot-et-Garonne et l'Aisne) sur les deux années et un juste en 2017 (Finistère Nord (INRA)). Il explique entre 6 et 23% de la variabilité avec le sous-ensemble de la collection et reste significatif avec l'ensemble de la collection mais avec un pourcentage explicatif plus faible (4-7%). La classe génotypique la plus favorable est la classe AAAB. Il n'y a qu'un seul individu dans la classe AAAA.

Deux autres SNPs sont détectés sur le chromosome 4. Le SNP solcap_snp_c2_56837 est détecté

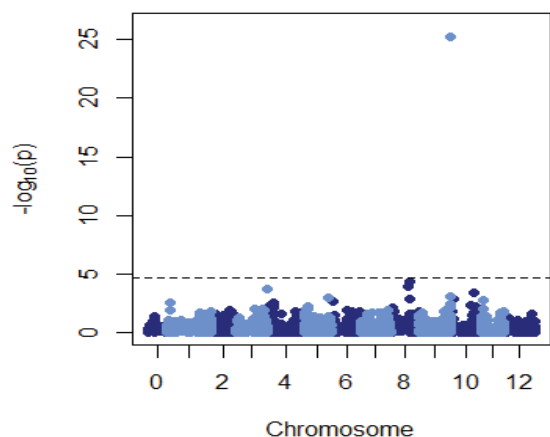


Figure 11 : Manhattan plot obtenu avec le modèle simple dominant (site du Finistère Nord (INRA), 2017).

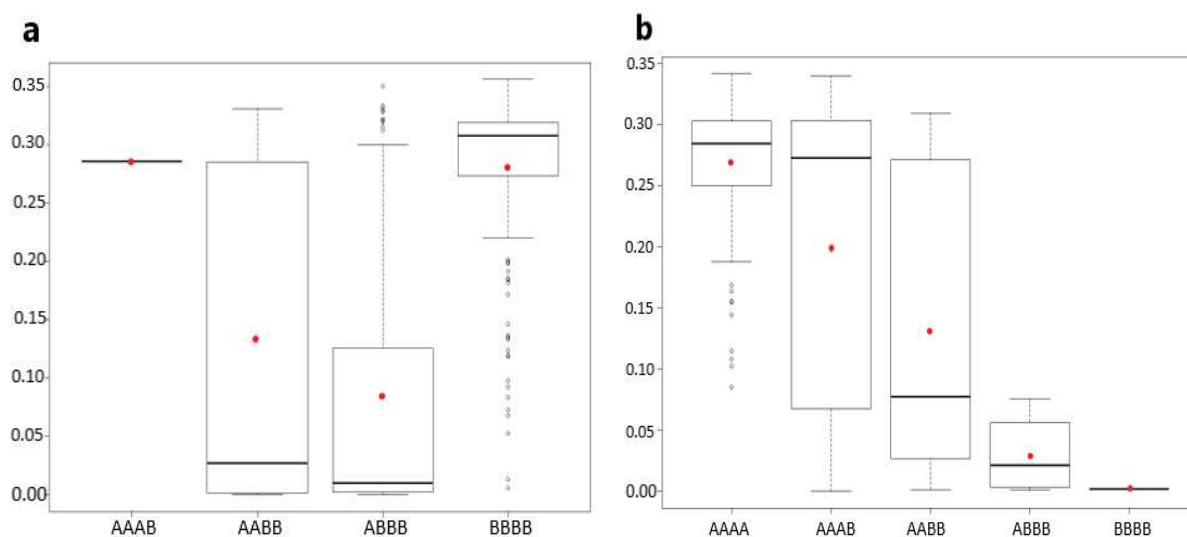


Figure 12 : Distribution des valeurs de $rAUDPC$ en fonction des classes génotypiques. Le point rouge représente la moyenne. La figure a pour le SNP *solcap_snp_c1_8560* (site : Finistère Nord (site INRA) 2017). La figure b pour le SNP *solcap_snp_c2_56418* (site Finistère Nord 2017).

Tableau 2: Présentation des valeurs de $rAUDPC$ moyenne des différentes classes génotypiques pour le SNP *c1_8560* du chromosome 9 sur les sites et les années où il a été détecté.

	site	année	classe génotypique			
			AAAB	AABB	ABBB	BBBB
Effectif			1	7	79	186
rAUDPC moyenne	Aisne	2017	0.20	0,12	0,1	0,17
		2018	0.20	0,05	0,04	0,11
	Finistère Nord	2017	0.30	0,15	0,11	0,27
		2018	0.28	0,1	0,06	0,2
	Lot-et-Garonne	2017	0.15	0,11	0,08	0,14
		2018	0.33	0,14	0,07	0,29
	INRA Ploudaniel	2017	0.29	0,13	0,09	0,28
		2018	0.36	0,14	0,1	0,26

sur les deux sites du Finistère Nord en 2017 et un autre site en 2018 (Lot-et-Garonne). La classe génotypique la plus favorable est BBBB. Le deuxième SNP est le solcap_snp_c2_34867 détecté sur deux sites (Finistère Nord (INRA) et Finistère Nord) en 2017. La classe favorable pour ce SNP est AAAA. Ce SNP se situe dans un gène qui code pour une protéine kinase PBS1 (PGSC0003DMG402007955).

Sur le chromosome 9, deux autres SNP sont détectés (Tableau 3). Ces deux SNP conservent des effets significatifs sur la variabilité phénotypique dans la collection totale. Le SNP solcap_snp_c2_30008 est détecté sur deux sites en 2017 (les deux sites du Finistère Nord) et sur un site en 2018 (Lot-et-Garonne). La classe génotypique favorable est AAAA. Ce SNP se trouve dans un gène qui code pour une sous-unité d'une flavoprotéine (PGSC0003DMG400017180). Le deuxième SNP détecté, solcap_snp_c2_29919, est détecté sur les deux sites du Finistère Nord en 2017. La classe génotypique la plus favorable est AAAA. Ce SNP se situe dans un gène qui code pour une hydroxylase (desacetoxyvindoline 4-hydroxylase, PGSC0003DMG403017241).

Trois SNP sont détectés sur le chromosome 12 (Tableau 3). Le SNP solcap_snp_c2_18827 est détecté sur les deux sites du Finistère Nord en 2017. La classe génotypique favorable est AAAA. Ce SNP se situe dans un gène qui code pour une multidrug resistance-associated protein 1,3 (PGSC0003DMG400033641). Le SNP solcap_snp_c2_43895 est détecté sur les deux sites du Finistère Nord en 2017 et sur un site (Lot-et-Garonne) en 2018. La classe génotypique favorable est AAAA. Il se situe dans un gène qui code pour une protéine réagissant à la déshydratation (PGSC0003DMG400026461). Le dernier SNP, le solcap_snp_c2_45803, est détecté sur les mêmes sites que le solcap_snp_c2_43895. Ce SNP conserve un effet significatif lorsqu'il est testé sur toute la collection. La classe génotypique favorable est AAAA. Ce dernier SNP se situe dans un gène qui code pour une protéine LeArcA1 (PGSC0003DMG400015829).

Sur le chromosome 5, trois SNP ont été détectés (Tableau 3). Il n'y a qu'un seul de ces trois SNP qui conserve un effet significatif dans la collection totale, le solcap_snp_c1_14836 qui est signalé sur deux sites (Finistère Nord et l'Aisne) en 2018. Pour ce SNP les classes favorables sont AABB et AAAB. Il n'y a qu'un seul individu dans la classe AAAA. Ce SNP se trouve dans un gène qui code pour une protéine de reconnaissance à un ARN (PGSC0003DMG400025609). Le SNP solcap_snp_c2_47597 est détecté sur les deux sites du Finistère Nord en 2017 et sur un site (Lot-et-Garonne) en 2018. La classe génotypique la plus favorable est AAAA. Il se trouve dans un gène qui code pour une protéine kinase (PGSC0003DMG400017625). Le troisième SNP détecté sur ce chromosome est le solcap_snp_c2_22986. Il est détecté sur les mêmes sites que le solcap_snp_c2_47597. Les classes favorables sont les classes AABB et ABBB. Il n'y a qu'un seul individu dans la classe AAAB et aucun dans la classe AAAA. Il se situe dans un gène qui code pour une hélicase à ADN (PGSC0003DMG402030493).

Deux SNP sont détectés sur le chromosome 1. Le SNP solcap_snp_c2_30956 est le seul SNP à garder un effet significatif sur la variabilité lorsqu'il est testé sur toute la collection (tableau 3). Ce SNP est détecté sur les deux sites du Finistère Nord en 2017. La classe génotypique la plus favorable est AABB. Il n'y a aucun individu appartenant aux classes génotypiques AAAA et AAAB pour ce SNP. Le deuxième SNP, le solcap_snp_c2_2591 est détecté sur deux sites (l'Aisne et le Finistère Nord) en 2018. Les classes génotypiques les plus favorables sont ABBB et BBBB. Ce SNP est dans un gène qui code pour une protéine de liaison à l'ADN (PGSC0003DMG400022651).

Le SNP solcap_snp_c2_22184, détecté sur le chromosome 11 conserve un effet significatif dans la collection globale. Il explique environ 10% de cette variabilité. Il est détecté sur les deux

Tableau 3 : Description des SNP détectés le plus souvent dans l'analyse d'association 2. Les colonnes de la gauche vers la droite sont : les noms des SNP, le numéro de chromosome, la position physique, les deux allèles du SNP, le modèle avec lequel le SNP a été détecté dans l'analyse, le nombre de fois où le SNP a été détecté en fonction de l'année, le pourcentage d'explication du SNP avec le sous ensemble d'individus et le pourcentage d'explication avec l'analyse de variance sur toute la collection.

SNP	chr	modèle	Position (Mpb)	Nombre de détection		R ²	R ² avec tous les individus
				2017	2018		
solcap_snp_c2_56837	4	général	68.5	2	1	13-17%	
solcap_snp_c2_34867	4	général	68	2		11-16%	
solcap_snp_c2_32099	4	général	67.3	4	3	6-23%	4-7 %
solcap_snp_c2_22184	11	général	40.8	2	1	14-22%	5-8%
solcap_snp_c2_18827	12	général	51.8	2		16-18%	
solcap_snp_c2_43895	12	général	50.8	2	1	11-18%	
solcap_snp_c2_45803	12	général	40.1	2	1	9-16%	4-5%
solcap_snp_c2_30008	9	général	56.2	2	1	11-18%	9-10%
solcap_snp_c2_29919	9	général	56.1	2		13-17%	8-9%
solcap_snp_c1_14836	5	général	5.7		2	10-11%	11-12%
solcap_snp_c2_47597	5	général	5.9	2	1	9-13%	NS
solcap_snp_c2_22986	5	général	4.3	2	1	11-19%	NS
solcap_snp_c2_21717	2	simplex dominant	25.4		2	15-25%	9-12%
solcap_snp_c2_41055	3	général	32.9	2	1	10-15%	6%
solcap_snp_c2_51756	6	général	38.2	2		11-12%	NS
solcap_snp_c2_2591	1	général	59		2	5-9%	NS
solcap_snp_c2_30956	1	duplex dominant	88.5	2		9-12%	9-10%

sites du Finistère Nord en 2017 et sur un site (Lot-et-Garonne) en 2018. La classe génotypique la plus résistante est ABBB. Il n'y a aucun individu dans la classe BBBB. Sur les chromosomes 2 et 3, les SNP détectés (Tableau 3) conservent aussi leur effet significatif dans la collection totale. Ce qui n'est pas observé pour le SNP détecté sur le chromosome 6. Le SNP du chromosome 2 est détecté sur deux sites (l'Aisne et le Finistère Nord (INRA)) en 2018. La classe génotypique la plus favorable est AAAA. Sur le chromosome 3, le SNP `solcap_snp_c2_41055` est détecté sur les deux sites du Finistère Nord en 2017 et sur un site (Lot-et-Garonne) en 2018. La classe la plus favorable est BBBB. Ce SNP se situe dans un gène qui code pour une ubiquitine (PGSC0003DMG400021791). Le SNP, `solcap_snp_c2_51756`, du chromosome 6 est détecté sur les deux sites du Finistère Nord en 2017 et sur un site (Lot-et-garonne) en 2018. La classe génotypique favorable est ABBB. Il n'y a pas de génotypes dans les classes AAAA et AAAB. Ce SNP est dans un gène qui code pour une protéine de la famille des pyruvate kinase (PGSC0003DMG400019605).

Déséquilibre de liaison

Sur le chromosome 9, en cumulant les deux analyses d'association, cinq SNP différents ont été détectés. La figure 13, nous montre qu'il existe deux régions dans lesquelles les SNP sont en déséquilibre de liaison. Les SNP issus de l'analyse 1 se trouvent dans une région et ceux de l'analyse 2 dans la deuxième région. Les trois SNP de l'analyse 1 (`solcap_snp_c2_27763`, `solcap_snp_c1_8560` et `solcap_snp_c2_56418`) sont significativement corrélés, la corrélation la plus forte étant entre les SNP `c1_8560` et `c2_56418` (Fig.13). Les deux autres SNP détectés lors de la deuxième analyse (`solcap_snp_c2_29919` et `solcap_snp_c2_30008`) présentent un coefficient de corrélation significatif et très élevé, 0.81. Ce qui est en accord avec leur position, ces deux SNP sont très proches sur le génome (0.084 Mpb de distance, Fig.10). Ces deux SNP montrent une corrélation significative avec un des SNP de l'analyse 1, le `c2_56418` mais avec des valeurs de coefficients assez faibles. Cependant le test n'est pas significatif avec les deux autres SNP de l'analyse 1.

Sur le chromosome 4, en cumulant les deux analyses d'association, cinq SNP différents ont été détectés. Sur la figure 14, il est possible d'observer deux zones dans lesquelles les corrélations entre paires de SNP sont significatives. La zone la plus haute à gauche sur le graphique 14 contient les SNP de la première analyse. Ces deux SNP (`c2_36060` et `c1_10762`) ont un coefficient de corrélation négatif élevé, -0.66, ce qui est cohérent avec leur position car ils sont assez proches (0.22 Mpb, Fig.10). Les trois SNP de l'analyse 2 qui se situent dans la deuxième zone, en bas à droite sur le graphique, ont des coefficients de corrélation significatifs. Le coefficient de corrélation le plus élevé est entre les SNP `c2_32099` et `c2_34867` avec 0.65 qui sont séparés par 0.64 Mpb.

Sur le chromosome 5, en cumulant les deux analyses, quatre SNP sont détectés. La figure 15, montre que deux des trois marqueurs détectés dans l'analyse 2 sont en déséquilibre de liaison (le SNP `c2_22986` et le `c2_47597`). Le marqueur détecté dans la première analyse (`c2_23049`) est en déséquilibre avec deux des marqueurs de l'analyse 2 : `c1_14836` et `c2_47597`. Ces 4 SNP sont situés dans la même région du chromosome (Fig.10).

Sur le chromosome 12, les trois SNP détectés ont des coefficients de corrélation significatifs. Le SNP `c2_43895` est plus éloigné des deux autres (distant de 10.6 Mpb, Fig.10) mais est très corrélé avec eux : le coefficient est de 0.69 avec le `c2_18827` et de 0.67 avec le `c2_45803`.

Sur le chromosome 8, les deux SNP détectés lors de la première analyse ne présentent pas un coefficient de corrélation significatif. Ils ne sont donc pas en déséquilibre de liaison ce qui est cohérent avec la distance qui les sépare (7.5 Mpb, Fig.10). Sur le chromosome 1, les deux SNP



Figure 13 : Matrice de corrélation d'une sélection de SNP du chromosome 9 (en vert SNP de l'analyse 1 et en rouge SNP de l'analyse 2). Les points bleus représentent les valeurs significatives (<0.01), les chiffres sont les coefficients de corrélations. A gauche, les distances entre les SNPs sont indiquées.



Figure 14 : Matrice de corrélation d'une sélection de SNP du chromosome 4 (en vert SNP de l'analyse 1 et en rouge SNP de l'analyse 2). Les points bleus représentent les valeurs significatives (<0.01), les chiffres sont les coefficients de corrélations. Les distances entre les SNPs sont indiquées dans la colonne de gauche.

de la deuxième analyse sont physiquement très éloignés (29.5 Mpb, Fig.10). Ils ne sont pas non plus en déséquilibre de liaison, leur coefficient de corrélation n'est pas significatif.

Discussion

L'analyse de structuration de la population de géniteurs améliorés indique une structuration en 2 sous-populations. L'analyse des programmes d'origine de ces géniteurs montre que cette structuration en deux sous-populations peut avoir une explication biologique (Fig.9). En effet chacune de ces sous-populations contient principalement des individus issus de deux programmes différents : *G. pallida* et mildiou. Ce sont les deux programmes les plus représentés dans la population toutefois leur proportion dans chacune des sous-populations est bien distincte. Cette structuration selon le programme d'origine s'explique par l'utilisation de parents génétiquement différents au sein de ces programmes.

Cependant, cette analyse de structuration montre également que de nombreux individus ont des fonds génétiques mélangés (Fig.8). Le pourcentage d'appartenance à une des deux sous-populations n'est pas toujours très supérieur à 50%. Ce qui montre que la structuration de la population de géniteurs améliorés est plutôt faible. Ces résultats sont en accord avec les résultats d'études précédentes qui mettent en évidence l'absence de structuration parmi les clones de variétés de pomme de terre européennes (Stich et al. 2013). Ces données de structuration ont toutefois été intégrées dans l'analyse d'association.

La distribution des rAUDPC (fig. 5 et 6) montre qu'il existe une variabilité selon les sites. En 2018, sur le site du Pas-de-Calais, les conditions climatiques chaudes et sèches ont été peu favorables au développement de la maladie. Ce qui explique une distribution des valeurs de rAUDPC vers des valeurs faibles, y compris pour des témoins sensibles (Fig.8). Ainsi, les données obtenues dans cet essai sont très peu discriminantes. Elles n'ont donc pas été utilisées pour l'analyse d'association.

La première série d'analyses d'association faite avec les données de 2018 détecte moins de SNP significatifs qu'avec les données 2017. Globalement, dans les essais 2018, le mildiou a progressé très vite en début d'épidémie. Puis celle-ci a stagné rapidement du fait de conditions climatiques peu favorables. Ainsi, la résistance partielle n'a pas pu s'exprimer pleinement (représentation plus faible des valeurs intermédiaires de rAUDPC) (Fig. 5 et Fig. 6). Ceci peut donc expliquer la plus faible détection de SNP sur cette année-là.

Le nombre de SNP détectés varie selon les sites (Fig. 10). Les sites de l'Aisne et du Lot-et-Garonne sont les sites dans lesquels le plus de SNP ont été détectés. La majorité de ces SNP ont été identifiés avec les données de 2017. La gamme de distribution des rAUDPC en 2017 pour ces sites montre qu'il y a une plus forte représentation des valeurs intermédiaires de rAUDPC par rapport aux autres sites (Fig. 5). Sur ces deux sites, une inoculation artificielle a dû être réalisée car les conditions météorologiques n'étaient pas favorables au démarrage de l'épidémie et ensuite à la progression de la maladie. Ainsi, tous les SNP détectés sur ces sites ne sont probablement pas tous associés à la résistance au mildiou.

Le site du Pas-de-Calais est le site où le moins de SNP sont détectés (Fig. 10). Cet essai a également dû être inoculé de façon artificielle. L'épidémie a démarré tardivement et les notations ont été rapidement rendues compliquées par la sénescence des plantes et l'apparition d'une autre maladie fongique (l'alternaria). Ainsi l'épidémie est moins discriminante, ce qui peut expliquer la faible détection de SNP et qu'aucun SNP associé au QTL à effet fort du chromosome 9 ne soit détecté.



Figure 15 : Matrice de corrélation d'une sélection de SNP du chromosome 5 (en vert SNP de l'analyse 1 et en rouge SNP de l'analyse 2). Les points bleus représentent les qvalues significatives (<0.01), les chiffres sont les coefficients de corrélations. Les distances entre les SNP sont indiquées dans la colonne de gauche.

Les données du site du Finistère nord, correspondant à l'essai avec une couverture fongique partielle, ne permettent pas de détecter un plus grand nombre de SNP ou des SNP différents de l'autre site du Finistère Nord (essai INRA). Un des objectifs de ce dispositif était de regarder s'il était possible d'observer l'expression de résistances partielles qui s'exprimeraient avec des pressions de mildiou plus faibles.

La deuxième série d'analyses d'association a été faite avec moins d'individus, il est donc plus difficile d'identifier le modèle génétique permettant d'expliquer l'effet des SNP détectés car certaines classes génotypiques sont moins représentées. Ainsi, les SNP identifiés lors de cette analyse ont été principalement détectés avec le modèle général. Cette analyse permet d'identifier des SNP différents de ceux de la première série d'analyses. Certains SNP sont identifiés sur les mêmes chromosomes (4, 5, 9), d'autres sur d'autres chromosomes (1, 3, 6, 11 et 12), (Tableaux 1 et 3). Lorsque que l'effet de ces SNP est analysé sur l'ensemble de la collection, cet effet est plus faible et même parfois n'est plus significatif (Tableau 3). Nous pouvons penser que les QTL associés à ces SNP ont des effets faibles, masqués par l'effet des QTL à effet plus fort du chromosome 9.

D'autre part, l'analyse génétique des SNP identifiés montre que certaines classes génotypiques sont peu ou pas représentées, comme par exemple les classes AAAA et AAAB pour le SNP `solcap_snp_c1_8560` (Tableau 2), même si on prend en compte l'ensemble de la collection de géniteurs. Plusieurs hypothèses peuvent expliquer cette observation. Leurs phénotypes n'étaient peut-être pas favorables et ces individus n'ont pas été conservés dans la collection ou ces génotypes sont peut-être létaux. D'autre part, nous avons observé que les allèles favorables de ces SNP avaient une fréquence plus élevée dans la sous-population 2, ce qui est bien en accord avec le fait que cette sous-population comprend majoritairement des géniteurs issus du programme de résistance au mildiou.

Les SNP identifiés qui présentent le pouvoir explicatif le plus fort sont 3 SNP situés à l'extrémité du bras long du chromosome 9 (Tableau 1). Etant très corrélés entre eux, ils marquent probablement un même QTL. C'est ce QTL qui semble le plus stable parmi ceux détectés. Il est détecté sur le plus de sites et sur les deux années, donc dans des conditions épidémiologiques différentes. Deux de ces marqueurs (`c2_56418` et `c2_27763`) ont déjà été identifiés dans une autre analyse de la résistance au mildiou dans une population de clones tétraploïdes d'Amérique latine (Lindqvist-Kreuze et al. 2014). Pour le SNP `solcap_snp_c2-56418`, les auteurs indiquent que dans leur étude la résistance liée à ce marqueur n'est pas stable dans tous les environnements. Les classes hétérozygotes sont également les plus résistantes, ce qui est en accord avec nos résultats. Cette résistance, dans leur population, proviendrait d'un géniteur qui porte des gènes R non connus (Lindqvist-Kreuze et al. 2014). Dans la collection analysée ici, les génotypes porteurs de cette résistance sont issus majoritairement des programmes de résistance au mildiou (69%), ce qui était attendu.

Les SNP du chromosome 9 détectés lors de la deuxième analyse marquent probablement un autre QTL que celui identifié lors de la première analyse car ils sont situés entre 56.1 et 56.2 Mpb. Une étude du déséquilibre de liaison sur une collection de pomme de terre a montré que le déséquilibre de liaison diminuait de moitié à partir de 860 000pb dans cette région du chromosome 9 (Sharma et al. 2018). Ce qui confirme l'hypothèse que les deux analyses identifient deux QTL différents sur le chromosome 9. En effet ces deux régions sont séparées par 3 998 000 pb.

Le bras long du chromosome 9 porte deux clusters de gènes de résistance (Jo et al. 2011). Un de ces clusters, situé vers 59.6 Mpb, contient différentes copies d'un gène R provenant de l'espèce sauvage *S. venturii* : *Rpi-vnt1*. Le marqueur `solcap_snp_c2_56418` est en position plus

télomérique que le gène *Rpi-vnt1* mais sa position (60.2 Mpb) reste de proche de ce gène (Lindqvist-Kreuze et al. 2014). Les gènes de résistance *R8* et *R9a*, issus de *S. demissum* sont aussi cartographiés sur cette partie du chromosome 9 (Jo et al. 2011; 2015). Il y a donc plusieurs gènes situés dans cette extrémité du chromosome 9 qui peuvent contribuer à la résistance au mildiou observée dans la collection de géniteurs améliorés. Cette résistance serait indépendante de la maturité des plantes (Mosquera et al. 2016). Dans nos essais, les hôtes différentiels porteurs de gènes *R8* et *R9* ne présentent pas ou peu de symptômes. Ainsi, ce sont peut-être ces gènes qui s'expriment dans ces géniteurs. Toutefois, il est possible que ces hôtes différentiels contiennent d'autres gènes R non connus. D'autre part, l'étendue des valeurs de rAUDPC observée dans les classes favorables pour les marqueurs que nous avons identifiés peut refléter le fait que d'autres facteurs génétiques contribuent à cette résistance (Fig.12).

Les SNP détectés sur le chromosome 5 se situent sur le bras court, dans une région où Mosquera et al (2016) ont identifié des SNP associés à la résistance au mildiou. Le bras court du chromosome 5 est connu pour porter le gène *R1* et des QTL liés à des QTL de maturité (Leonards-Schippers et al. 1992; Visker et al. 2003; Mosquera et al. 2016). Toutefois le gène *R1* est contourné (Fry, 2008) et l'hôte différentiel portant *R1* s'est montré très sensible dans nos essais. Le SNP identifié dans l'analyse 1 se trouve dans un gène qui code pour une protéine de résistance au mildiou. Ainsi, la résistance associée aux marqueurs identifiés sur ce chromosome est probablement en partie due à ce gène.

Les SNP détectés sur le chromosome 4 sont présents sur le bras long du chromosome dans deux régions distantes de 8.2 Mpb et marquent donc probablement deux QTL différents (Fig. 10). Le chromosome 4 est porteur de plusieurs gènes R (*Rpi-blb3*, *R2*...) mais qui sont cartographiés sur le bras court du chromosome (Park et al. 2005). Mosquera et al (2016) ont identifié un QTL associé à la résistance au mildiou entre 71.0-72.5 Mpb sur le bras long. Le meilleur gène candidat que les auteurs ont identifié dans cette région est le gène HMGCR (Mosquera et al. 2016). Ce gène code pour une enzyme impliquée dans la voie du mévalonate qui fournit des précurseurs pour la synthèse des triterpènes et isoprénoides (Mosquera et al. 2016). Les SNP que nous avons identifiés dans l'analyse 2 sont distants de 3 Mpb de cette région, il semble donc peu probable que le QTL identifié soit le même. Par contre Draffehn et al (2013) ont identifié des SNP associés à la résistance au mildiou dans les deux régions que nous avons identifiées dans cette étude.

Les deux SNP du chromosome 8, distants de 7.5 Mpb, marquent deux régions différentes. Le bras long de ce chromosome porte un cluster de gènes R : *Rpi-blb1* issu de *S. bulbocastanum*, *Rpi-pt1* et *Rpi-st1* issus de *S. stoloniferum* (Wang et al. 2008; van der Vossen et al. 2003). Les résistances associées à ces SNP sont peut-être dues à la présence d'un ou plusieurs de ces gènes car ils se situent sur le même bras chromosomique.

Le SNP détecté sur le chromosome 11 se situe à l'extrémité du bras court du chromosome. Dans cette région un cluster de gènes R issus de *S. demissum* a été cartographié. Ce locus est constitué de plusieurs gènes dont *R3a*, *R3b*, *Rpi-Smira1*, *R10* et *R11* (El-Kharbotly et al. 1994; Bradshaw et al. 2006; Huang et al. 2004; Tomczyńska et al. 2014). La résistance identifiée dans la collection de géniteurs améliorés peut être conférée par un ou plusieurs de ces gènes. Les gènes *R3*, *R10* et *R11* sont contournés et nous avons bien observé des symptômes importants sur les hôtes différentiels porteurs des gènes *R3* et *R11*. Par contre l'hôte différentiel porteur du gène *R10* a été moins attaqué. De plus il a été rapporté que le comportement de ce gène R était proche de l'expression d'un QTL (Bradshaw et al. 2006).

Les deux SNP détectés sur le chromosome 1 ne sont pas en déséquilibre de liaison. Ce qui est cohérent avec la distance physique qui les sépare (29.5 Mpb, Fig. 10). Un de ces SNP, le SNP

solcap_snp_c2_30956, se situe dans une région dans laquelle Mosquera et al (2016) ont identifié un QTL de résistance au mildiou (88 – 89 Mpb). Un gène proche de cette région (PGSC0003DMG400006186, 87.6 Mpb) a aussi été détecté comme associé à la résistance dans une autre étude (Muktar et al. 2015). Ce gène code pour une hydroxyde pyruvate impliquée dans la photorespiration.

Ainsi, certains des SNPs identifiés dans cette étude semblent marquer des régions avec des gènes R. La plupart de ces gènes sont contournés mais ils peuvent conserver un effet résiduel et donc exprimer une résistance similaire à une résistance quantitative (Stewart, Bradshaw, Pande 2003). Il a aussi été observé que des QTL et des gènes R se trouvent dans les mêmes régions (Rauscher et al. 2010). Les facteurs génétiques identifiés dans cette étude sont complexes. Ces différentes résistances pourraient être intégrées dans les programmes de sélection afin de créer des variétés qui les cumulent. D'autres part, certains de ces SNP sont présents dans des gènes annotés sur la séquence de la pomme de terre. Une étude plus approfondie de ces gènes doit encore être réalisée afin d'évaluer leur rôle possible dans l'expression de la résistance.

Conclusion et perspectives

Cette étude a permis d'identifier des QTL associés à la résistance au mildiou au sein de la collection de géniteurs améliorés du centre de ressources biologiques de l'INRA.

Les facteurs génétiques identifiés dans cette étude sont complexes. Ces différentes résistances pourraient être intégrées dans les programmes de sélection afin de créer des variétés qui les cumulent. Ainsi, Les SNP identifiés pourront être utilisés en sélection assistée par marqueur (SAM). Il faut s'intéresser dans un premier temps aux SNP associés aux QTL les plus explicatifs. Le SNP solcap_snp_c2_8560, identifié sur le chromosome 9, est associé à la résistance la plus stable. Pour la mise en œuvre de la SAM, une première phase va consister à regarder si ces SNP peuvent bien être utilisés sur du matériel génétique différent. Le génotypage de ces marqueurs SNP pourra être réalisé avec la technologie KASPar basée sur l'utilisation de trois amorces : deux sont spécifiques de chacun des allèles du SNP et la troisième est commune. Après amplification, la lecture de la fluorescence peut être faite par exemple sur le Light Cycler 480 de Roche. Cette technique est adaptée pour génotyper quelques SNP sur un grand nombre d'individus, ce qui est réalisé en SAM.

La comparaison de la position des SNP détectés avec les facteurs de résistance identifiés dans la bibliographie est compliquée. La comparaison pourrait être plus facile si les marqueurs identifiés dans la bibliographie étaient positionnés sur la séquence du génome.

Les analyses réalisées n'ont pas pris en compte l'influence de la maturité. Il serait alors intéressant de réaliser une analyse d'association avec les notes de maturité des génotypes ou de calculer des MCR (maturity corrected resistance) qui sont des valeurs de rAUDPC corrigées selon la maturité. La comparaison des SNP détectés pourrait permettre de savoir si les QTL détectés pour la résistance au mildiou sont indépendants ou non de la maturité des plants.

Dans le cadre de ce projet les souches de mildiou présentes sur chaque site ont été analysées. Il pourra être intéressant d'étudier si certains SNP identifiés sont associés à une résistance vis-à-vis de certaines populations de mildiou.

Bibliographie

- ANDRIVON, PELLÉ et ELLISSÈCHE, 2006. Assessing resistance types and levels to epidemic diseases from the analysis of disease progress curves: Principles and application to potato late blight. *American Journal of Potato Research*. 2006. Vol. 83, n° 6, pp. 455-461. DOI 10.1007/BF02883506.
- BALDING, David J., 2006. A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*. 2006. Vol. 7, n° 10, pp. 781-791. DOI 10.1038/nrg1916.
- BRADSHAW, BRYAN, LEES, MCLEAN et SOLOMON-BLACKBURN, 2006. Mapping the R10 and R11 genes for resistance to late blight (*Phytophthora infestans*) present in the potato (*Solanum tuberosum*) R-gene differentials of Black. *Theoretical and Applied Genetics*. Vol. 112, n° 4, pp. 744-751. DOI 10.1007/s00122-005-0179-9.
- COLON, TURKENSTEEN, PRUMMEL, BUDDING et HOOGENDOORN, 1995. Durable resistance to late blight (*Phytophthora infestans*) in old potato cultivars. *European Journal of Plant Pathology*. Vol. 101, n° 4, pp. 387-397. DOI 10.1007/BF01874852.
- CORBIÈRE, CHATOT, MABON et MARIETTE, 2016. Mildiou en France: *P. infestans* sous surveillance. *Phytoma*. 2016. N° 690, pp. 30-36.
- DANAN, VEYRIERAS et LEFEBVRE, 2011. Construction of a potato consensus map and QTL meta-analysis offer new insights into the genetic architecture of late blight resistance and plant maturity traits. *BMC Plant Biology*. 2011. Vol. 11, n° 1, pp. 16. DOI 10.1186/1471-2229-11-16.
- DOUCHES, HIRSCH, MANRIQUE-CARPINTERO, MASSA, COOMBS, HARDIGAN, BISOGNIN, DE JONG et BUELL, 2014. The Contribution of the Solanaceae Coordinated Agricultural Project to Potato Breeding. *Potato Research*. Vol. 57, n° 3-4, pp. 215-224. DOI 10.1007/s11540-014-9267-z.
- DRAFFEHN, LI, KREZDORN, DING, LÜBECK, Jens, STRAHWALD, MUKTAR, WALKEMEIER, ROTTER et GEBHARDT, 2013. Comparative transcript profiling by SuperSAGE identifies novel candidate genes for controlling potato quantitative resistance to late blight not compromised by late maturity. *Frontiers in Plant Science*. Vol. 4. DOI 10.3389/fpls.2013.00423.
- EARL et VONHOLDT, 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*. Vol. 4, n° 2, pp. 359-361. DOI 10.1007/s12686-011-9548-7.
- EL-KHARBOTLY, LEONARDS-SCHIPPERS, HUIGEN, JACOBSEN, PEREIRA, A., STIEKEMA, SALAMINI et GEBHARDT, 1994. Segregation analysis and RFLP mapping of the R1 and R3 alleles conferring race-specific resistance to *Phytophthora infestans* in progeny of dihaploid potato parents. *MGG Molecular & General Genetics*. Vol. 242, n° 6. DOI 10.1007/BF00283432.
- ESNAULT, CHAUVEAU, BÉRARD, BOLAND, LE PASLIER, BRUNEL et CHAUVIN, 2012. *Diversity analysis of a potato (Solanum tuberosum L. subsp. tuberosum) core collection using the SolCAP chip*.
- EVANNO, REGNAUT et GOUDET, 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*. 2005. Vol. 14, n° 8, pp. 2611-2620. DOI 10.1111/j.1365-294X.2005.02553.x.
- FELCHER, COOMBS, MASSA, HANSEY, HAMILTON, John P., VEILLEUX, BUELL et DOUCHES, 2012. Integration of Two Diploid Potato Linkage Maps with the Potato Genome Sequence. *PLoS ONE*. Vol. 7, n° 4, pp. e36347. DOI 10.1371/journal.pone.0036347.

FORBES, G. A., CHACON, M. G., KIRK, H. G., HUARTE, M. A., VAN DAMME, M., DISTEL, S., MACKAY, G. R., STEWART, H. E., LOWE, R., DUNCAN, J. M., MAYTON, H. S., FRY, W. E., ANDRIVON, D., ELLISSECHE, D., PELLE, R., PLATT, H. W., MACKENZIE, G., TARN, T. R., COLON, L. T., BUDDING, D. J., LOZOYA-SALDANA, H., HERNANDEZ-VILCHIS, A. et CAPEZIO, S., 2005. Stability of resistance to *Phytophthora infestans* in potato: an international evaluation. *Plant Pathology*. Vol. 54, n° 3, pp. 364-372. DOI 10.1111/j.1365-3059.2005.01187.x.

FRY, 2008. *Phytophthora infestans*: the plant (and R gene) destroyer. *Molecular Plant Pathology*. 2008. Vol. 9, n° 3, pp. 385-402. DOI 10.1111/j.1364-3703.2007.00465.x.

HAMILTON, HANSEY, WHITTY, STOFFEL, MASSA, Alicia N, VAN DEYNZE, DE JONG, DOUCHES et BUELL, 2011. Single nucleotide polymorphism discovery in elite north american potato germplasm. *BMC Genomics*. 2011. Vol. 12, n° 1. DOI 10.1186/1471-2164-12-302.

HAVERKORT, BOONEKAMP, HUTTEN, JACOBSEN, LOTZ, L. A. P., KESSEL, VISSER et VAN DER VOSSSEN, 2008. Societal Costs of Late Blight in Potato and Prospects of Durable Resistance Through Cisgenic Modification. *Potato Research*. Vol. 51, n° 1, pp. 47-57. DOI 10.1007/s11540-008-9089-y.

HIRSCH, HIRSCH, FELCHER, COOMBS, ZARKA, Dan, VAN DEYNZE, DE JONG, VEILLEUX, JANSKY, BETHKE, DOUCHES et BUELL, 2013. Retrospective View of North American Potato (*Solanum tuberosum* L.) Breeding in the 20 th and 21 st Centuries. *G3 Genes|Genomes|Genetics*. Vol. 3, n° 6, pp. 1003-1013. DOI 10.1534/g3.113.005595.

HUANG, VLEESHOUWERS, WERIJ, HUTTEN, VAN ECK, VISSER et JACOBSEN, 2004. The R3 Resistance to *Phytophthora infestans* in Potato is Conferred by Two Closely Linked R Genes with distinct Specificities. *Molecular Plant-Microbe Interactions*. Vol. 17, n° 4, pp. 428-435. DOI 10.1094/MPMI.2004.17.4.428.

JEFFREY, JINKS et GRINDLE, 1962. Intraracial variation in *Phytophthora infestans* and field resistance to potato blight. *Genetica*. Vol. 32, n° 1, pp. 323-338. DOI 10.1007/BF01816101.

JO, ARENS, KIM, JONGSMA, VISSER, JACOBSEN et VOSSSEN, 2011. Mapping of the S. demissum late blight resistance gene R8 to a new locus on chromosome IX. *Theoretical and Applied Genetics*. Vol. 123, n° 8, pp. 1331-1340. DOI 10.1007/s00122-011-1670-0.

JO, VISSER, JACOBSEN et VOSSSEN, 2015. Characterisation of the late blight resistance in potato differential MaR9 reveals a qualitative resistance gene, R9a, residing in a cluster of Tm-2 2 homologs on chromosome IX. *Theoretical and Applied Genetics*. Vol. 128, n° 5, pp. 931-941. DOI 10.1007/s00122-015-2480-6.

LEONARDS-SCHIPPERS, GIEFFERS, SALAMINI et GEBHARDT, 1992. The R1 gene conferring race-specific resistance to *Phytophthora infestans* in potato is located on potato chromosome V. *MGG Molecular & General Genetics*. Vol. 233, n° 1-2, pp. 278-283. DOI 10.1007/BF00587589.

LINDQVIST-KREUZE, GASTELO, PEREZ, FORBES, DE KOEYER et BONIERBALE, 2014. Phenotypic Stability and Genome-Wide Association Study of Late Blight Resistance in Potato Genotypes Adapted to the Tropical Highlands. *Phytopathology*. Vol. 104, n° 6, pp. 624-633. DOI 10.1094/PHYTO-10-13-0270-R.

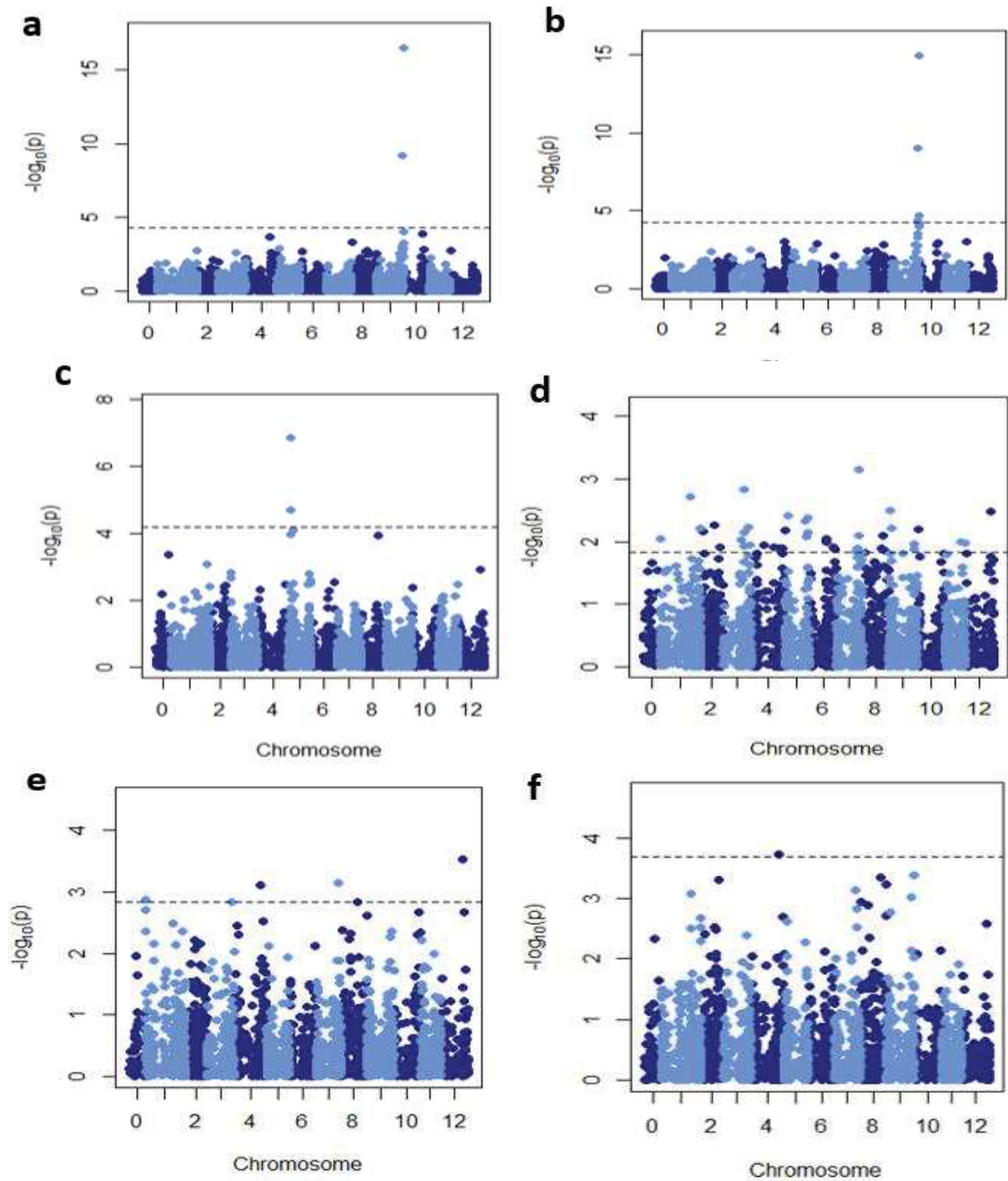
MARHADOUR, Sylvie, MÉAR, PELLÉ, ABIVEN, AUROUSSEAU, DUBREUIL, LE HINGRAT et CHAUVIN, 2012. Markers assisted selection for late blight resistance in tetraploid potato, experience about the use of the use of the potato genome sequence. In : *9th Solanaceae Conference, From bench to innovative applications*. Neuchâtel, Suisse. 26 août 2012.

- MARIETTE, MABON, CORBIÈRE, BOULARD, GLAIS, I., MARQUER, PASCO, MONTARRY et ANDRIVON, 2016. Phenotypic and genotypic changes in French populations of *Phytophthora infestans*: are invasive clones the most aggressive? *Plant Pathology*. Vol. 65, n° 4, pp. 577-586. DOI 10.1111/ppa.12441.
- MOSQUERA, ALVAREZ, JIMÉNEZ-GÓMEZ, MUKTAR, PAULO, Maria João, STEINEMANN, LI, DRAFFEHN, HOFMANN, LÜBECK, STRAHWALD, TACKE, HOFFERBERT, WALKEMEIER et GEBHARDT, 2016. Targeted and Untargeted Approaches Unravel Novel Candidate Genes and Diagnostic SNPs for Quantitative Resistance of the Potato (*Solanum tuberosum* L.) to *Phytophthora infestans* Causing the Late Blight Disease. *PLOS ONE*. 2016. Vol. 11, n° 6, pp. e0156254. DOI 10.1371/journal.pone.0156254.
- MUKTAR, LÜBECK, STRAHWALD et GEBHARDT, 2015. Selection and validation of potato candidate genes for maturity corrected resistance to *Phytophthora infestans* based on differential expression combined with SNP association and linkage mapping. *Frontiers in Genetics*. Vol. 6. DOI 10.3389/fgene.2015.00294.
- PARK, GROS, SIKKEMA, VLEESHOUWERS, MUSKENS, Marielle, ALLEFS, JACOBSEN, VISSER et VAN DER VOSSEN, 2005. The Late Blight Resistance Locus Rpi-blb3 from *Solanum bulbocastanum* Belongs to a Major Late Blight R Gene Cluster on Chromosome 4 of Potato. *Molecular Plant-Microbe Interactions*. Vol. 18, n° 7, pp. 722-729. DOI 10.1094/MPMI-18-0722.
- PELLÉ, BOZEC, ANDRIVON, SOYER et ELLISSÈCHE, 2005. Evaluation au champ de la résistance de la pomme de terre au mildiou du feuillage. In *Cahier des techniques*. 2016. N° spécial : Méthodes d'appréciation du comportement variétal vis-à-vis des bioagresseurs, p.73-78.
- PRITCHARD, STEPHENS et DONNELLY, 2000. Inference of Population Structure Using Multilocus Genotype Data. pp. 15.
- RAUSCHER, SIMKO, MAYTON, BONIERBALE, SMART, Christine D., GRÜNWARD, GREENLAND et FRY, 2010. Quantitative resistance to late blight from *Solanum berthaultii* cosegregates with R Pi-ber: insights in stability through isolates and environment. *Theoretical and Applied Genetics*. Vol. 121, n° 8, pp. 1553-1567. DOI 10.1007/s00122-010-1410-x.
- ROSYARA, DE JONG, DOUCHES et ENDELMAN, 2016. Software for Genome-Wide Association Studies in Autopolyploids and Its Application to Potato. *The Plant Genome*. Vol. 9, n° 2, pp. 0. DOI 10.3835/plantgenome2015.08.0073.
- ROUSSELLE, ROBERT et CROSNIER, 1996. *La pomme de terre: Production, amélioration, ennemis et maladies, utilisations*. ISBN 978-2-7380-0676-9.
- SCHMITZ CARLEY, COOMBS, DOUCHES, BETHKE, PALTA, NOVY et ENDELMAN, 2017. Automated tetraploid genotype calling by hierarchical clustering. *Theoretical and Applied Genetics*. Vol. 130, n° 4, pp. 717-726. DOI 10.1007/s00122-016-2845-5.
- SHARMA, MACKENZIE, MCLEAN, DALE, DANIELS et BRYAN, 2018. Linkage Disequilibrium and Evaluation of Genome-Wide Association Mapping Models in Tetraploid Potato. *G3 Genes|Genomes|Genetics*. Vol. 8, n° 10, pp. 3185-3202. DOI 10.1534/g3.118.200377.
- SPOONER, GHISLAIN, SIMON, JANSKY et GAVRILENKO, 2014. Systematics, Diversity, Genetics, and Evolution of Wild and Cultivated Potatoes. *The Botanical Review*. 2014. Vol. 80, n° 4, pp. 283-383. DOI 10.1007/s12229-014-9146-y.
- STEWART, BRADSHAW et PANDE, 2003. The effect of the presence of R-genes for resistance to late blight (*Phytophthora infestans*) of potato (*Solanum tuberosum*) on the underlying level of field resistance. *Plant Pathology*. Vol. 52, n° 2, pp. 193-198. DOI 10.1046/j.1365-3059.2003.00811.x.

- STICH, URBANY, HOFFMANN et GEBHARDT, 2013. Population structure and linkage disequilibrium in diploid and tetraploid potato revealed by genome-wide high-density genotyping using the SolCAP SNP array. *Plant Breeding*. Vol. 132, n° 6, pp. 718-724. DOI 10.1111/pbr.12102.
- STOREY et TIBSHIRANI, 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences*. Vol. 100, n° 16, pp. 9440-9445. DOI 10.1073/pnas.1530509100.
- TOMCZYŃSKA, STEFAŃCZYK, CHMIELARZ, KARASIEWICZ, KAMIŃSKI, Piotr, JONES, LEES et ŚLIWKA, 2014. A locus conferring effective late blight resistance in potato cultivar Sárpo Mira maps to chromosome XI. *Theoretical and Applied Genetics*. Vol. 127, n° 3, pp. 647-657. DOI 10.1007/s00122-013-2248-9.
- VAN DER PLANK, 1971. Stability of resistance to *Phytophthora infestans* in cultivars without R genes. *Potato Research*. 1971. Vol. 14, n° 4, pp. 263-270. DOI 10.1007/BF02355989.
- VAN DER VOSSSEN, SIKKEMA, HEKKERT, GROS, STEVENS, Patricia, MUSKENS, WOUTERS, PEREIRA, STIEKEMA et ALLEFS, 2003. An ancient R gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato. *The Plant Journal*. Vol. 36, n° 6, pp. 867-882. DOI 10.1046/j.1365-3113X.2003.01934.x.
- VISKER, KEIZER, VAN ECK, JACOBSEN, COLON et STRUIK, 2003. Can the QTL for late blight resistance on potato chromosome 5 be attributed to foliage maturity type? *Theoretical and Applied Genetics*. Vol. 106, n° 2, pp. 317-325. DOI 10.1007/s00122-002-1021-2.
- VOORRIPS, GORT et VOSMAN, 2011. Genotype calling in tetraploid species from bi-allelic marker data using mixture models. *BMC Bioinformatics*. Vol. 12, n° 1. DOI 10.1186/1471-2105-12-172.
- WANG, ALLEFS, VAN DEN BERG, VLEESHOUWERS, VAN DER VOSSSEN et VOSMAN, 2008. Allele mining in *Solanum*: conserved homologues of *Rpi-blb1* are identified in *Solanum stoloniferum*. *Theoretical and Applied Genetics*. Vol. 116, n° 7, pp. 933-943. DOI 10.1007/s00122-008-0725-3.
- XU, PAN, CHENG, ZHANG, MU, NI, ZHANG, YANG, LI, WANG, ORJEDA, GUZMAN, TORRES, LOZANO, PONCE, MARTINEZ, DE LA CRUZ, CHAKRABARTI, PATIL, SKRYABIN, KUZNETSOV, RAVIN, KOLGANOVA, BELETSKY, MARDANOV, DI GENOVA, BOLSER, MARTIN, LI, YANG, KUANG, HU, XIONG, BISHOP, SAGREDO, MEJÍA, ZAGORSKI, GROMADKA, GAWOR, SZCZESNY, HUANG, ZHANG, LIANG, HE, LI, HE, XU, ZHANG, XIE, DU, QU, BONIERBALE, GHISLAIN, DEL ROSARIO HERRERA, GIULIANO, PIETRELLA, PERROTTA, FACELLA, O'BRIEN, FEINGOLD, BARREIRO, MASSA, DIAMBRA, WHITTY, VAILLANCOURT, LIN, MASSA, GEOFFROY, LUNDBACK, DELLAPENNA, ROBIN BUELL, SHARMA, MARSHALL, WAUGH, BRYAN, DESTEFANIS, NAGY, MILBOURNE, THOMSON, FIERS, JACOBS, NIELSEN, SØNDERKÆR, IOVENE, TORRES, JIANG, VEILLEUX, BACHEM, DE BOER, BORM, KLOOSTERMAN, VAN ECK, DATEMA, TE LINTEL HEKKERT, GOVERSE, VAN HAM et VISSER, 2011. Genome sequence and analysis of the tuber crop potato. *Nature*. Vol. 475, n° 7355, pp. 189-195. DOI 10.1038/nature10158.

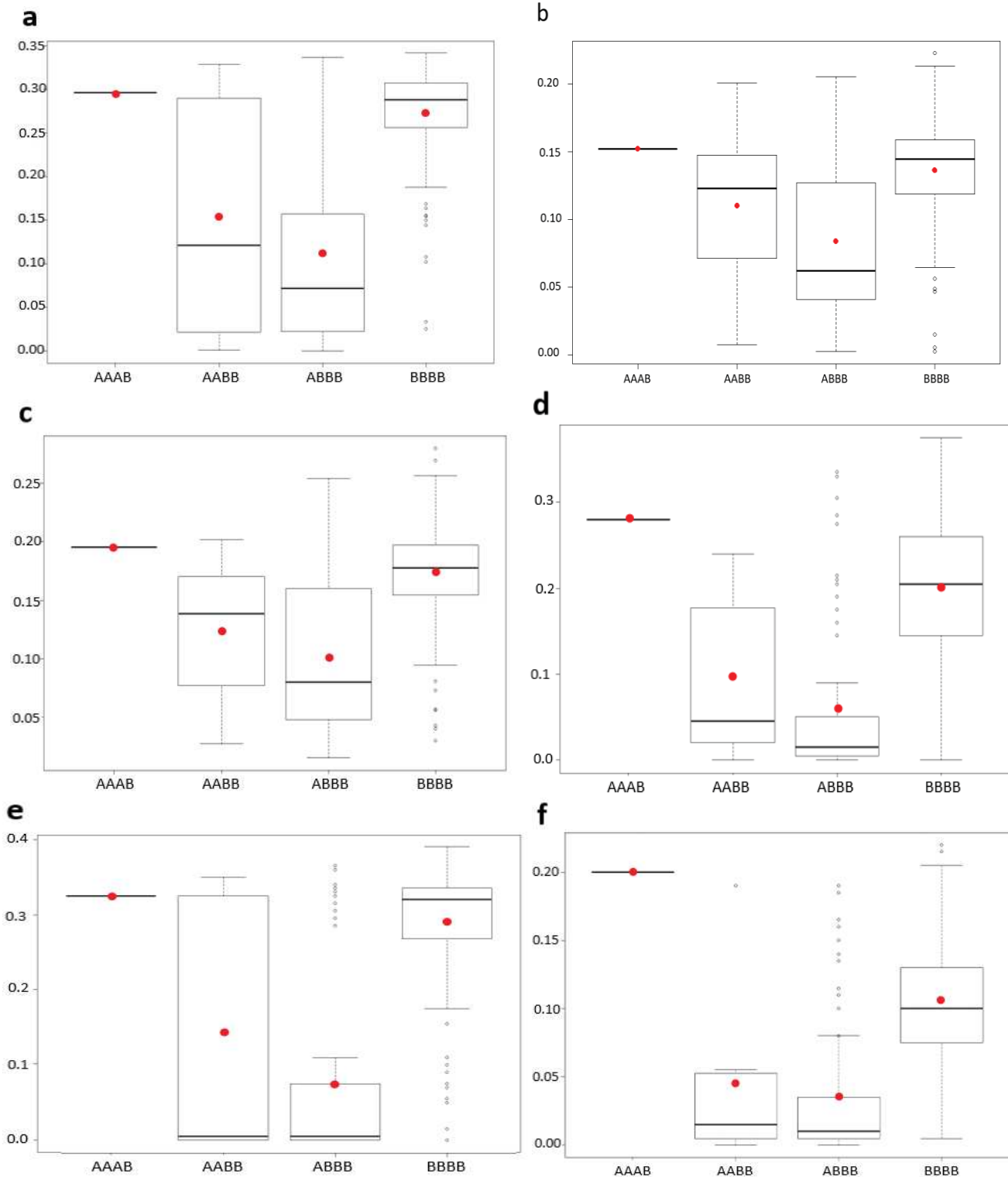
Annexes

Annexe I : Manhattan plot des SNP détectés par la première analyse selon leur modèle.



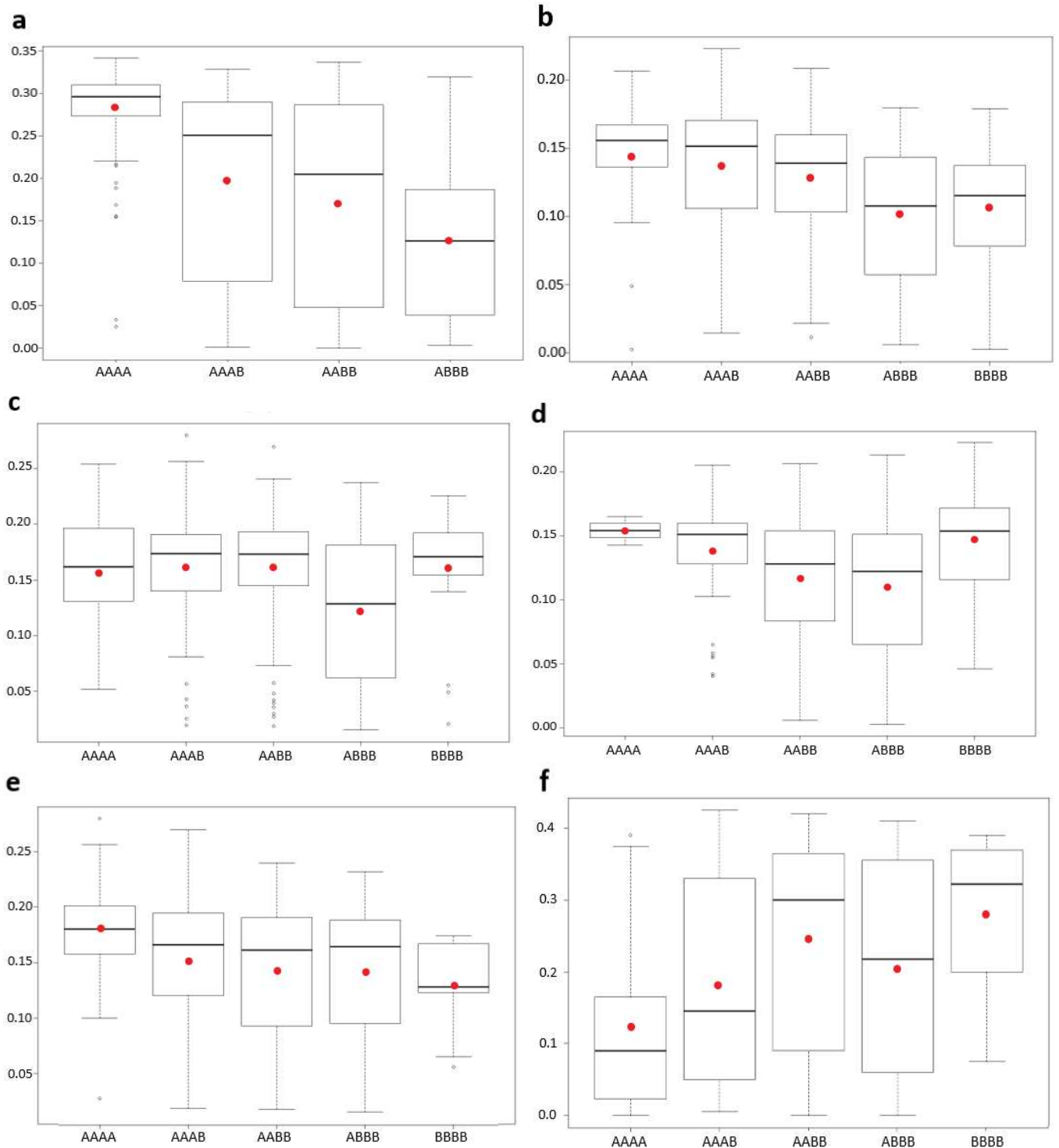
a : Pour le SNP, c2_56418 (chromosome 9), détecté avec un modèle additif (site du Finistère Nord (INRA) en 2017). **b** : Pour les SNP, C2_56418 et C2_27763 (chromosome 9), détectés avec un modèle additif (site du Finistère Nord en 2017). **c** : Pour le SNP, c2_23049 (chromosome 5), détecté avec un modèle additif (site du Pas-de-Calais en 2017). **d** : Pour le SNP, c2_18945 (chromosome 8), détecté avec un modèle duplex dominant avec l'allèle de référence (site de l'Aisne en 2017). **e** : Pour les SNP, c2_41467 (chromosome 8) et c1_10762 (chromosome 4), détecté avec un modèle duplex dominant avec l'allèle alternatif (site du Finistère Nord (INRA) en 2018). **f** : Pour le SNP, c2_36060 (chromosome 4), détecté avec un modèle duplex dominant avec l'allèle de référence (site du Finistère Nord (INRA) en 2017). Pour les modèles duplex dominants l'allèle de référence est A et l'allèle alternatif : B.

Annexe II : Distribution des valeurs de rAUDPC en fonction des classes génotypiques du SNP solcap_snp_c1_8560 du chromosome 9




Le point rouge représente la moyenne. Les graphiques de **a** à **c** représentent la distribution de différents sites en 2017 : **a** : Finistère Nord ; **b** Lot-et-Garonne ; **c** l'Aisne. Les graphiques de **d** à **f** représentent la distribution sur différents sites en 2018 : **d** : Finistère Nord ; **e** : Lot-et-Garonne ; **f** : l'Aisne.

Annexe III : Distribution des valeurs de rAUDPC en fonction des classes géotypiques des SNP détectés par la première analyse.



Le point rouge représente la moyenne. **a** : pour le SNP c2_27763 du chromosome 9 (Site du Finistère Nord, 2017). **b** : Pour le SNP c2_23049 du chromosome 5 (Site du Lot-et-Garonne, 2017). **c** : Pour le SNP c2_18945 du chromosome 8 (Site de l'Aisne, 2017). **d** : Pour le SNP c2_41467 du chromosome 8 (Site du Lot-et-Garonne, 2017). **e** : Pour le SNP c2_36060 du chromosome (Site de l'Aisne, 2017). **f** : pour le SNP c1_10762 du chromosome 4 (site du Finistère Nord (INRA , 2018).

	Diplôme et Mention : Master Biologie, Agrosciences Parcours : Amélioration, Production et Valorisation du Végétal Option : Génétique, Génomique et Amélioration des Plantes Responsable d'option : Mélanie JUBAULT
Auteur(s) : Nadège LE GARREC Date de naissance* : 07/02/1993	Organisme d'accueil : INRA Adresse : Keraïber 29260 Ploudaniel
Nb pages : 20 Annexe(s) : 3	Maître de stage : Florence ESNAULT
Année de soutenance : 2019	
Titre français : Identification de marqueurs associés à une résistance stable au mildiou de la pomme de terre par génétique d'association	
Titre anglais : Genome wide association study to identify markers linked to a stable resistance of the potato to late blight	
<p>Résumé : Une des maladies qui causent le plus de dégâts sur la culture de pomme de terre est le mildiou, causée par l'oomycète : <i>Phytophthora infestans</i>. La lutte chimique utilisée pour faire face à cet agent pathogène contribue au fort Indice de Fréquence de Traitement (IFT) de cette culture. Le plan Ecophyto II a pour objectif de diminuer l'utilisation de produits phytosanitaires. L'identification de résistances génétiques quantitatives, qui sont plus stables, constitue donc un enjeu majeur.</p> <p>Mon sujet de stage s'insère dans le projet PotStaR qui vise à identifier des sources de résistance stable au mildiou et des marqueurs moléculaires associés à ces résistances stables. Mon stage a porté sur la mise en œuvre d'une stratégie de génétique d'association sur une collection de géniteurs améliorés issus de différents programmes de recherche de l'INRA de Ploudaniel. Les données de phénotypage proviennent de cinq sites d'expérimentation et les données de génotypage d'un génotypage avec la puce SolCAP 8303 SNP. Pour commencer une étude de structuration génétique de la population a été faite. Ensuite différentes analyses d'association ont permis d'identifier des marqueurs associés à la résistance. Une analyse du déséquilibre de liaison a été faite pour les marqueurs qui se situent dans les mêmes régions génomiques. Le QTL le plus stable identifié par cette étude se situe sur le chromosome 9.</p>	
<p>Abstract : One of the most destructive disease for potato crop is late blight caused by an oomycete : <i>Phytophthora infestans</i>. Fungicides are massively used to fight against this pathogen, which contributes to the high treatment frequency index of this crop. The Ecophyto II plan aims at reducing the use of pesticides. The identification of quantitative genetic resistance, which is more stable, represents therefore a major issue.</p> <p>My internship subject is part of the PotStaR project whose purpose is to identify stable sources of resistance to the late blight and molecular markers linked to these stable resistance. My internship was about the development of a strategy of association mapping on a collection of improved breeding lines originating from research programs from INRA of Ploudaniel. The phenotypic data come from five trials and the genotypic data from the use of the 8303 SNP SolCAP array. First, a study of the population structure was performed. Then markers linked to late blight resistance have been identified. An analysis of linkage disequilibrium was carried out for the markers that are located in the same genomic region. The most stable QTL identified in this study is located on chromosome 9.</p>	
Mots-clés : Pomme de terre; Mildiou; Génétique d'association; Résistance quantitative; Puce SolCAP Key Words : Potato; Late Blight ; Genome-wide association; Quantitative resistance; SolCAP array	

* Élément qui permet d'enregistrer les notices auteurs dans le catalogue des bibliothèques universitaires