



Recherche de marqueurs moléculaires de la résistance aux bio-agresseurs (anthracnose et verticilliose) chez la luzerne pérenne

Bernadette Julier

► To cite this version:

Bernadette Julier. Recherche de marqueurs moléculaires de la résistance aux bio-agresseurs (anthracnose et verticilliose) chez la luzerne pérenne: Compte-rendu final du contrat de branches C2008-16. [Contrat] INRA; ACVF Luzerne. 2011. hal-02987000

HAL Id: hal-02987000

<https://hal.inrae.fr/hal-02987000>

Submitted on 3 Nov 2020

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

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

Recherche de marqueurs moléculaires de la résistance aux bio-agresseurs (anthracnose et verticilliose) chez la luzerne pérenne

Contrat de Branches C2008-16 - Luzerne

Rapport à l'issue de la troisième année (Octobre 2011)

I. Identification d'un gène de résistance à l'anthracnose (*Colletotricum trifolii*) chez la luzerne (*Medicago sativa*)

Enjeux économiques et questions scientifiques

La résistance aux maladies constitue un enjeu majeur de l'amélioration de la luzerne (*Medicago sativa*), une légumineuse fourragère majeure en France et dans le monde. L'intégration d'innovations technologiques permettant une accélération du progrès génétique est indispensable. Les connaissances acquises chez la légumineuse modèle *Medicago truncatula* peuvent probablement être transposées chez la luzerne. Dans ce projet, le transfert entre *M. truncatula* et *M. sativa* est un enjeu fort, puisqu'il s'agit de démontrer, pour un caractère simple, que le déterminisme génétique de la variabilité phénotypique est le même chez les deux espèces.

L'amélioration de la résistance à l'anthracnose de la luzerne est nécessaire pour maintenir la culture de cette espèce. Cependant, il est bien établi qu'une diversité de niveaux de résistance au sein d'une culture est nécessaire pour que les pathogènes ne contournent pas les résistances des plantes. Il convient donc de disposer d'outils pour sélectionner des variétés ayant un niveau de résistance fort mais pas total.

L'objectif principal est de prouver que le polymorphisme allélique chez la luzerne pour le gène RCT1, décrit pour gouverner la résistance au *Colletotricum trifolii* chez *M. truncatula* (Yang et al. 2007 ; Yang et al. 2008), est relié au niveau de résistance à cette même maladie. Le polymorphisme permettra de définir des marqueurs moléculaires utilisables en sélection. De tels marqueurs permettront de piloter le niveau de résistance des variétés de luzerne (les variétés sont des populations synthétiques, génétiquement hétérogènes), en évitant d'instaurer un monomorphisme pour ce gène, qui aurait pour conséquence certaine le contournement de la résistance par le pathogène.

Démarche

La démarche repose sur une approche « gène candidat » faisant l'hypothèse que le gène de résistance RCT1 identifié chez *M. truncatula* est aussi responsable de la résistance chez la luzerne pour la même maladie. L'approche retenue est celle de la BSA (Bulk Segregant Analysis) dans laquelle on crée des groupes d'individus ayant les mêmes caractéristiques phénotypiques (ici résistance ou sensibilité), et pour lesquels on cherche des marqueurs moléculaires qui différencient les deux groupes et qui seront liés au caractère phénotypique. Ici, les marqueurs n'appartiendront qu'au gène RCT1. La validation de ce gène sera effectuée en utilisant la génétique d'association. Cette méthode est particulièrement adaptée aux espèces fourragères dont les variétés sont des populations synthétiques, très variables et peu structurées (Flajoulot et al. 2005), et chez lesquelles le déséquilibre de liaison est très petit (Auzanneau et al. 2007).

Etat des connaissances

Le gène de résistance a été cloné chez *M. truncatula* (3297 pb) à partir de la localisation du QTL de résistance sur le chromosome 4 (Yang et al. 2007 ; Yang et al. 2008). Le gène (cDNA) homologue de RCT1 chez *M. sativa* (3378 pb) a aussi été séquencé (Yang et al. 2008), le pourcentage de similitude

de ce gène entre ces deux espèces est de 90,6%. Sur la base de données NCBI, nous avons recensé trois clones BAC de *M. truncatula* comprenant le gène RCT1. Ce gène est composé de 5 exons et 4 introns et sa longueur est de 6 kb (5' et 3' UTR, UnTranslated Region, inclus). Yang et al. (2008) ont comparé les séquences du gène (introns exclus) entre deux lignées pures de *M. truncatula*, l'une résistante (Jemalong) et l'autre sensible (F83005.5). Au total, 27 SNP (Single Nucleotide Polymorphism), une insertion/délétion de 2 pb, une insertion/délétion de 48 pb dans le 5' UTR ainsi qu'une insertion/délétion de 199 pb dans le 3' UTR ont été identifiés. La seule différence présentée dans ce papier sur une dizaine de lignées est une excision alternative de l'intron 4 chez les lignées résistantes, l'intron 4 est en grande majorité excisé chez les lignées sensibles. Le gène RCT1 confère une résistance à l'anthracnose à large spectre quand il est transféré dans des clones sensibles de *M. sativa*. Pour rappel, le contexte génétique de *M. truncatula* et *M. sativa* est totalement différent. La plante modèle, *M. truncatula*, a été choisie pour sa génétique simple (diploïde et autogame) et la possibilité d'obtenir des lignées pures. A l'opposé, la luzerne a une génétique complexe, elle est autotétraploïde et allogame. Les variétés de luzerne sont des populations synthétiques très variables et peu structurées. La résistance à l'anthracnose est dominante sur la sensibilité (Yang et al ; 2007). Un individu sensible de luzerne a donc quatre allèles de sensibilité (SSSS) alors que l'individu résistant peut avoir un à quatre allèles de résistance et zéro à trois allèles de sensibilité (RRRR, RRRS, RRSS, RSSS).

Acquis des années 1 et 2 (2008-09 et 2009-10)

Nous avons séquencé le gène RCT1 chez un individu de luzerne. Il est sensiblement plus long chez la luzerne que chez *M. truncatula*. Nous avons réalisé des tests de résistance chez 8 variétés de luzerne, permettant d'isoler chez chaque variété un pool de 15 individus sensibles et un autre pool de 15 individus résistants, ces pools constituant le matériel de base pour l'approche BSA. Sur ces individus, le gène RCT1, amplifié en deux portions de 6.5 et 8 kb, a été séquencé en utilisant un séquenceur de nouvelle génération à haut débit. Une première analyse des données avait montré qu'on avait amplifié le gène ciblé et que le polymorphisme est très important à l'intérieur de chaque pool.

Objectif de l'année 3 (2010-11)

Devant la somme des données de séquences et le grand niveau de polymorphisme révélé, il était nécessaire de traiter ces données avec des outils de la bioinformatique. Pour cela, il fallait mobiliser un collègue bioinformaticien de Toulouse (Jérôme Gouzy, LIPM, INRA). A Lusignan, un CDD de niveau ingénieur devait être recruté pour finaliser les analyses.

Résultats 2010-11

Traitement bioinformatique des données

La séquence du gène obtenue après clonage a servi de référence. Cette séquence de référence a servi à l'assemblage des séquences brutes issues du séquenceur nouvelle génération. Ceci a permis de créer des séquences consensus pour chaque pool de 15 individus (J. Gouzy, Toulouse), séquences qu'on peut assimiler à des allèles (ou plutôt à des fragments d'allèles).

Le polymorphisme de la région amplifiée étant très important, on s'est concentré sur les régions codantes dans lesquelles tout changement est susceptible d'avoir une incidence en modifiant directement la protéine (sauf mutations synonymes). Tous les alignements ont été effectués sous Bioedit.

Identification des exons

Pour cela, les exons décrits sur *Medicago truncatula* (gi|193584699, Yang 2007) ont été alignés individuellement sur l'ADNc de *Medicago sativa* (gi|193584701, Yang 2007).

Les exons ainsi déterminés ont été alignés sur les séquences chimères.

Référence+ (sens de lecture du gène)

Position exon 1 : 7839 à 8473

Position exon 2 : 10100 à 11226

Position exon 3 : 11550 à 11824

Position exon 4 : 12046 à 12873

Position exon 5 : 14140 à 14679

Référence- (brin complémentaire du gène)

Position exon 1 : 8257 à 8891

Position exon 2 : 5504 à 6630

Position exon 3 : 4906 à 5180

Position exon 4 : 3857 à 4684

Position exon 5 : 2051 à 2590

Les exons ont été identifiés sur les séquences chimères des alignements à l'aide de « features » pour une meilleure lisibilité.

Alignement des séquences nucléotidiques des exons

Les alignements des séquences ADN pour chaque exon ont été obtenus à partir des alignements des séquences consensus à la fois des individus sensibles et résistants d'une même variété contre la séquence référence. Pour cela, seules les séquences consensus donnant suffisamment d'information pour l'exon concerné ont été conservées. De même, ces séquences ont été recadrées pour ne contenir que l'exon.

Traduction en protéines

Les séquences des alignements précédents ont été nettoyées de tous leurs « gaps » et traduites en protéines. Ces fragments de protéine déduits ont été alignés avec la protéine de *Medicago sativa* enregistrée dans NCBI (gi|193584702). Un grand nombre de séquences nucléotidiques possède des insertions/délétions de petites tailles qui modifient le cadre de lecture rendant impossible l'alignement. Cela pouvant être dû à une mauvaise qualité des séquences, ces séquences ont été modifiées par l'ajout de gaps pour retrouver le même cadre de lecture que la protéine enregistrée.

La position des polymorphismes est notée à partir de la séquence de *M. sativa* enregistrée sur NCBI (gi|193584702).

Tous les « allèles » identifiés étaient communs aux différentes variétés et aux pools Résistant et Sensible. Aucun n'est donc particulièrement « candidat » pour expliquer le polymorphisme étudié.

On s'attendait pourtant à identifier des allèles spécifiques des pools Résistants ou Sensibles, c'est le principe même de la méthode BSA choisie. Ce résultat est donc négatif. Néanmoins, la méthode de contigage a pu induire une perte de la notion de « nombre d'allèles » dans chacun des pools. Or, une erreur de phénotypage peut conduire à placer un individu résistant dans un pool sensible. Dans ce cas, l'allèle de résistance qu'il porte sera identifié dans le pool sensible. Pour préciser la situation, un génotypage individuel des plantes contenues dans les pools a été réalisé pour les sites majeurs de polymorphisme.

Définition de marqueurs des délétions de la région codante

Les séquences de toutes les variétés ont été réunies dans un seul alignement par exon. A partir de ces alignements, 5 sites de délétions potentielles ont été identifiés.

On suppose que l'allèle résistant possède la séquence la plus longue possible et donc aucune délétion.

Quatre couples d'amorces PCR ont été définis à l'aide de Primer-BLAST sur NCBI à partir des séquences obtenues. Les séquences qui y sont enregistrées ont été obtenues à partir des alignements

nucléotidiques dont seules les régions suffisamment conservées pouvant servir à la définition des amorces ont été gardées (figure 1).

Les amorces ont été définies à l'extérieur des zones pouvant être sujettes à une délétion (Table 1). Le but est de pouvoir observer un polymorphisme de taille. Trois sites de polymorphisme décalent le cadre de lecture, sur l'exon 1, l'exon 3 et l'exon 5, impliquant une modification profonde de la protéine. En particulier, la délétion dans l'exon 1 supprime l'ATG de début de gène, elle devrait conduire à l'absence de la protéine (perte totale de fonction). Les deux autres sites, sur l'exon 3 et une l'exon 5, ne modifient pas le cadre de lecture mais modifient la protéine et donc peut-être son activité.

Les produits PCR des 30 individus (Résistants et Sensibles) des 8 variétés étudiées ont été obtenus pour les 4 couples d'amorces et ont été déposés sur gel d'acrylamide dans un séquenceur Licor. Les tailles des bandes ont été lues sur SAGA. Les résultats ont été lus en tant que présence/absence de chaque allèle ou en tant que dose de chaque allèle.

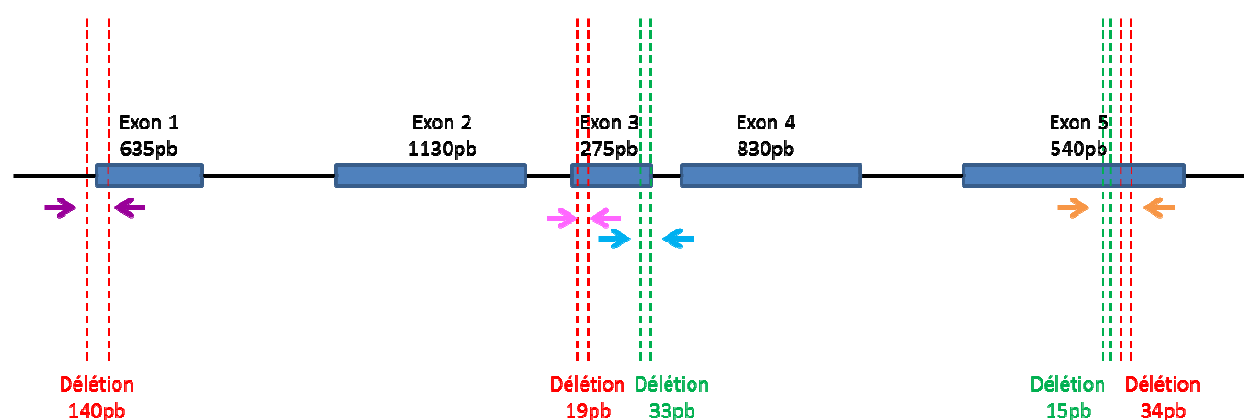


Figure 1. Délétions pouvant affecter les exons du gène RCT1-like et amorces PCR correspondantes

Les délétions notées en rouge décalent le cadre de lecture. La protéine alors produite ne correspondrait plus à la protéine telle que décrite dans la bibliographie (sur *Medicago truncatula*).

Au contraire, les délétions notées en vert ne modifient pas le cadre de lecture.

Table 1. Couples d'amorces permettant de révéler le polymorphisme de taille généré par les délétions

amorce	Sequence 5' → 3'	Site ciblé	Taille fragment amplifié (pb)
RCT1-1_F	ACAAAGCACTTCCCAACTTCACAGA	Délétion 1	104 ou 244
RCT1-1_R	TCCAGCGAAGCCAGATCGTTCT		
RCT1-2_F	ACAGGGAACAAAAGCTGTTGAGGGA	Délétion 2	106 ou 125
RCT1-2_R	GCTGTACACCAGAGAGTTGCAGC		
RCT1-3_F	TGCTGCAACTCTCTGGTGTACAGC	Délétion 3	243 ou 276
RCT1-3_R	CGTGCTTTGGACGCAACACTCA		
RCT1-4_F	AGTCCCTCAAGTGGGAAGGCCGT	Délétion 4 et 5	4 tailles possibles : -291 (1e délétion) -272 (2e délétion) -306 (aucune délétion) -257 (2 délétions à la fois, non observé lors du séquençage)
RCT1-4_R	TGCTCCAATATTTGCCGACTGGT		

Les amorces Forward contenaient en plus une queue M13

Programme PCR utilisé

95°C	5 mn	
95°C	30 sec	} x 8
62°C	-1°/cycle 30sec	
72°C	30 sec	
95°C	30 sec	} x 17
55°C	30 sec	
72°C	30 sec	
72°C	5 mn	
10°C	Forever	

Quels sont les attendus ?

La résistance est dominante sur la sensibilité. Cela signifie que les individus sensibles ne porteront que des allèles de sensibilité, ils sont donc de génotype SSSS. Par contre, les individus résistants peuvent porter des allèles de sensibilité, ils peuvent être RRRR, RRRS, RRSS ou RSSS. Cette notation avec des allèles R et S est schématique, en effet, on ne sait pas s'il existe plusieurs allèles de sensibilité ou plusieurs allèles de résistance. On peut supposer que plusieurs allèles de sensibilité seront observés car un allèle inactif est susceptible d'accumuler d'autres mutations.

Au total, dans les pools S, sauf mauvais classement d'un individu (un individu R), on ne doit trouver que des allèles de sensibilité. Au contraire, dans les pools R, on doit trouver à la fois des allèles de sensibilité et de résistance. Leur proportion variera en fonction de la fréquence moyenne de la résistance phénotypique, en supposant que les variétés sont à l'équilibre panmictique (cf Julier et Bournoville, Euphytica 2004) :

Si P est la fréquence des individus résistants, et qu'on note s la fréquence de l'allèle de sensibilité (r , la fréquence de l'allèle de résistance étant $r = 1 - s$), alors $P = 1 - s^4$.

Autre point d'importance : on ne sait pas si les allèles de résistance / sensibilité sont les mêmes dans les différentes variétés. D'après des études antérieures montrant le faible niveau de différenciation entre variétés, on peut supposer que le déterminisme de la résistance est le même dans toutes les variétés.

Enfin, dans un contexte où on identifie plusieurs sites polymorphes, il serait souhaitable d'avoir leur phase. En effet un génotype ABBB au site 1 et CDDD au site 2 peut être sous deux configurations au niveau des allèles :

AC	ou	AD
BD		BC
BD		BD
BD		BD

Si on considère que seule la configuration AC permet la résistance, alors le premier individu est résistant alors que le second est sensible. Le génotypage tel qu'il a été réalisé ne permet pas cette finesse d'analyse.

Résultats sur les fréquences alléliques dans les variétés

En tenant compte des doses alléliques déterminées pour chaque individu, les effectifs de chaque allèle ont été calculés et sont présentés dans les tables 2 à 5 pour chaque site polymorphe. On a pratiqué des tests de χ^2 pour tester si lorsqu'un allèle est présent il y a autant d'individus sensibles que résistants.

Les effectifs peu importants pour les allèles non majoritaires n'ont pas permis d'effectuer ces tests pour tous les sites (moins de 5 par classe). Seuls les sites de l'exon 1 et du début de l'exon 3 ont pu être testés.

Pour le site de l'exon 1 (Table 2), seul l'allèle 261 est fréquent. Les autres allèles sont trop peu fréquents pour être responsables de la résistance ou de la sensibilité. Pour les 3 allèles, les 2 pools de chaque variété ont des fréquences très similaires. L'allèle 122 qui correspond à la délétion de l'ATG est peu fréquent et réparti uniformément dans les pools et les variétés. Cependant, le test Chi² pour cet allèle 122, effectué sur le cumul de toutes les variétés, est significatif ($P = 0,06$) alors qu'il n'est pas significatif pour l'allèle 264 ($P = 0,56$).

Table 2. Nombre d'allèles de taille 122, 261 et 264 de la région polymorphe de l'exon 1 pour chaque variété, en tenant compte de la dose de chaque allèle

Variété	Résistance	Nb individus	Allèle 122	Allèle 261	Allèle 264	Donnée manquante
Bar1	R	17	3	59	6	0
	S	17	1	56	7	1
Bar2	R	17	1	56	11	0
	S	17	0	52	12	1
Cannelle	R	17	4	50	6	2
	S	17	1	61	6	0
Galaxie	R	14	6	41	1	2
	S	17	4	52	8	1
Kali	R	17	9	48	11	0
	S	17	3	56	5	1
Marshall	R	17	1	48	13	1
	S	7	0	24	4	0
Prunelle	R	17	2	54	8	1
	S	17	3	61	4	0
Symphonie	R	14	2	49	5	0
	S	17	1	47	8	3
Total	R	130	28	405	61	6
	S	126	13	409	54	7

NB : ces tailles d'allèles incluent la queue M13 qui comprend 20 nucléotides

Concernant le site du début de l'exon 3, deux allèles ont été observés, l'allèle 147 étant plus fréquent que l'allèle 124 (Table 3). Tous deux sont de fréquence identique dans chaque pool et chaque variété (pour l'allèle 124, $P = 0,4$).

Pour ce site, parmi les 19 individus qui n'ont pas donné de produit PCR, 15 sont résistants et seulement 4 sont sensibles ($P = 0,01$). On peut supposer qu'il existe un autre « allèle » qui consiste en une mutation dans le site d'hybridation d'une des 2 amorces. Cet allèle pourrait éventuellement être lié à la résistance mais c'est un allèle rare qui ne peut pas expliquer à lui seul le phénotype de résistance/sensibilité. Cet allèle pourrait être la mutation G/T à la position 11558 (position notée à partir de la référence+). Dans les alignements de séquences consensus (« contigs »), cette mutation est présente chez 4 séquences issues de groupes résistants et 2 séquences de groupes sensibles.

Table 3. Nombre d'allèles de taille 124 et 147 de la région polymorphe du début de l'exon 3 pour chaque variété, en tenant compte de la dose de chaque allèle

Variété	Résistance	Nb individus	Allèle 124	Allèle 147	Donnée manquante
Bar1	R	17	11	57	0
	S	17	7	53	1
Bar2	R	17	12	52	1
	S	17	9	58	0
Cannelle	R	17	11	41	4
	S	17	12	56	0
Galaxie	R	14	11	37	2
	S	17	11	57	0
Kali	R	17	13	55	0
	S	17	15	53	0
Marshall	R	17	13	47	2
	S	7	7	21	0
Prunelle	R	17	9	47	3
	S	17	12	52	1
Symphonie	R	14	10	34	3
	S	17	10	50	2
Total	R	130	90	370	15
	S	126	84	404	4

Table 4. Nombre d'allèles de taille 262, 294, 295, 297 et 298 de la région polymorphe de la fin de l'exon 3 pour chaque variété, en tenant compte de la dose de chaque allèle

Variété	Résistance	Nb individus	Allèle 262	Allèle 294	Allèle 295	Allèle 297	Allèle 298	Donnée manquante
Bar1	R	17	3	0	65	0	0	0
	S	17	0	4	61	0	3	0
Bar2	R	17	0	4	43	0	1	5
	S	17	0	0	68	0	0	0
Cannelle	R	17	0	0	51	2	3	3
	S	17	0	0	60	0	0	2
Galaxie	R	14	0	0	46	1	1	2
	S	17	0	0	32	0	0	9
Kali	R	17	0	0	40	0	0	7
	S	17	0	0	47	1	0	5
Marshall	R	17	0	0	48	0	0	5
	S	7	0	0	12	0	0	4
Prunelle	R	17	0	0	64	0	0	1
	S	17	0	0	65	0	3	0
Symphonie	R	14	0	0	56	0	0	0
	S	17	0	0	52	0	0	2
Total	R	130	3	4	413	3	5	23
	S	126	0	4	405	1	5	22

Pour le site de la fin de l'exon 3, cinq allèles ont été observés. Quatre (allèles 262, 294, 297 et 298) sont très peu fréquents et ne peuvent pas expliquer le polymorphisme. L'allèle 295 est très fréquent. Il montre des différences de fréquence entre les pools R et S de certaines variétés (Bar2 par exemple), mais ces différences ne sont pas dans le même sens pour toutes les variétés (comparaison Bar2 et Marshall). Le couple d'amorces génère des données manquantes en grand nombre, mais ces données manquantes sont réparties également dans les deux types de pools.

Table 5. Nombre d'allèles de taille 278, 290, 310 et 325 de la région polymorphe de l'exon 5 pour chaque variété, en tenant compte de la dose de chaque allèle

Variété	Résistance	Nb individus	Allèle 278	Allèle 290	Allèle 310	Allèle 325	Donnée manquante
Bar1	R	17	0	4	0	52	3
	S	17	0	2	0	54	3
Bar2	R	17	0	2	0	54	3
	S	17	0	9	0	47	3
Cannelle	R	17	1	4	0	48	4
	S	17	1	5	0	58	1
Galaxie	R	14	0	3	0	37	4
	S	17	0	3	1	48	4
Kali	R	17	0	6	1	45	4
	S	17	0	0	1	43	6
Marshall	R	17	0	1	1	42	6
	S	7	0	0	0	16	3
Prunelle	R	17	0	3	1	32	8
	S	17	0	3	3	46	4
Symphonie	R	14	0	3	0	53	0
	S	17	0	6	1	61	0
Total	R	130	0	26	3	363	32
	S	126	1	28	6	373	24

Pour le site de l'exon 5, l'allèle 325 est très fréquent, l'allèle 290 est assez rare et les deux allèles 278 et 310 sont très rares. Le test Chi² est non significatif ($P = 0.55$). Sur ce site, des données manquantes sont enregistrées, avec probablement des polymorphismes dans la séquence des amorces. Cependant, la fréquence de ces données manquantes est similaire chez les individus résistants et les individus sensibles.

En complément à des tests Chi qui n'ont pas pu être menés sur toutes les unités expérimentales (variété x site) pour des raisons d'effectifs, on a fait des analyses de variances basées sur le nombre d'individus résistants et sensibles en fonction de la présence/absence des différents allèles. Puis, on a fait des régressions linéaires pour voir l'effet de la dose des allèles sur la fréquence d'individus résistants. Ces tests ont été effectués sous SAS. Les analyses de variances ont été faites à l'aide de la procédure GLM et les régressions linéaires avec la procédure REG.

Ces deux types d'analyse ne permettent pas de voir d'effet d'un allèle particulier sur la résistance. Seules les doses des allèles 122 du site de l'exon 1 et 147 du site du début de l'exon 3 semblent liées à la résistance dans le cas d'une régression linéaire. Cependant, ces allèles étant minoritaires par rapport à l'allèle « sauvage » (allèle 261 pour l'exon 1 et allèle 147 pour le début de l'exon 3), toutes les doses possibles (0, 1, 2, 3 et 4) ne sont pas présentes.

Par ailleurs, ces deux types d'analyses sont adaptés à l'étude de variables quantitatives. Ces analyses ne sont donc pas bien adaptées à l'étude du caractère qualitatif résistant ou sensible.

Tentative de dénombrement des séquences issues du 454 par allèle

Parallèlement à l'utilisation de marqueurs moléculaires, on a essayé de retrouver le nombre de séquences issues du séquençage 454 correspondant à chaque allèle pour les 5 sites de délétions/insertions identifiés. Pour cela, des séquences de quelques centaines de pb ont été définies pour chaque allèle de chaque site. Ces séquences sont les mêmes que celles ayant servies à la définition des amorces PCR. Par ailleurs, les fichiers contenant les séquences de base issues du séquençage 454 ont été importés dans Bioedit en tant que bases de données locales. Un Blast local a

alors été fait pour chaque « allèle » dans chaque base de données. Les séquences ayant une E value < 1,0E-100 ont été dénombrées, en autorisant ou non les gaps.

Les fréquences alléliques obtenues par cette méthode ne correspondent pas aux fréquences réelles déterminées sur Licor avec les marqueurs moléculaires. Le critère de E value n'étant pas le plus pertinent dans le choix des séquences et la proportionnalité entre le nombre de séquences produites par le 454 et la fréquence allélique n'étant pas vérifiée, ces résultats n'ont pas été pris en compte.

Déséquilibre de liaison

Le déséquilibre de liaison a été testé entre les sites 1 et 2, seuls sites ayant des allèles non majoritaires assez fréquents.

Toutes les combinaisons entre ces deux sites ont testées. Toutefois, les combinaisons impliquant l'allèle 147 du site du début de l'exon 3 n'ont pas pu donner lieu à des calculs car l'allèle 147 est toujours présent (accompagné ou non par l'allèle 124).

Parmi les autres combinaisons, seule la combinaison allèle 261-allèle 124 présente un déséquilibre de liaison (χ^2 , $P = 0,03$; Fischer, $P = 0,04$). Cependant, il est à noter que la classe « absence de 261 et présence de 124 » n'est représentée que par 3 individus et non au moins 5 comme recommandé.

Les combinaisons 122-124 (χ^2 , $P = 0,3$; Fischer, $P = 0,58$) et 264-124 (χ^2 , $P = 0,82$; Fischer, $P = 0,89$) ne présentent pas de déséquilibre de liaison.

Discussion

Alignement

Les alignements ont été effectués presque uniquement sur les régions codantes car le moindre changement dans ces régions est susceptible de modifier la protéine produite. Cependant, le polymorphisme dans les introns peut également avoir une grande incidence sur la protéine produite, notamment les signaux de début et de fin d'épissage. Or, ceux-ci ont été peu étudiés, en grande partie parce qu'il est difficile de prévoir la conséquence du polymorphisme dans les introns.

On note dans ces alignements beaucoup de petits polymorphismes (petites insertions/délétions, SNP,...). Mais il est difficile de savoir si ces mutations sont réelles ou dues à un problème de qualité du séquençage 454. On a donc préféré s'intéresser aux mutations que l'on retrouve plusieurs fois dans des séquences différentes (idéalement de variétés différentes).

L'idéal aurait été de pouvoir déterminer les fréquences de chaque mutation observées directement à partir des données de séquençage. Cependant, la réunion des séquences en contigs, nécessaire à l'alignement, empêche de dénombrer les séquences ayant un polymorphisme donné. Par ailleurs, on ne peut pas être sûr que les données de séquençage respectent les fréquences alléliques réelles. Ceci ne permet donc pas de connaître l'importance des différents polymorphismes. De même, ce problème de dénombrement empêche également la caractérisation du jeu de séquences de l'échantillon (calcul de la diversité nucléotidique et autres indices). Toutefois, l'échantillonnage des individus n'ayant pas été fait au hasard dans les différentes variétés (15 plantes sensibles et 15 résistantes), il y aurait eu un biais. Ce n'était d'ailleurs pas le but de l'étude.

Amplification des sites polymorphes

Les 4 couples d'amorces ont permis d'amplifier des fragments d'ADN. Les différents polymorphismes sont moins présents que prévus mais ils existent et en fréquence non négligeable pour certains.

D'autres allèles sont visibles mais le plus souvent en fréquence faible. Seul l'allèle 264 du site 1 a une fréquence de près de 12%.

Les doses de chaque allèle ont pu être déterminées mais la méthode même manque de précision.

Analyse

On n'observe pas beaucoup de déséquilibre de liaison. Ce n'est pas étonnant pour autant vu les marqueurs utilisés (peu fréquents et peu discriminants) et la taille du gène (plusieurs milliers de pb). Un déséquilibre aurait pu s'expliquer par la faible distance entre certains sites et l'appartenance à un même gène (distance entre sites modérée, fonction du gène actuelle ou passée).

Effet des allèles : on ne voit presque rien. Mais, on note quand même que les individus ayant l'allèle 122 du site 1, correspondant à la délétion du codon d'initiation, semblent plus souvent résistants. Il existe dans la littérature des gènes de résistance qui conduisent à une perte de fonction.

Hypothèses

On n'a trouvé aucun polymorphisme dans le gène RCT1-like expliquant la résistance à l'anthracnose. Toutes les étapes du protocole peuvent alors être remises en cause :

- La stratégie BSA basée sur 8 variétés : il aurait peut-être mieux valu prendre une seule variété pour maximiser le nombre d'individus et minimiser le polymorphisme entre individus.
- La souche utilisée et le protocole de test de résistance ne sont pas les mêmes que ceux de la publication américaine.
- Le séquençage nouvelle génération : en poolant les individus, on perd l'information individuelle ; de plus, cette technique conduit à de nombreux petits fragments dont on a perdu la phase. La qualité du séquençage est moins bonne que le séquençage Sanger, on peut attribuer une partie du polymorphisme (SNP ou petits indels) à des erreurs de séquençage.
- Les alignements sont délicats du fait du polymorphisme. Les contigages induisent une perte des informations de nombres de séquences de chaque type.
- Les sites polymorphes identifiés sont réduits à ceux présents dans les exons. D'après les résultats obtenus à l'aide de ces marqueurs, les sites d'insertion/délétion étudiés ne sont pas responsables de la sensibilité observée ou du moins ne l'expliquent pas suffisamment. Peut-être que les sites responsables de la résistance sont présents dans les introns ou en amont ou en aval de la séquence codante.

Cependant, il n'y a pas vraiment de protocole alternatif à proposer. En effet, le séquençage de type Sanger aurait conduit à réduire le nombre de données de séquences, en choisissant quelques portions du gène et quelques individus pour établir le polymorphisme. Le choix de certaines régions au gène aurait été périlleux à cause de la longueur du gène, ce qui est confirmé par l'absence de déséquilibre de liaison entre les 2 sites testés.

Il est également possible que le gène RCT1 soit régulé par un autre gène ou que la protéine produite subisse des modifications post-transcription (dont épissage alternatif qui a été mis en évidence sur *M. truncatula*).

Il existe peut-être un autre gène de résistance qui a un effet plus important que RCT1 chez les variétés observées.

Le fait qu'un ou plusieurs autres couples gènes R/Avr soient en jeu dans le phénomène de résistance pourrait entraîner une absence de forte sélection dans le gène RCT1 et ainsi expliquer l'existence de nombreux allèles différents. L'utilisation d'une souche et d'un protocole différent de ceux ayant permis d'identifier RCT1 conforterait cette hypothèse.

La présence de nombreux allèles différents pourrait s'expliquer par un modèle évolutif de type « guerre de tranchées » (grande diversité allélique à la fois du gène de résistance et du gène d'avirulence correspondant, ce qui permet le maintien d'un équilibre dans le temps). Cependant, dans ce cas, pour le gène de résistance, supposé RCT1, ayant subi plusieurs cycles de sélection, le nombre d'allèles aurait dû être fortement diminué. Les tests phénotypiques étant fait avec la même souche, ce qui correspond à un seul allèle pour « avr-rct1 », un seul allèle du gène de résistance aurait dû être sélectionné.

L'allèle 122 de l'exon 1 semble lié à la résistance. Cet allèle est délété dans l'ATG du début du gène, il conduit donc à une protéine inactive. Soit, par déséquilibre de liaison, cette délétion est liée à une résistance, même si ce n'est pas ce polymorphisme qui explique la résistance dans le matériel étudié. Soit la résistance à l'anthracnose est acquise par une perte de fonction, un cas déjà décrit dans la littérature sur d'autres couples espèce végétale – pathogène. Dans cette situation, le polymorphisme identifié, même s'il n'explique pas globalement la résistance dans la population étudiée, peut générer une résistance. Une validation de cette hypothèse pourrait être recherchée en réalisant des polycross entre individus portant ou ne portant pas cette délétion et en comparant leurs niveaux de résistance. Les marqueurs pour réaliser ce test sont disponibles grâce à cette étude. La seule précaution à prendre est de génotyper un grand nombre d'individus dans une population puisque ce polymorphisme est rare.

Dépenses budgétaires d'octobre 2010 à septembre 2011

Nous avons recruté une personne en CDD de niveau IE pendant 3 mois, Ludovic Alaux, pour compléter les 15 mois de CDD déjà réalisés en 2009-10 par Isabelle Meusnier.

Publication

Les résultats ont été synthétisés et présentés au Congrès Eucarpia 2011, section Fodder crops & Amenity grasses, sous forme de poster. Un texte de 4 pages a été rédigé, il sera publié sous forme d'ouvrage par l'éditeur Springer (en annexe) :

B. Julier, I. Meusnier, L. Alaux, S. Flajoulot, P. Barre, J. Gouzy. Role of RCT1 gene in anthracnose resistance in alfalfa

Une publication dans une revue à comité de lecture est en préparation.

Perspectives

Il serait judicieux de génotyper un grand nombre d'individus (500 par exemple) sur une variété qui contient une fréquence relativement élevée de l'allèle 122 sur le site 1 de l'exon 1 (Galaxie, par exemple). Il faudrait ensuite réaliser deux polycross, chacun contenant 10 à 30 plantes, soit portant l'allèle 122, si possible en plusieurs doses, soit ne portant pas l'allèle 122. La descendance de ces polycross pourrait alors être testée pour la résistance à l'anthracnose. Si les deux polycross étaient différents en résistance, cela signifierait que la perte de fonction du gène RCT1 induit la résistance. Dans cette situation, les perspectives les plus optimistes du projet auraient été atteintes.

II. Etude du déterminisme génétique de la résistance à la verticilliose (*Verticillium albo-atrum*) chez la légumineuse modèle *Medicago truncatula*.

Enjeux économiques et questions scientifiques

La verticilliose est une maladie vasculaire causée par un champignon pathogène du sol, *Verticillium albo-atrum*, et caractérisée par un flétrissement des parties aériennes des plantes hôtes. En Europe, ce champignon constitue l'une des principales causes de maladies chez la luzerne pérenne et est à l'origine de pertes économiques très importantes. La capacité de *V. albo-atrum* de survivre dans le sol ainsi que sa facilité d'infection par les racines en font un pathogène difficile à combattre, la lutte génétique par sélection de variétés tolérantes apparaissant comme une approche prometteuse.

Démarche scientifique

La démarche repose sur une analyse sans *a priori* par une approche de génétique de transmission. Cette stratégie consiste à analyser des populations de lignées recombinantes (RILs) issues de croisements entre parents sensible et résistant afin de mener une approche QTL visant à identifier les zones du génome impliquées dans l'expression de la tolérance. Ces zones sont ensuite réduites, par analyse de plantes présentant des recombinaisons dans les QTL potentiels. Enfin, la séquence du génome couvrant le QTL est analysée finement pour mettre en évidence des gènes susceptibles de concourir à la tolérance de *M. truncatula* à *V. albo-atrum*.

Acquis des années 1 et 2 (2008-09 et 2009-10)

Afin de pouvoir entreprendre une étude du contrôle génétique de la résistance à *Verticillium albo-atrum* chez la légumineuse modèle *Medicago truncatula*, un nouveau test de résistance à la verticilliose a tout d'abord été mis au point. Le protocole adopté pour l'inoculation des plantes par le pathogène, basé sur le trempage des racines dans une solution de spores (adapté de Vailleau et al, 2007) est apparu facile à mettre en œuvre, robuste, fiable et reproductible au cours des nombreux essais réalisés chez deux des partenaires du projet (Annexe 2). Ce test s'est révélé plus efficace que le test par fauchage utilisé jusqu'alors pour la luzerne. Une échelle pour la notation des symptômes a également été établie.

Cinq des six lignées impliquées dans des croisements à l'origine de populations de RILs déjà cartographiées ont montré une réponse différentielle vis-à-vis du pathogène (Tableau 6).

Croisement	Population de RILs	p-value
A17 (R) X F83005.5 (S)	LR5	$8,7.10^{-7}$ ***
A17 (R) X DZA315.16 (S)	LR4	$5,4.10^{-5}$ ***
F83005.5 (S) X DZA45.5 (R)	LR3	$9,2.10^{-7}$ ***
DZA315.16 X DZA45.5 (R)	LR1	$3,1.10^{-2}$ **
A17 (R) X A20 (S)	LR6	0,06 (sub-significatif)

Tableau 6: Liste des 5 populations de lignées recombinantes dont les lignées parentales montrent une réponse contrastée à la verticilliose à la suite d'une inoculation par fauchage. La p-value représente le niveau de signification du test de comparaison multiple des moyennes. (R): parent résistant, (S): parent sensible.

Les populations de RILs LR5 (A17 X F83005.5) et LR4 (A17 X DZA315.16) issues du croisement entre les lignées les plus significativement différentes quant à leur réponse à *V. albo-atrum* ont été

analysées. Pour chacune de ces populations, le phénotypage correspondant à l'obtention d'une cinétique plante-à-plante de symptômes de flétrissement foliaire au cours du temps, a été effectué dans deux laboratoires différents (INP-ENSAT et Barenbrug). Au total, 160 RILs LR5 en génération F9 ou F10 et 113 RILs LR4 en génération F8 ou F9 ont été analysées à raison d'environ 6 à 8 plantes par lignées.

Une analyse génétique préliminaire a été menée sur les données de phénotypage obtenues à l'ENSAT sur la population LR5. Dans un premier temps, une modélisation de la cinétique de la maladie selon une courbe logistique permettant de décrire l'évolution des symptômes en fonction des 3 paramètres Asym (score maximum de maladie), Xmid (temps nécessaire pour atteindre 50% du score maximum de maladie), et scal (temps nécessaire pour passer de 50% à 75% du score maximum de maladie) a été réalisée. Puis, une recherche de QTL de résistance a été effectuée sur le paramètre Asym, seul paramètre montrant une variation statistiquement significative entre les RILs LR5. Cette première analyse, réalisée à l'aide du logiciel PlabQTL en Composite Interval Mapping avec 9 cofacteurs, a révélé un QTL majeur sur le chromosome 7 avec un LOD score (rapport de vraisemblance) très élevé de 11,63 et expliquant 27,9% de la variation phénotypique. L'intervalle de confiance de ce QTL est de 6 cM. Les marqueurs MTE85 et MTE126 flanquant le QTL sont ancrés sur la carte physique et sont situés sur les BACs AC146554 et AC135311 (Version MtV3.0). Cette zone comprend 31 BACs et 2 régions non séquencées.

Une évaluation phénotypique de la biodiversité naturelle de *M. truncatula* en réponse à *V. albo-atrum* a été réalisée grâce à l'étude d'une collection de 96 lignées (CoreColl 96) pouvant représenter l'ensemble de la variabilité génétique existant au sein de cette espèce modèle (Ronfort et al., 2006). Cette analyse a permis de mettre en évidence une très grande variabilité de réponse des différentes lignées par rapport au champignon pathogène tant en termes de vitesse de déclenchement que de vitesse de progression ou d'index maximum de maladie. Au sein de chacun des 4 sous-groupes identifiés par une étude de structure génétique de population (Ronfort et al., 2006), on compte des lignées présentant différents niveaux de résistance.

Objectifs de l'année 3 (2010-11)

Au cours de cette dernière année du contrat de branche, le premier objectif était de confirmer l'existence d'un QTL majeur sur le chromosome 7 de *Medicago truncatula* contrôlant la résistance à la verticilliose dans la population de RILs LR5. Ce QTL a été identifié à partir d'une expérience effectuée à l'ENSAT et doit être précisé par l'analyse de données phénotypiques obtenues au cours de 2 autres répétitions biologiques indépendantes (expériences réalisées chez Barenbrug et à l'ENSAT).

Un second objectif était de détecter les QTL de résistance pour la population LR4 et de comparer le ou les *loci* identifiés à ceux mis en évidence à partir de la population LR5. Il est à noter que la source de résistance est identique dans les deux croisements, le parent résistant étant la lignée A17 de référence qui a été séquencée (Young et al. 2011).

Un troisième objectif était de tenter de réduire l'intervalle de confiance (IC) du QTL sur le chromosome 7 grâce à une cartographie fine du *locus*.

Il a également été décidé d'étendre l'analyse génétique par transmission à un croisement impliquant une source génétique de résistance différente de celle des populations LR4 et LR5. A cet effet, la population LR3 impliquant le parent résistant DZA45.5 et le parent sensible F83005.5 a été phénotypée par l'entreprise R2n.

Enfin, la caractérisation de la variabilité de tolérance à la verticilliose au sein de la biodiversité naturelle de *Medicago truncatula* a également été poursuivie par une seconde évaluation de la CoreColl96 chez R2n.

Résultats 2010-11

Etude génétique de la résistance à *Verticillium albo-atrum* chez *Medicago truncatula*.

Afin d'identifier des QTL contrôlant la résistance à la verticilliose chez *M. truncatula*, les populations de RILs LR5 (A17 X F83005.5) et LR4 (A17 X DZA315.16) issues du croisement entre les lignées les plus significativement différentes quant à leur réponse à *V. albo-atrum* ont été analysées.

Evaluation phénotypique des RILs LR4 et LR5

Pour chacune de ces populations, trois évaluations phénotypiques indépendantes ont été effectuées: le phénotypage de la population LR4 a été réalisé dans trois laboratoires différents (INP-ENSAT et Barenbrug en 2009-2010 et R2n en 2011); pour la population LR5, 2 répétitions ont été faites à l'ENSAT et une chez Barenbrug. Au total, 160 RILs LR5 en génération F9 ou F10 et 113 RILs LR4 en génération F8 ou F9 ont été analysées à raison d'environ 6 à 8 plantes par lignée.

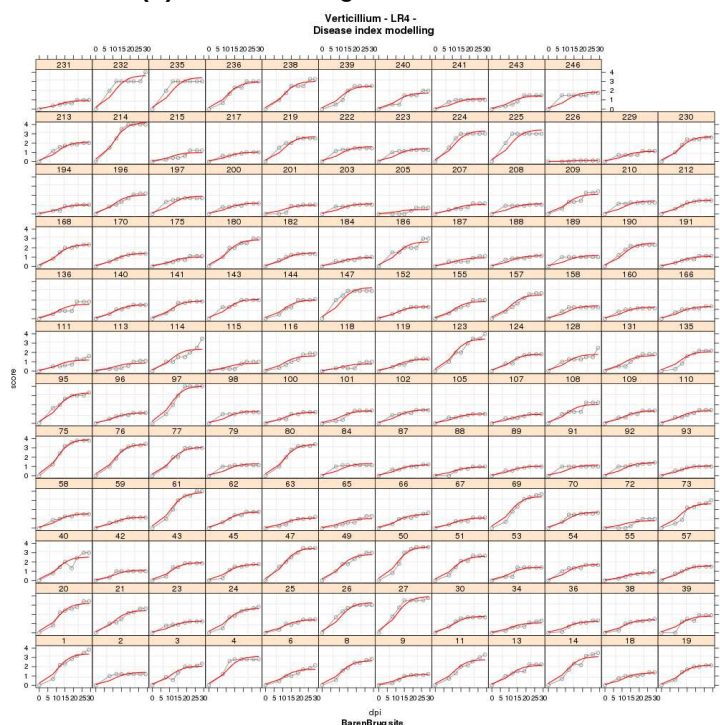
Les plantes ont été inoculées selon le protocole par trempage des racines dans la solution de spores précédemment établi (Annexe 2). Le suivi des symptômes a été réalisé tous les deux jours entre 0 et 27 jours après inoculation selon une échelle de notation allant de 0 à 4 où un score de 0 correspond aux plantes qui ne présentent pas de symptômes, un score de 4 aux plantes mortes.

Modélisation de la cinétique de maladie des RILs LR4 et LR5 par courbe logistique

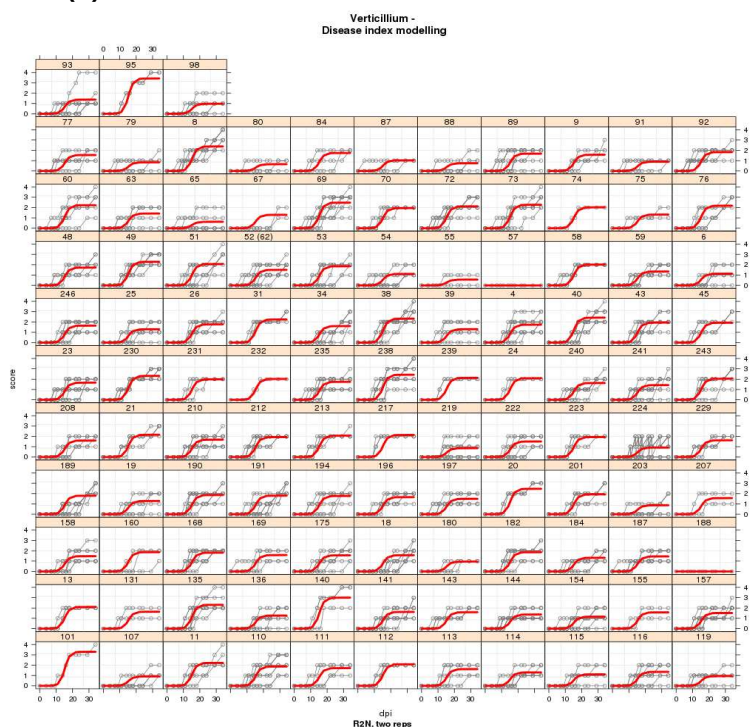
La cinétique de la maladie a été modélisée selon une courbe logistique pour chacune des populations de RILs et chacune des répétitions biologiques (ENSAT, R2n, Barenbrug). La modélisation de la courbe de progression de la maladie (Gilligan, 1990) a été réalisée en utilisant des modèles mixtes non linéaires (Lindstrom and Bates, 1990) tels qu'implémentés dans le package nlme du logiciel R (<http://cran.r-project.org>).

Figure 2: Modélisation par courbe logistique de la cinétique de maladie des RILs LR4 pour les expériences réalisées chez Barenbrug (A) et R2n (B) et des RILs LR4 chez Barenbrug (C) et à l'ENSAT (D).

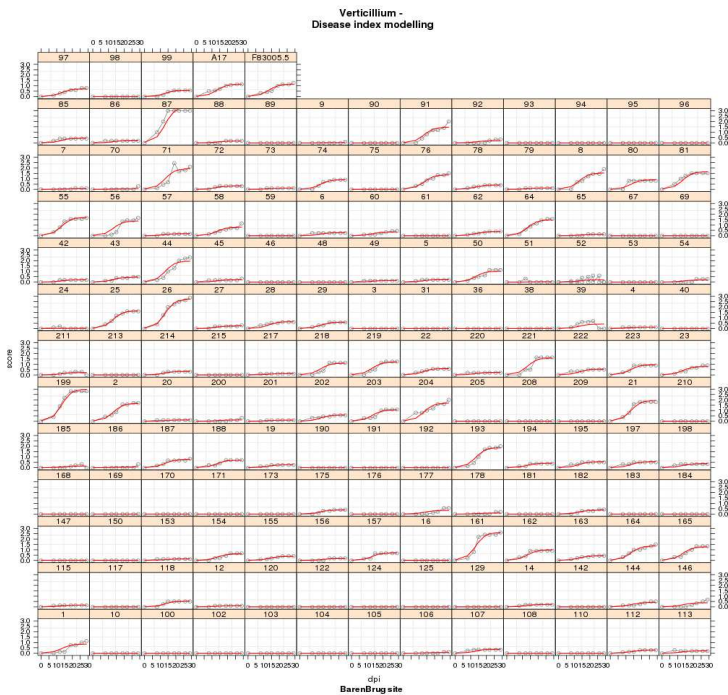
(A) LR4 Barenbrug



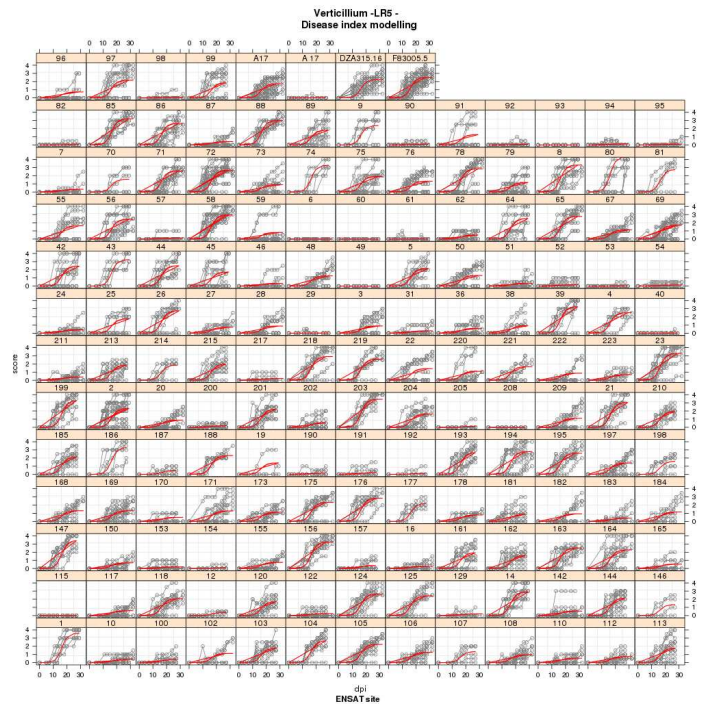
(B) LR4 R2n



(C) LR5 Barenbrug



(D) LR5 ENSAT



Quelle que soit la population de RILs considérée, ces analyses révèlent des profils de cinétique très différents pour les données obtenues chez Barenbrug comparées à celles de l'ENSAT et R2n (figure 2). La pression d'infection semble avoir été moins forte pour les expériences de Barenbrug, menant à un nombre de plantes malades et une gravité des symptômes moindres que dans les autres répétitions biologiques. Ce problème semble lié aux conditions de préparation de l'inoculum de spores qui n'était pas préparé au moment même de l'inoculation mais plusieurs jours avant, ce qui entraîne une perte de virulence du pathogène. Les données de phénotypage obtenues chez Barenbrug ont par la suite été retirées de l'analyse génétique.

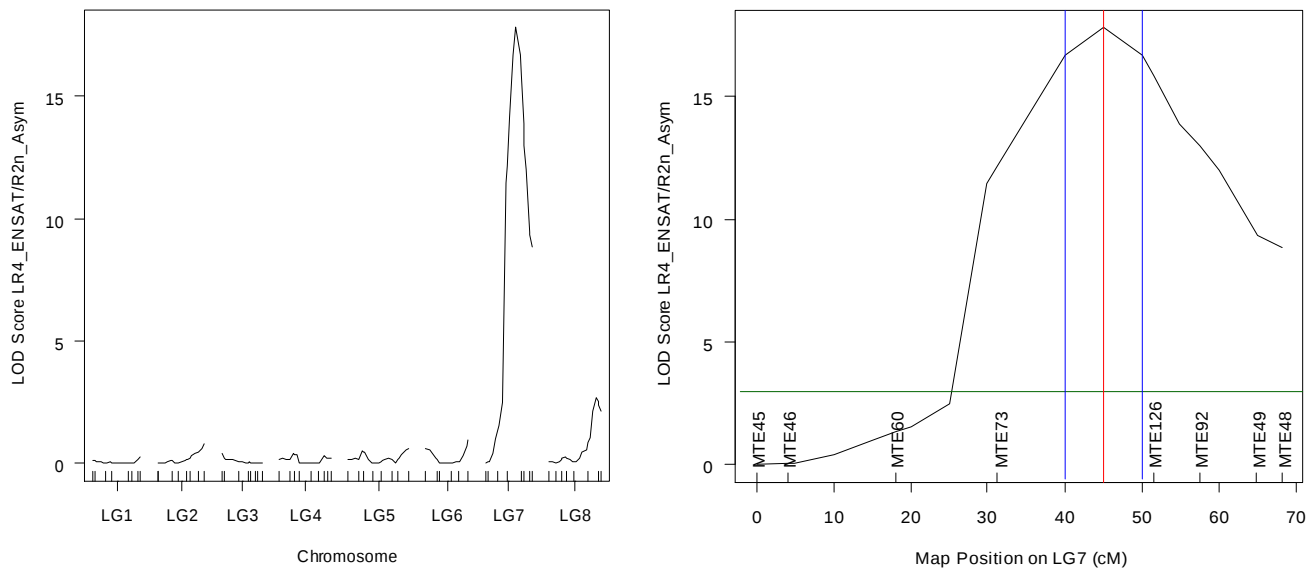
Une analyse statistique par ANOVA et comparaison multiple de moyennes effectuée sur les différents paramètres calculés par la modélisation logistique a permis d'identifier le paramètre Asym (index maximum de maladie) comme étant le seul paramètre montrant une variation significative entre les RILs pour chacune des populations de RILs étudiées et quelle que soit la répétition biologique considérée (ENSAT ou R2n).

Détection d'un QTL majeur de résistance sur le chromosome 7 pour LR4 et LR5.

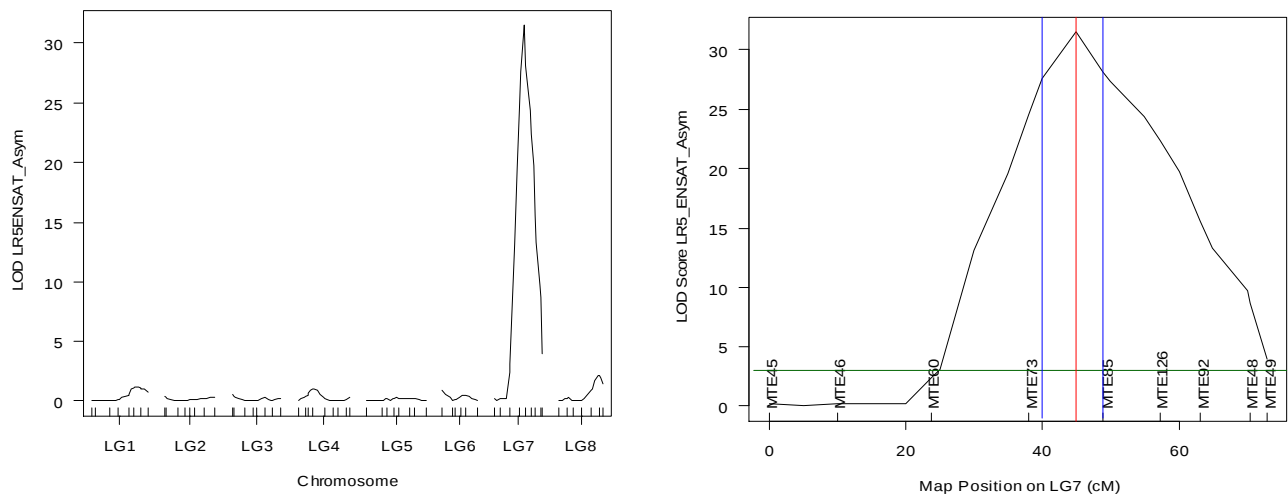
Une recherche de QTL de résistance pour les populations de RILs LR4 et LR5 a été effectuée sur le paramètre Asym, seul paramètre montrant une variation statistiquement significative entre les RILs. Chez la légumineuse modèle, l'existence d'une carte génétique consensus utilisant des marqueurs cadres communs aux cartes génétiques des populations LR4 et LR5 permet une comparaison directe des QTLs identifiés suite à l'étude de ces deux populations. Cette analyse génétique, réalisée à l'aide du package qtl du logiciel R en Multiple QTL Mapping (MQM) avec un cofacteur sur le LG7, a révélé pour les 2 populations de RILs considérées, un unique QTL majeur sur le chromosome 7 avec un LOD score (rapport de vraisemblance) très élevé (de l'ordre de 31.437 pour LR5 et de 17.794 pour LR4) (figure 3). Ce locus explique respectivement 29% et 44% de la variation phénotypique existant au sein des RILs LR4 et LR5.

Figure 3: Un QTL majeur de résistance à la verticilliose sur le LG7 identifié à partir du paramètre Asym calculé pour les populations de RILs LR4 (A) et LR5 (B) (cartes génétiques cadre CGPM).

(A) LR4 (analyse combinée sur les phénotypes obtenus à l'ENSAT et chez R2n).



(B) LR5 ((analyse combinée sur les phénotypes obtenus au cours des 2 répétitions faites à l'ENSAT).



L'intervalle de confiance de ce QTL (LOD Score - 1) est de 10 cM pour LR4, les marqueurs les plus proches étant MTE73 et MTE126. Pour LR5, l'IC est de 6cM est les marqueurs flanquants sont MTE85 et MTE126.

Le même QTL majeur de résistance sur le chromosome 7 a également été identifié suite à la détection de QTL réalisée sur les valeurs d'AUDPC (Area Under Disease Progression Curve, Shaner and Finney, 1977) des RILs LR4 et LR5.

Vers la réduction de l'intervalle de confiance (IC) du QTL majeur sur le chromosome 7.

La réduction de l'intervalle de confiance du QTL majeur identifié précédemment nécessite le phénotypage et le génotypage fins des RILs recombinant dans la zone d'intérêt. Vingt-trois lignées de ce type ont été identifiées dans la population LR5. Le génotype de 3 autres lignées à l'un des marqueurs flanquants reste à vérifier. De plus, parmi celles-ci, 4 lignées contenant de l'hétérozygotie résiduelle au niveau du QTL ont été identifiées et multipliées en serre dans l'optique de dériver des HIFs (Heterogeneous Inbred Lines).

Dans un premier temps, le phénotype des RILs recombinant sous le QTL a été validé au cours d'une nouvelle inoculation (figure 4 et Annexe 3). Cette expérience a été réalisée à partir d'un nombre plus élevé de plantes pour chaque lignée (~10 plantes/lignée). Les 10 RILs apparues comme étant les plus sensibles et les plus résistantes au cours de la première expérience de phénotypage de la population LR5 ont également été re-testées en tant que contrôles.

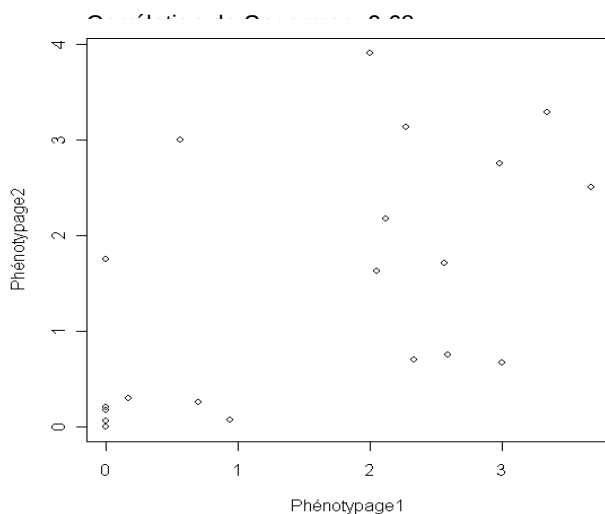


Figure 4: Corrélation du paramètre Asym (index maximum de maladie) des RILs LR5 recombinant sous le QTL majeur sur le LG7 entre deux évaluations phénotypiques indépendantes.

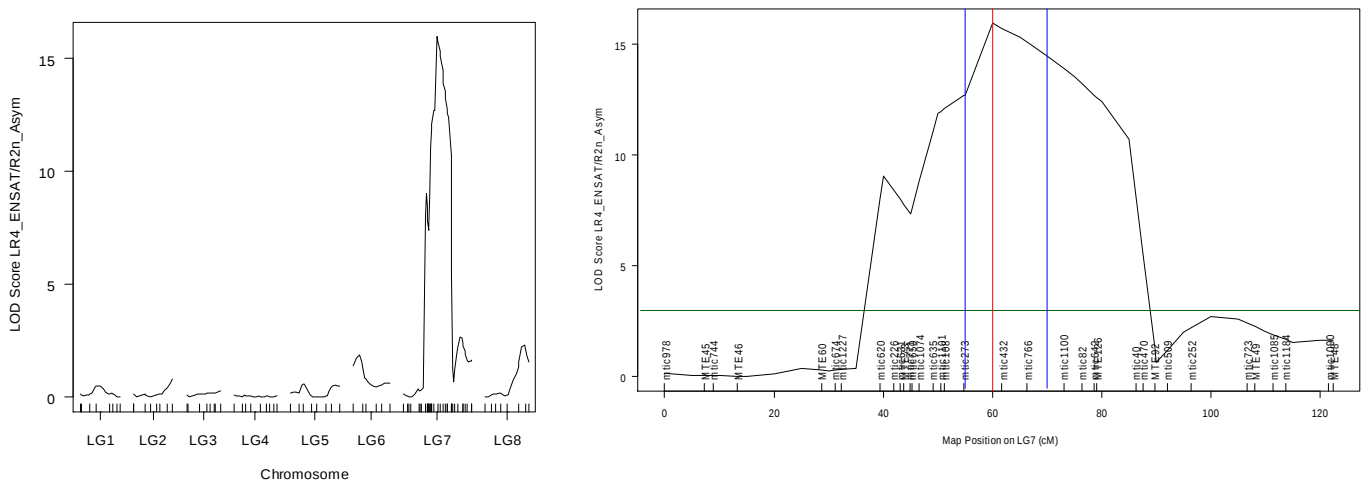
Afin de densifier la cartographie génétique sous le QTL majeur, 13 nouveaux marqueurs SSR régulièrement espacés dans la zone d'intérêt ont été désignés en s'appuyant sur la séquence génomique de la lignée de référence A17. Parmi ceux-ci, cinq apparus polymorphes et co-dominants entre les lignées parentales A17 et F83005.5 de la population LR5 ont été utilisés pour génotyper les RILs recombinant sous le QTL afin de préciser les points de recombinaison (Annexe 3). La cartographie fine du *locus* devra être poursuivie en génotypant sur toutes les RILs LR5 de nouveaux marqueurs désignés à intervalles réguliers sous le QTL afin de densifier la carte dans la région génomique d'intérêt. Une nouvelle détection de QTL pourra alors être réalisée sur la carte dense du LG7 ainsi obtenue dans le but de réduire l'IC du QTL.

De nouveaux recombinants dans la zone d'intérêt pourront également être générés à partir des 4 RILs LR5 encore hétérozygotes sous le QTL. L'intervalle de confiance du QTL étant de 6 cM, cette étape nécessitera le criblage génotypique de 500 à 1000 individus avec 2 ou 3 marqueurs microsatellites régulièrement répartis le long de la zone d'intérêt. Une fois identifiés, ces nouveaux recombinants pourront être finement phénotypés et génotypés afin de réduire encore plus l'IC du QTL.

Enfin, la même démarche pourra être initiée pour la population LR4 (identification de RILs recombinant ou encore hétérozygotes sous le QTL, création de nouveaux recombinants, génotypage fin du *locus* avec de nouveaux marqueurs). Parmi les 13 nouveaux marqueurs SSR désignés précédemment dans la région génomique d'intérêt, sept sont apparus polymorphes et co-dominants entre les 2 lignées parentales A17 et DZA315.16 de la population LR4. Si nécessaire, ils pourront donc être génotypés sur l'ensemble des RILs de cette population dans le but de densifier la carte génétique. Toutefois, pour cette population, nous disposons déjà d'une carte densifiée pour le LG7 contenant 10 marqueurs situés entre les marqueurs flanquants MTE73 et MTE126. Une nouvelle détection de QTL effectuée sur la carte dense du LG7 nous a permis d'affiner la localisation du QTL

en réduisant l'intervalle de confiance, les marqueurs les plus proches du pic étant désormais les SSRs mtic273 et mtic 1100 (figure 5).

Figure 5: Réduction de l'IC du QTL majeur de résistance à la verticilliose par analyse d'une carte dense du LG7 de la population LR4.



Annotation fonctionnelle des gènes prédits dans les régions génomiques d'intérêt

Chez la légumineuse modèle, il est possible de faire un lien direct entre la carte génétique et la carte physique, la plupart des marqueurs génotypés sur les populations de RILs étant cartographiés sur le génome de la lignée de référence A17.

Toutefois, dans notre cas, seuls 2 marqueurs flanquants le QTL ont pu être localisés sur le génome. Ainsi, le marqueur MTE85 flanquant gauche du QTL pour la population LR5 est situé sur le BAC AC146554 (Version MtV3.5). Le marqueur flanquant droit MTE126 n'a pas pu être ancré sur un BAC du chromosome 7. De même, le marqueur de la carte LG7 dense mtic273 flanquant le QTL LR4 à gauche est situé sur le BAC AC234395 alors qu'à ce jour, le marqueur de la carte LG7 dense mtic1100 flanquant le QTL à droite n'est pas encore ancré sur le génome.

Etant donné les informations parcellaires concernant la carte physique de la région d'intérêt actuellement à notre disposition, il nous est pour l'instant impossible d'apporter des données plus détaillées sur les BACs que couvre le QTL et la fonction potentielle des gènes prédits dans la zone.

La réduction de l'IC du QTL à partir des ressources disponibles (lignées recombinant sous le QTL, nouveaux marqueurs dans la zone) et des données combinées pour les 2 populations de RILs LR4 et LR5 doit être poursuivie afin d'affiner encore la zone d'intérêt et de s'assurer que l'on dispose bien de la séquence pour pouvoir ensuite envisager raisonnablement une annotation fonctionnelle détaillée des gènes prédits dans la zone. Cette analyse permettra à terme de vérifier l'existence ou non de RGA (Resistance Gene Analog) dans la région ciblée et d'éditer une liste de gènes candidats potentiels pouvant conférer la résistance. Nous pouvons tout de même noter que ce QTL majeur n'est pas situé dans l'une des régions génomiques riches en RGA identifiées *in silico* par Ameline-Torregrosa et al. (2008a).

Analyse phénotypique de la population de RILs LR3

Une nouvelle analyse génétique visant à terme la détection de nouveaux QTL de résistance a été initiée à partir de la population de RILs LR3 issue du croisement entre la lignée résistante DZA45.5 et la lignée sensible F83005.5. Ceci pourra peut-être permettre de mettre en évidence un contrôle génétique de la résistance à *V. albo-atrum* différent de celui déjà identifié par l'analyse des populations LR4 et LR5 pour lesquelles le parent résistant est la lignée de référence A17.

Une évaluation phénotypique des RILs LR3 a été réalisée, dans ce but, au sein de l'entreprise R2n selon le protocole d'inoculation racinaire. Afin de déterminer la tolérance/sensibilité des RILs, le

pourcentage de plantes mortes pour chaque lignée a été noté 33 jours après inoculation (dpi) en 2 blocs (figure 6).

L'analyse statistique des données selon un modèle linéaire généralisé et l'analyse de la table de déviance révèle un effet génotype très net sur le pourcentage de mortalité ($\text{Pr}(> F) < 2\text{e-}16$ ***) et pas d'effet bloc.

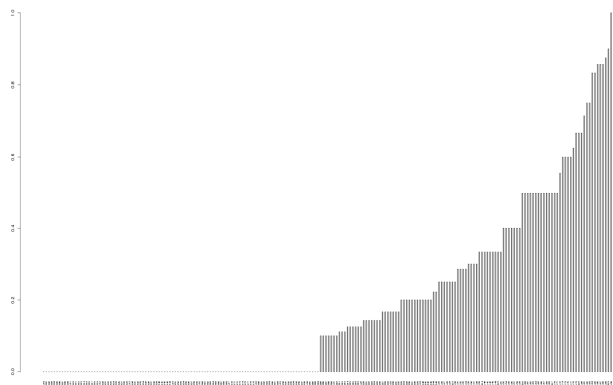


Figure 6: Pourcentage de plantes mortes à 33 dpi pour chaque RILs de la population LR3 corrigé de l'effet bloc.

La détection de QTL pourra être réalisée par l'ENSAT dès que les données de cartographie génétique nécessaires à cette analyse auront été fournies par l'INRA. Les résultats seront ensuite communiqués aux partenaires de ce Contrat de Branches 'Résistance aux bio-agresseurs'.

Méta-analyse

Cette étape du projet consiste à comparer l'ensemble des résultats des analyses de QTL de résistance à *V. albo-atrum* réalisées au cours de ce projet sur les 3 populations de RILs LR3, LR4 et LR5. Les croisements LR4 et LR5 impliquant le même parent résistant A17, nous pouvons supposer que le QTL majeur de résistance identifié sur le chromosome 7 pour ces deux populations correspond à la même source génétique de résistance et donc au(x) même(s) gène(s). Les données bibliographiques concernant d'autres pathosystèmes impliquant la légumineuse modèle *M. truncatula* et utilisant les mêmes croisements ont également été comparées aux résultats de notre analyse. Le QTL majeur de résistance à la verticilliose identifié suite à notre étude des populations de RILs LR4 et LR5 ne colocalise avec aucun des QTLs de résistance identifiés jusqu'ici sur les mêmes populations mais dans d'autres interactions racinaires de type fongique (*Aphanomyces euteiches*, Djébali et al., 2009) ou bactérien (*Ralstonia solanacearum*, Vaillau et al., 2007), ou encore dans des interactions avec des pathogènes foliaires (*Colletotrichum trifolii*, *Erysiphe pisi*, Ameline-Torregrosa et al, 2008b).

Les résultats concernant la population LR3 sont quant à eux en attente. La même analyse de comparaison et de synthèse sera réalisée dès lors que la détection de QTL aura été faite. Si nécessaire, le logiciel MCQTL adapté pour ce type d'analyse pourra être utilisé (Jourjon et al, 2005).

Analyse de la variabilité naturelle de la tolérance à la verticilliose chez *Medicago truncatula*.

Une évaluation phénotypique de la biodiversité naturelle de *M. truncatula* en réponse à *V. albo-atrum* a été envisagée grâce à l'étude d'une collection de 96 lignées (CoreColl96) pouvant représenter l'ensemble de la variabilité génétique existant au sein de cette espèce modèle (Ronfort et al., 2006). Une grande partie de ces lignées sont également des lignées parentales de population de RILs.

La variabilité de la tolérance à la verticilliose existant au sein des lignées de la Core Collection 96 a été évaluée au cours de deux expériences indépendantes réalisées à l'ENSAT et chez R2n.

La modélisation des courbes de maladie a permis de mettre en évidence une très grande variabilité de réponse des différentes lignées par rapport au champignon pathogène tant en termes de vitesse de

déclenchement que de vitesse de progression ou d'index maximum de maladie (figure 7). Ces résultats peuvent laisser supposer l'existence de différents contrôles génétiques de la résistance à la verticilliose chez la légumineuse modèle.

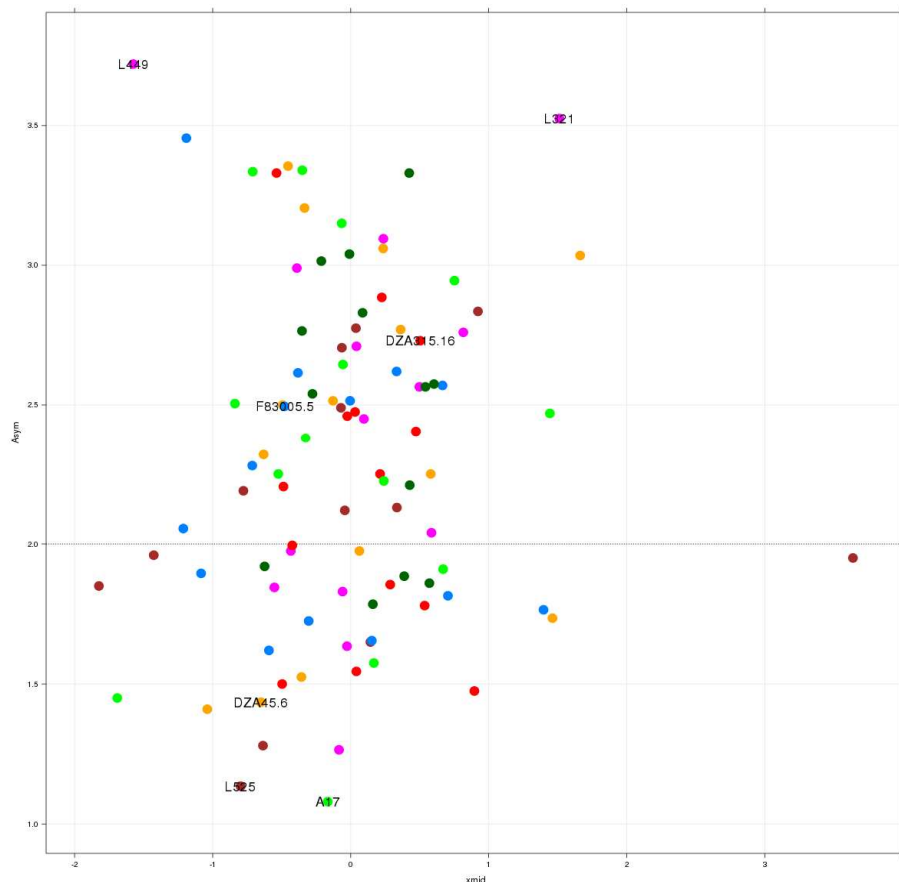


Figure 7. Niveau de tolérance des lignées de la CoreColl96 de *Medicago truncatula* représenté en fonction des paramètres de modélisation de la courbe de maladie *Asym* (Index maximum de maladie) et *xmid* (Temps nécessaire pour atteindre 50% d'*Asym*).

Comme attendu, les lignées parentales A17 et DZA45.6 apparaissent comme étant résistantes et F83005.5 et DZA315.16 sensibles. Cette évaluation d'un plus grand nombre de lignées a également permis d'identifier de nouvelles lignées montrant des réponses encore plus contrastées vis-à-vis de *V. albo-atrum*. La lignée L449 apparaît par exemple comme étant la plus sensible (développement très rapide et très intense des symptômes de flétrissement). A l'avenir, il pourrait donc être intéressant de tester la descendance de nouveaux croisements entre les lignées les plus différentielles. Cependant, compte-tenu de l'effet significatif du site (ENSAT ou R2n) sur les paramètres *Asym* et *xmid* révélé par les analyses statistiques des données, il serait judicieux d'effectuer, préalablement à une nouvelle analyse lourde de détection de QTL, une troisième évaluation phénotypique indépendante des lignées de la CoreColl96 à la verticilliose pour garantir la robustesse des données de résistance/sensibilité des lignées parentales choisies.

Les résultats préliminaires sur les données des deux premières expériences réalisées à l'ENSAT et chez R2n ont tout de même permis d'identifier, par régression jointe, 4 lignées 'typiques':

- L449, stablement très sensible
- L525, stablement très résistante
- L738 (A17), très résistante, d'autant plus résistante que la condition est plus favorable à la maladie (donc interaction)
- L513, 'le' spécimen type de l'interaction : sensible dans un site, résistante dans l'autre.

De plus, au sein de chacun des 4 sous-groupes identifiés par une étude de structure génétique de population (Ronfort et al., 2006), on compte des lignées présentant différents niveaux de résistance. Ce résultat nous permet d'envisager de réaliser dans le futur une étude de génétique d'association sans avoir à tenir compte de la structure génétique existant au sein de l'espèce.

CONCLUSION – BILAN DES RESULTATS DELIVRES

Les travaux portant sur l'étude génétique de la résistance à la verticilliose chez *Medicago truncatula* menés au cours du contrat de branche 'Résistance aux bioagresseurs' auront permis de:

- Développer un test précoce de résistance à la verticilliose chez la légumineuse modèle fiable, robuste, facile à mettre en œuvre. Ce protocole a d'ores et déjà été transféré avec succès à la luzerne cultivée *Medicago sativa* par les partenaires privés impliqués dans ce projet.
- Identifier 4 populations de RILs potentiellement intéressantes pour mener une analyse génétique visant à détecter les QTL impliqués dans le contrôle de la résistance à la verticilliose. Il est à noter que les lignées parentales de ces croisements peuvent correspondre à différentes sources génétiques de résistance.
- Détecter un QTL majeur de résistance à la verticilliose situé sur le chromosome 7 du génome de *Medicago truncatula* conférant la résistance aux RILs de 2 populations indépendantes et expliquant jusqu'à 44% de la variation phénotypique.
- Mettre en évidence une large variabilité de réponse à la verticilliose au sein de la biodiversité de la légumineuse modèle. Combinés aux données de re-séquençage des lignées de la CoreColl96 obtenues dans le cadre du projet NSF MtHapMap, ces résultats phénotypiques pourront être à l'avenir utilisés dans une approche de génétique d'association visant à identifier de manière exhaustive l'ensemble des régions génomiques impliquées dans la résistance à la verticilliose.

Communications et prévision de publication

Les résultats obtenus sur la caractérisation du pathosystème *M. truncatula*/ *V. albo-atrum* ont fait l'objet de communications sous forme de posters à l'occasion du Meeting 'From foliar to root-interacting pathogens and symbionts' en février 2011 (annexe 4) et du 'Model Legume Congress' organisé par le laboratoire SP2 de l'INP-ENSAT en mai 2011 (annexe 5).

De plus, une publication scientifique en cours de rédaction clôturera ce programme. Cet article devrait être soumis d'ici la fin de l'année 2011 au MPMI journal (Molecular Plant-Microbe Interaction), journal scientifique à comité de lecture international.

Ben C., Montanari S., Toueni M., Amatya S., Negahi A., Gras M.C. Noël D., Gentzbittel L., Rickauer M. (2011). Genetic traits and regulatory mechanisms involved in resistance of *Medicago truncatula* to *Verticillium albo-atrum*. Model Legume Congress, Sainte-Maxime, France, 15-19 Mai 2011, p 39 (presentation orale)

Ben C., Montanari S., Amatya S., Gras M.C. Noël D., Gentzbittel L., Rickauer M. (2011). Genetic traits and regulatory mechanisms involved in resistance **of the model legume plant *Medicago truncatula* towards *Verticillium albo-atrum***. Meeting 'From foliar to root-interacting pathogens and symbionts'. Goettingen, février 2011 (poster)

Références bibliographiques

Ameline-Torregrosa C, Wang BB, O'Bleness MS, Deshpande S, Zhu H, Roe B, Young ND, Cannon SB, [Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*](#). Plant Physiol. (2008) 146(1):5-21.

- Ameline-Torregrosa C, Cazaux M, Danesh D, Chardon F, Cannon SB, Esquerré-Tugayé MT, Dumas B, Young ND, Samac DA, Huguet T, Jacquet C., [Genetic dissection of resistance to anthracnose and powdery mildew in *Medicago truncatula*](#). *Mol Plant Microbe Interact.* (2008) 21(1):61-9.
- Djébali N, Jauneau A, Ameline-Torregrosa C, Chardon F, Jaulneau V, Mathé C, Bottin A, Cazaux M, Pilet-Nayel ML, Baranger A, Aouani ME, Esquerré-Tugayé MT, Dumas B, Huguet T, Jacquet C., [Partial resistance of *Medicago truncatula* to *Aphanomyces euteiches* is associated with protection of the root stele and is controlled by a major QTL rich in proteasome-related genes.](#) *Mol Plant Microbe Interact.* (2009) 22(9):1043-55.
- Gilligan CA (1990) Comparison of disease progress curves. *New Phytol.* 115, 223-242
- [Jourjon MF](#), [Jasson S](#), [Marcel J](#), [Ngom B](#), [Mangin B](#). *MCQTL: multi-allelic QTL mapping in multi-cross design. Bioinformatics (2005) 21 (1): 128-130.*
- Lindstrom ML, Bates DM. (1990) Nonlinear mixed effects models for repeated measures data *Biometrics.* 46(3):673-87.
- Shaner G, Finney RE (1977). The effect of nitrogen fertilization on the expression of slow-mildewing resistance in Knox wheat. *Phytopathology* 67:1051-1056.

Role of RCT1 gene in anthracnose resistance in alfalfa

B. Julier¹, I. Meusnier^{1,4}, L. Alaux¹, S. Flajoulot², P. Barre¹, J. Gouzy³

¹ INRA, UR 4, Unité de Recherche Pluridisciplinaire Prairies et Plantes Fourragères, Le Chêne, RD 150, BP 80006, 86600 Lusignan, France

² Jouffray-Drillaud, La Litière, 86600 Saint-Sauvant, France

³ INRA-CNRS, Laboratoire des Interactions Plantes-Microorganismes, 31326 Castanet Tolosan, France

⁴ INRA, UMR-Centre de Biologie pour la Gestion des Populations, Campus international de Baillarguet - CS 30 016, 34988 Montpellier sur Lez Cedex, France (present address)

Abstract Anthracnose (*Colletotrichum trifolii*) is a severe disease in alfalfa (*Medicago sativa*). The RCT1 gene, isolated from the model legume *M. truncatula*, is a candidate gene to explain genetic variation for anthracnose resistance in alfalfa. A bulk segregant analysis was carried out to test this hypothesis: for eight alfalfa varieties, 15 resistant (R) plants and 15 susceptible (S) plants were selected and DNA was extracted. The whole gene including the upstream and downstream regions (a total of 14 kb) was amplified by PCR for each individual and the 15 R plants (accordingly S plants) were pooled for each variety. Sequencing was carried out using the next generation sequencer 454 of Roche. The sequence reads, that averaged 295 bp, were assembled to produce consensus sequences that can be considered as alleles. Considering the five exons of the gene, five regions contained clear deletion/insertion polymorphism but these polymorphisms were present in both the pools R and S. Individual genotyping for these indels indicated that different alleles were present but no allele was related to the phenotype. These polymorphic regions in RCT1 seemed not to explain the variation of anthracnose resistance in alfalfa.

Keywords *Medicago sativa*, *Medicago truncatula*, *Colletotrichum trifolii*, candidate gene, bulk segregant analysis, next-generation sequencing

1. Introduction

Anthracnose (*Colletotrichum trifolii*) causes severe yield losses in alfalfa (*Medicago sativa*) forage production (Barnes et al. 1969; Raynal et al. 1989). This fungus attacks stems and crowns and kills individual plants, contributing to stand decline. Three races have been described (Ariss and Vandemark 2007). Alfalfa resistance to anthracnose is conferred by two independent dominant genes, *An1* and *An2* (Elgin Jr and Ostazeski 1985). Resistance to race 1 and likely race 4 is conferred by *An1* (Mackie et al. 2003), and to race 2 by *An2*. Breeding has focused on the selection of resistant varieties, using tests in controlled conditions based on artificial inoculations (Gondran 1984; O'Neill 1991). Despite an oligogenic inheritance, quantitative methods are used to select resistant plants, specifically because of the autotetraploidy of alfalfa makes the genetic analyses more complex. QTL studies were conducted to localise the resistance genes on the genome, both on alfalfa (Mackie et al. 2007) and on the model plant *M. truncatula* (Ameline-Torregrosa et al. 2008; Yang et al. 2007) in which susceptible and resistant plants were identified. Partially consistent results were obtained with QTLs found on chromosome 4 but also a QTL with strong effect on chromosome 8 in alfalfa.

On *M. truncatula*, the diploidy and genomics tools give the opportunity to conduct fine mapping strategies aiming at discovering the gene responsible of a QTL. Fine mapping analysis on a large population resulted in the identification of RCT1 gene, a TIR-NBS-LRR resistance gene (Yang et al. 2008). The function of this gene was not validated on *M. truncatula* because of the difficulty to transform some genotypes. But RCT1 was introduced as a transgene in susceptible alfalfa genotypes,

and conferred resistance to anthracnose (Yang et al. 2008). Transgenic varieties are not agreed by European society, and gene flow risks are high in alfalfa because of its allogamy, inducing transgene escapes to cultivated or sub-spontaneous populations. Thus, the use of RCT1 to breed resistant GMO cultivars is not adapted to European market.

Given the results on the gene RCT1 by Yang et al. (2008), we hypothesized that the heterologous of this gene in alfalfa could be responsible for anthracnose resistance in this species. The objective of this study was to test the role of RCT1 sequence polymorphism in alfalfa resistance to anthracnose. The polymorphisms related to resistance could be used to develop markers available to breeders to select for plants carrying high doses of resistant alleles. The relying hypothesis is that a gene identified to explain trait variation in the model species would explain the variation explained in alfalfa for the same trait (Julier and Meusnier 2010).

2. Material and methods

RCT1 gene in *M. truncatula* spans more than 3.5 kb. Upstream and downstream regions as well as intronic sequences showed polymorphisms that could be related to resistance expression (Yang et al. 2008). As a consequence, we sequenced the whole gene, including non-coding regions. Because of the low sequence conservation in non-coding sequences but the high sequence conservation in coding regions between alfalfa and *M. truncatula*, we defined primers in the adjacent genes of RCT1, in *M. truncatula*.

The heterozygosity of alfalfa implies to sequence the four alleles of the gene. Direct sequencing of small gene portions is feasible (Pierre et al. 2011) but time-consuming and expensive. The Next Generation Sequencers offers the possibility to sequence a large set of genomic sequences at a reasonable cost. This method was used to get the complete gene sequence of RCT1, with its four alleles, in a set of genotypes.

Eight varieties with different resistance level were chosen and anthracnose tests were carried out on 100 seedlings of each variety with the strain C86-2 of *C. trifolii*. For each variety, 15 resistant (R) plants and 15 susceptible (S) plants were selected, DNA was extracted. Two primer pairs were defined to amplify the whole gene and the upstream and downstream regions (a total of about 14 kb). A bulk segregant analysis (BSA) was carried out with the objective to identify gene polymorphism that was associated to resistance (Michelmore et al. 1991). The 454 (Roche) sequencing technology was used. The PCR products of the 15 R plants (accordingly S plants) were pooled for each variety. DNA libraries were tagged for each variety and the libraries were then pooled. Up to 60 different alleles could be present in each pool. The full RCT1 gene was cloned and sequenced on one alfalfa genotype with Sanger sequencing method to produce a reference sequence.

The raw sequences were assembled based on the reference sequence to produce consensus sequences that could be considered as alleles. As a very large sequence polymorphism was present, the rest of analysis was restricted to the five exons of the gene. These consensus sequences were translated into proteins and R and S sequences were compared. In the regions showing deletions, primer pairs were defined to genotype the individuals of the R and S pools. The relationship between genotype and phenotype was analysed.

3. Results and discussion

Expected results: Resistance is dominant over susceptibility. It means that susceptible genotypes only carry alleles of susceptibility (S), their genotype is SSSS. By contrast, the resistant genotypes contain at least one allele of resistance (R) but may carry alleles of susceptibility: they may be RRRR, RRRS, RRSS or RSSS. This notation with alleles R and S is schematic because we do not know how many alleles of susceptibility or resistance could exist. If susceptibility corresponds to inactive allele, several alleles of susceptibility will be observed due to accumulation of mutations. To summarize, the S pools should only contain alleles of susceptibility, except if a resistant individual has been misclassified as susceptible. The R pools should contain both alleles of resistance and susceptibility, their proportion depending on the frequency of resistance alleles in the variety (Julier et al. 2004).

A last point could modify the analysis: the alleles of resistance or susceptibility could be different in the varieties. These varieties were bred by different breeders who use partly different genetic

backgrounds. However, the low level of genetic differentiation among European varieties is favorable to the hypothesis that the genetic determinant of resistance is the same in all varieties.

Resistance test: The 8 varieties were evaluated for their resistance. As expected, Marshall was the most resistant variety. Kali that was chosen for its low level of resistance had an unexpected 42% of resistant plants. The other six varieties, known as of intermediate resistance level, had among 20 and 47% of resistant plants. In all varieties, 15 resistant plants and 15 susceptible plants were chosen, except in Marshall in which only 7 susceptible plants were identified.

Table 1 Evaluation of the level of resistance of the 8 alfalfa cultivars to anthracnose, expressed as the percentage of resistant plants.

Variety	% of resistant plants
Galaxie	20.0
Bar 2	27.0
Prunelle	28.6
Kali	41.8
Canelle	42.0
Symphonie	44.9
Bar 1	47.4
Marshall	83.9

RCT1 sequencing: The whole gene and the upstream and downstream regions was cloned and sequenced for a single individual with Sanger method. It gave a 14.5 kb sequence, with the same structure in 5 exons as on *M. truncatula*. Among this individual, polymorphism was identified, with some long insertions/deletions. A reference sequence was built and exons were identified.

With the NGS method, the whole gene and the upstream and downstream regions were sequenced for 240 individuals (potentially 960 alleles). A total of 570 819 sequences was obtained with a mean size of 295 bp. A reference sequence was built and exons were identified. Raw sequences were assembled based on the reference sequence to produce consensus sequences that can be considered as alleles. As a very large sequence polymorphism was present, the rest of analysis was restricted to the five exons of the gene. Five regions contained clear deletion/insertion polymorphism: one in the ATG region of exon 1, two in exon 3 and two in exon 5 (Fig. 1).

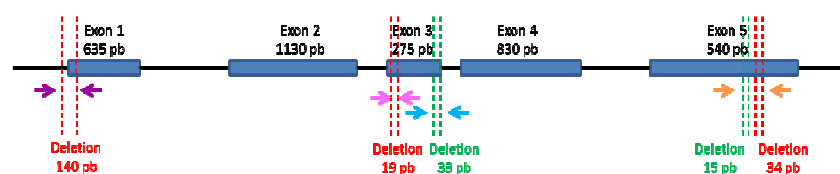


Fig. 1. Deletions identified in the exons of RCT1 gene in alfalfa. In red, the deletions induce a reading frameshift. In green, they do not produce a reading frame shift. The arrows indicate the position of the primers defined to amplify the gene portions.

In all pools, both deleted and non-deleted alleles were found, including in the S pools. As these alleles corresponded to sequence assemblies, this information was not sufficient to conclude, the misclassification of a single R individual in a S pool could induce this situation. However, this case should be rare because only the plants showing clear phenotypes were used to build the pools.

Polymorphism at the individual level: For the deletions on exon 1, beginning and end of exon 3, primers were defined to amplify a portion of less than 300 bp (Figure 1). As the two deletions of exon 5 were close, a single primer pair was defined. Polymorphism was evidenced for each region with 2 to 5 different alleles. One allele per region was highly frequent, whatever the variety or the R or S pool. The difference in allele frequency between R and S plants was never significant for the frequent

alleles. For the deletion of exon 1 that was the most promising because of ATG deletion, the deleted allele (102 bp) had a too low frequency to explain the polymorphism for resistance (Table 2) even if the χ^2 test reached $P = 0.06$.

Table 2 Number of alleles of 102, 241 et 244 bp in the polymorphic region of exon 1.

Phenotype	Number of individuals	Allele 102	Allele 241	Allele 244	Missing data
R	130	28	405	61	6
S	126	13	409	54	7

4. Conclusion

The polymorphic regions of RCT1 exons identified after the sequencing of the whole gene with a bulk-segregant analysis seemed not to explain the variation of anthracnose resistance in alfalfa. Either non coding sequences of this gene explain the resistance or another gene is involved. In this specific case, the gene detected in *M. truncatula* for anthracnose resistance may not be the one responsible to the same trait in alfalfa. Indeed, another QTL study indicated that the QTL position for anthracnose resistance was outside the position of RCT1 (Cazaux 2008). This results is less positive than a previous one on a CONSTANS-like gene involved in flowering date and stem elongation (Herrmann et al. 2010) that showed that translational genetics may be efficient. However, this study shows the power of next-generation sequencing technologies on the genome analysis of a polyploid and heterozygous species.

Acknowledgements We thank the French Ministry of Agriculture for financial support (Contrat de Branches 2008-11) and ACVF for scientific and technical contribution. We thank J. Lluch and O. Bouchez from “Plateforme Génomique” of Toulouse (France) for 454 sequencing and S. Fouilloux from BioGeves at Surgères (France) for DNA quantification.

References

- Ameline-Torregrosa, C.; Cazaux, M.; Danesh, D.; Chardon, F.; Cannon, S.B.; Esquerre-Tugaye, M.T.; Dumas, B.; Young, N.D.; Samac, D.A.; Huguet, T.; Jacquet, C. 2008. Genetic dissection of resistance to anthracnose and powdery mildew in *Medicago truncatula*. MPMI 21, 61-69.
- Ariss, J.J.; Vandemark, G.J. 2007. Assessment of genetic diversity among nondormant and semidormant alfalfa populations using sequence-related amplified polymorphisms. Crop Sci. 47, 2274-2284.
- Barnes, D.K.; Ostazeski, S.A.; Schillinger, J.A.; Hanson, C.H. 1969. Effect of anthracnose (*Colletotrichum trifolii*) infection on yield, stand, and vigor of alfalfa. Crop Sci. 9, 344-346.
- Cazaux M (2008) Etude de la résistance de la légumineuse modèle *Medicago truncatula* à *Colletotrichum trifolii*, agent de l'anthracnose. PhD Thesis Université de Toulouse.
- Elgin Jr, J.H.; Ostazeski, S.A. 1985. Inheritance of resistance to race 1 and race 2 anthracnose in Arc and Saranac AR alfalfa. Crop Sci. 25, 861-865.
- Gondran J (1984) La verticilliose de la luzerne: détermination de l'agent causal, biologie du parasite, répartition géographique, dégâts et méthodes de lutte. Thèse Université U.E.R. Sciences de Poitiers.
- Herrmann, D.; Barre, P.; Santoni, S.; Julier, B. 2010. Association of a CONSTANS-LIKE gene to flowering and height in autotetraploid alfalfa. Theor. Appl. Genet. 121, 865-876.
- Julier, B.; Bournoville, R.; Landré, B.; Ecalte, C.; Carré, S. 2004. Genetic analysis of lucerne (*Medicago sativa* L.) seedling resistance to pea aphid (*Acyrtosiphon pisum* Harris). Euphytica 138, 133-139.
- Julier, B.; Meusnier, I. 2010. Alfalfa breeding benefits from genomics of *Medicago truncatula*. Ratarstvo i Povrtarstvo / Field Veg.Crop Res. 47, 395-402.
- Mackie, J.M.; Musial, J.M.; Armour, D.J.; Phan, H.T.T.; Ellwood, S.E.; Aitken, K.S.; Irwin, J.A.G. 2007. Identification of QTL for reaction to three races of *Colletotrichum trifolii* and further analysis of inheritance of resistance in autotetraploid lucerne. Theor. Appl. Genet. 114, 1417-1426.

- Mackie, J.M.; Musial, J.M.; O'Neill, N.R.; Irwin, J.A.G. 2003. Pathogenic specialisation within *Colletotrichum trifolii* in Australia, and lucerne cultivar reactions to all known Australian pathotypes. *Aust. J. Agric. Res.* 54, 829-836.
- Michelmore, R.W.; Paran, I.; Kesseli, R.V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Nat. Acad. Sci.* 88, 9828-9832.
- O'Neill, N. 1991. Anthracnose resistance. Eds CC Fox et al. Standard tests to characterise alfalfa cultivars, D1. North American Alfalfa Improvement Conference.
- Pierre, J.B.; Bogard, M.; Herrmann, D.; Huyghe, C.; Julier, B. 2011. A CONSTANS-like gene candidate that could explain most of the genetic variation for flowering date in *Medicago truncatula*. *Molecular Breed.* 28, 25-35.
- Raynal, G., Gondran, J., Bournoville, R., Courtillot, M. 1989. 'Ennemis et maladies des prairies.' INRA: Paris
- Yang, S.; Gao, M.; Deshpande, S.; Lin, S.; Roe, B.A.; Zhu, H. 2007. Genetic and physical localization of an anthracnose resistance gene in *Medicago truncatula*. *Theor. Appl. Genet.* 116, 45-52.
- Yang, S.; Gao, M.; Xu, C.; Gao, J.; Deshpande, S.; Lin, S.; Roe, B.A.; Zhu, H. 2008. Alfalfa benefits from *Medicago truncatula*: The RCT1 gene from *M. truncatula* confers broad-spectrum resistance to anthracnose in alfalfa. *Proc. Nat. Acad. Sci.* 105, 12164-12169.

Annexe 2: Protocole d'inoculation de *Verticillium albo-atrum* par trempage des racines dans une solution de spores.

The seeds, obtained by the pods, are scarified with sandpaper and then placed in little Petri dishes on a piece of blotting paper imbibed with sterile water. The dishes are kept at 4°C in obscurity for 2-3 days, to overcome seed dormancy, and then for 24 h at 25°C (or for 48 h at 20°C) to let them germinate.

The germinated seeds are transplanted in Jiffy pots (Jiffy, Lyon, France) and then placed in little greenhouses in the phytotron, with a day/night cycle of 16/8 h, with a temperature of 25/20°C and a RH (Relative Humidity) of 75%. The plants are inoculated when they are 10 days old.

V. albo-atrum V32 strain, the only wild strain available today in France, is grown on PDA medium (Potato-Dextrose Agar, Difco™: Potato starch +Dextrose) in Petri dishes, at a temperature of 24°C and in obscurity (12-15 days).

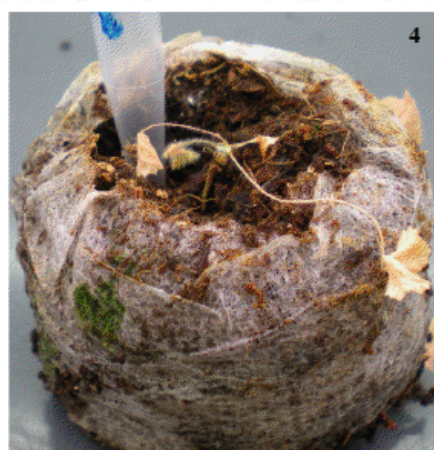
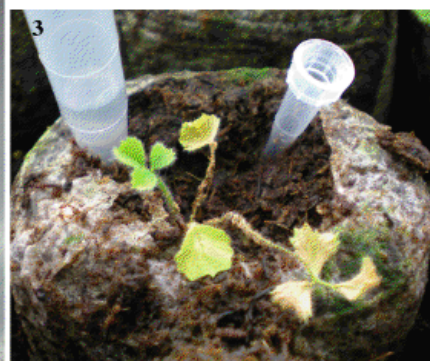
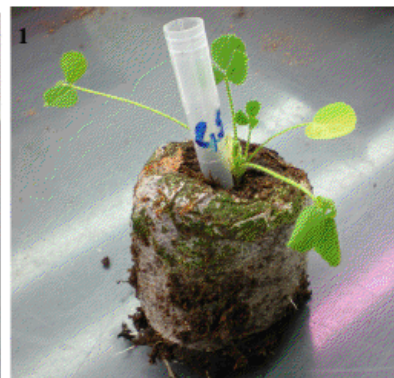
The solution of spores for inoculation is prepared by making a suspension of 10⁶ spores/mL of sterile water (concentration verified with the “Malassez cell”).

To make the spores absorption uniform in all the plants, the roots are clipped by cutting off a layer of about 1 cm at the bottom of each Jiffy pot. Then the Jiffy pots are dipped into the spores' solution for 30 minutes. At the end the Jiffy pots are placed on plates containing commercial potting soil and transferred in a phytotron with a day/night cycle of 16/8 h, with a temperature of 20/20°C and a RH of 57,9%.

The development of the symptoms is checked 2 days/week for ~30 days from the inoculation. The first symptoms appear about 7 to 10 days after the inoculation.

For the evaluation of the degree of the symptomatology, I used a disease severity index rating from 0 to 4 (Figure-Annexe 1):

- 0 = no symptoms;
- 1 = the first leaf is chlorotic;
- 2 = the first two leaves wilt;
- 3 = the entire plant wilts;
- 4 = the plant is dead.



Annexe 3: Données de cartographie fine et de phénotypage concernant les RILs LR5 recombinant sous le QTL sur le chromosome 7 et les lignées contrôles sensibles et résistants.

	LR5 F9 2	LR5 F9 5	LR5 F9 12	LR5 F9 14	LR5 F9 19	LR5 F9 21	LR5 F9 25	LR5 F9 26	LR5 F9 38	LR5 F9 39	LR5 F9 40	LR5 F9 42	LR5 F9 44	LR5 F9 45	LR5 F9 52	LR5 F9 53	LR5 F9 60	LR5 F9 61	LR5 F9 62	LR5 F9 64	LR5 F9 79	LR5 F9 81	LR5 F9 82	LR5 F9 89	LR5 F9 90	LR5 F9 91	LR5 F9 92
Mtic 1045 (MTE85)	H	A	A	B	B	B	B	B	A	B	A	B	B	B	B	A	B	A	A	B	B	A	A	B	A	B	A
Mtic 1431	H	A	A	B	B	B	B	B	A	?	?	B	B	B	B	A	A	A	A	B	?	A	A	B	A	B	A
Mtic 1426	H	?	?	B	B	B	B	B	A	B	B	?	B	B	?	A	A	A	B	B	?	A	?	B	A	?	A
Mtic 1423	H	A	A	B	B	B	B	B	?	B	?	B	B	B	B	A	B	B	A	B	B	?	A	B	A	B	A
Mtic 1425	B	A	A	B	B	B	B	A	B	A	B	A	B	B	B	A	B	?	A	B	B	?	A	B	A	B	A
Mtic 1424	B	?	?	?	?	?	?	?	B	?	?	?	?	?	B	A	A	A	B	B	?	?	?	?	?	?	H
Mtic 518 (MTE126)	?	A/H	A/H	B	B	B	B	A/H	B	A/H	B	A/H	B	A/H	B	A/H	?	A/H	A/H	B	B	A/H	A/H	B	A/H	B	A/H

Phenotyping 1

Medium score 27 dpi	2,17	2,17	0,42	2,75	3,00	2,50	0,83	2,83	1,75	3,50	0,06	0,83	1,75	0,70	0,00	0,08	0,00	0,00	0,25	0,67	0,42	2,80	0,10	1,40	0,20	2,08	0,17
n° plants	6	6	6	4	1	6	3	6	6	4	8	6	6	5	4	6	1	4	8	3	6	5	5	5	5	6	6

Phenotyping 2

Medium score 27 dpi	2.12	NA	2.98	0.56	3.67	2.08	2.23	0.00		0.00		3.06	2.33	0.29	0.00	0.00	0.00	0.70	3.00		NA	0.00	2.17	0.00	0.00	0.00
---------------------	------	----	------	------	------	------	------	------	--	------	--	------	------	------	------	------	------	------	------	--	----	------	------	------	------	------

	LR5 F9 94	LR5 F9 99	LR5 F9 100	LR5 F9 103	LR5 F9 104	LR5 F9 113	LR5 F9 115	LR5 F9 120	LR5 F9 142	LR5 F9 144	LR5 F9 147	LR5 F9 157	LR5 F9 161	LR5 F9 162	LR5 F9 181	LR5 F9 185	LR5 F9 190	LR5 F9 191	LR5 F9 193	LR5 F9 199	LR5 F9 200	LR5 F9 201	LR5 F9 202	LR5 F9 210	LR5 F9 213	LR5 F9 215
Mtic 1045 (MTE85)	A	B	B	B	B	B	B	B	A	B	A	A	A	B	H	?	A	B	A	B	A	B	?	B	H	B
Mtic 1431	A	B	B	B	A	B	B	B	A	B	A	A	A	A	A	A	A	B	A	B	A	B	A	B	H	B
Mtic 1426	A	B	B	A	?	?	B	B	A	B	?	A	?	B	B	A	B	B	?	B	?	B	A	B	?	B
Mtic 1423	A	B	B	B	H	B	B	B	A	B	A	A	A	A	H	H	B	B	B	A	B	A	B	A	A	B
Mtic 1425	B	B	B	B	A	A	B	A	A	B	?	A	?	B	B	A	B	B	A	B	A	B	?	B	A	?
Mtic 1424	B	?	?	?	A	?	?	?	?	?	?	A	?	A	?	?	?	?	?	?	A	B	A	B	A	?
Mtic 518 (MTE126)	B	B	B	A/H	A/H	?	B	A/H	A/H	B	A/H	B	A/H	A/H	A/H	A/H	B	B	A/H	?	A/H	B	A/H	B	A/H	A/H

Phenotyping 1

Medium score 27 dpi	0,00	0,75	0,31	2,11	3,90	2,19	0,00	1,25	0,00	1,71	3,29	0,17	1,63	1,31	0,17	1,93	0,20	0,29	3,13	3,25	0,29	2,14	0,07	1,79	1,64	2,25
n° plants	7	6	8	8	5	8	3	6	2	7	7	6	8	8	6	6	5	7	8	6	7	7	7	7	7	8

Phenotyping 2

Medium score 27 dpi	0.00	2.59		1.53	2.00	1.38	0.00	1.91		2.56	3.34	0.53	2.05	1.84	1.18		0.12	0.17	2.27	3.24	0.12	2.20	0.94	1.97	1.78	1.66
---------------------	------	------	--	------	------	------	------	------	--	------	------	------	------	------	------	--	------	------	------	------	------	------	------	------	------	------

Annexe 4. Résumé du poster présenté au Meeting 'From foliar to root-interacting pathogens and symbionts' en février 2011.

Genetic traits and regulatory mechanisms involved in resistance of the model legume plant *Medicago truncatula* towards *Verticillium albo-atrum*.

Ben C.¹, Montanari S.¹, Amatya S.¹, Gras M.C.², Noël D.², Gentzbittel L.¹, Rickauer M.¹

¹Université de Toulouse, INPT, UMR 5245 Ecologie fonctionnelle et Environnement (Ecolab), ENSAT, 18 chemin de Borde Rouge, BP 31607 Auzeville Tolosane, 31326 Castanet-Tolosan, France

²Tourneur Barenbrug Recherches, Negadis, 82600 Mas-Grenier, France

Infection of Alfalfa with *Verticillium albo-atrum* (Vaa) gives rise to vascular wilt disease causing serious losses in Europe. The alfalfa pathogen also infects the model legume plant *Medicago truncatula* under various conditions. Symptoms caused in *M. truncatula* were rated on a scale from 0 to 4, and analysed by a linear regression model of symptom evolution which characterises the phenotype by Asym (maximum disease index) and Xmid (time to reach 50% of Asym). Evaluation of a collection of 32 *M. truncatula* lines showed different degrees of susceptibility, from highly susceptible to fully resistant lines and differing by Asym and/or Xmid. Studies of a RILs population from the cross between lines Jemalong A17 and F83005.5 allowed us to identify a major QTL on chromosome 7 responsible for 45% of the resistance phenotype. This QTL in Jemalong A17 could be confirmed in another RILs population.

Gene regulation by microRNAs (miRNAs) has gained much interest during the last years. In order to identify regulatory RNAs in the interaction between *M. truncatula* and Vaa, we have produced libraries of small RNAs from control and infected roots of the resistant and susceptible line. More than 5 million reads were obtained by Solexa sequencing for each condition. Small 20-22nt RNAs unique sequences were identified after quality and redundancy trimming and are currently used for *in silico* prediction of potential miRNA genes on the *M. truncatula* genome MtV3.

In order to set up conditions for functional analyses, we investigated the role of miR393, a miRNA previously described to be involved in basal resistance against *P. syringae* in *A. thaliana*. We found that miR393 is induced strongly in the susceptible line after inoculation and that its corresponding target, MtTIR1, is repressed during infection. A first experiment showed that over-expression of miR393 in transformed roots leads to a delayed onset of disease symptoms. These preliminary results suggest that miR393 may be involved in root response to *Verticillium albo-atrum*.

Part of this work was supported by the Genoscope project "MIRMED", coordinated by Martin Crespi, ISV Paris.

Annexe 5. Résumé de la communication orale donnée lors du 'Model Legume Congress' en mai 2011.

Genetic traits and regulatory mechanisms involved in resistance of *Medicago truncatula* to *Verticillium albo-atrum*.

Ben C.¹, Montanari S.¹, Toueni M.¹, Amatya S.¹, Negahi A.¹, Gras M.C.², Noël D.², Gentzbittel L.¹, Rickauer M.¹

¹Université de Toulouse, INPT, UMR 5245 Ecologie fonctionnelle et Environnement (Ecolab), ENSAT, 18 chemin de Borde Rouge, BP 31607 Auzeville Tolosane, 31326 Castanet-Tolosan, France

²Tourneur Barenbrug Recherches, Negadis, 82600 Mas-Grenier, France

Infection of alfalfa with the soil-borne fungus *Verticillium albo-atrum* (Vaa) gives rise to vascular wilt disease causing serious losses in Europe. We found that Vaa also infects *Medicago truncatula* under various conditions. The evaluation of a collection of 32 *M. truncatula* lines showed different degrees of susceptibility, from highly susceptible to fully resistant lines and differing by maximum disease index and/or onset of symptom development. This indicates a high biodiversity in the disease response.

By studying a RILs population from a cross between lines Jemalong A17 (resistant) and F83005.5 (susceptible) we identified a major QTL on chromosome 7 which is responsible for 45% of the resistance phenotype. This QTL in Jemalong A17 was confirmed with a further RILs population with a different susceptible parent.

Vaa was transformed with the *GFP* marker gene in order to study root colonization by epifluorescence microscopy and confocal laser scanning microscopy. Observations of roots inoculated with the GFP-expressing strain showed that vigorous fungal growth occurs in the xylem vessels and parenchyma cells of the stele in line F83005.5, whereas it is only weak in line A17 and accompanied by accumulation of phenolic compounds in the root cells.

Analyses of a set of defence-related genes by qRT-PCR showed differential induction in roots of the resistant and susceptible line. In order to study the regulation of defence in this pathosystem, and to set up conditions for functional analyses, we investigated the role of miR393, a miRNA previously described to be involved in basal resistance against *Pseudomonas syringae* in *Arabidopsis thaliana*. We found that miR393 is transiently and significantly induced in the resistant line after inoculation. The corresponding target gene *MtTIR1*, is induced in both the resistant and susceptible line, but much stronger in the susceptible one, which is in agreement with its regulation by miR393. A first experiment showed that over-expression of miR393 in *Agrobacterium rhizogenes*-transformed roots led to a delayed onset of disease symptoms. Taken together, our preliminary results suggest that miR393 may be involved in the regulation of root responses to *Verticillium albo-atrum*.

Part of this work was supported by "Contrat de branches C2008-16 Luzerne" of the Ministry of Agriculture, coordinated by B. Julier.