



**HAL**  
open science

# Evolution moléculaire sous pression de sélection et implication dans la reconnaissance avr<sub>lm3</sub>/rlm<sub>3</sub> du gène d'avirulence avr<sub>lm4-7</sub> chez *Leptosphaeria maculans*

Guillaume Daverdin

► **To cite this version:**

Guillaume Daverdin. Evolution moléculaire sous pression de sélection et implication dans la reconnaissance avr<sub>lm3</sub>/rlm<sub>3</sub> du gène d'avirulence avr<sub>lm4-7</sub> chez *Leptosphaeria maculans*. Sciences agricoles. Université Paris Sud - Paris XI, 2011. Français. NNT : 2011PA112022 . tel-00676204

**HAL Id: tel-00676204**

**<https://theses.hal.science/tel-00676204>**

Submitted on 4 Mar 2012

**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.

Université Paris-Sud XI  
Ecole Doctorale Science du Végétal : du Gène à l'Ecosystème



## THESE

pour l'obtention du titre de

**Docteur de l'Université Paris-Sud XI**  
**Ecole Doctorale Science du Végétal : du Gène à l'Ecosystème**

présentée et soutenue publiquement  
le 04 Mars 2011

par

**Guillaume DAVERDIN**

**Evolution moléculaire sous pression de sélection et implication dans la  
reconnaissance *AvrLm3/Rlm3* du gène d'avirulence *AvrLm4-7* chez  
*Leptosphaeria maculans***

devant le jury composé de :

Mme. Cécile FAIRHEAD  
M. Didier THARREAU  
M. Benoit MOURY  
Mme. Martine LEFLON  
Mme. Marie Hélène BALESSENT

Université Paris-sud XI  
CIRAD Montpellier  
INRA Avignon  
CETIOM Grignon  
INRA Versailles-Grignon

Présidente  
Rapporteur  
Rapporteur  
Examinatrice  
Directrice de thèse



Université Paris-Sud XI  
Ecole Doctorale Science du Végétal : du Gène à l'Ecosystème



## THESE

pour l'obtention du titre de

**Docteur de l'Université Paris-Sud XI**  
**Ecole Doctorale Science du Végétal : du Gène à l'Ecosystème**

présentée et soutenue publiquement  
le 04 Mars 2011

par

**Guillaume DAVERDIN**

**Evolution moléculaire sous pression de sélection et implication dans la  
reconnaissance *AvrLm3/Rlm3* du gène d'avirulence *AvrLm4-7* chez  
*Leptosphaeria maculans***

devant le jury composé de :

Mme. Cécile FAIRHEAD  
M. Didier THARREAU  
M. Benoit MOURY  
Mme. Martine LEFLON  
Mme. Marie Hélène BALESSENT

Université Paris-sud XI  
CIRAD Montpellier  
INRA Avignon  
CETIOM Grignon  
INRA Versailles-Grignon

Présidente  
Rapporteur  
Rapporteur  
Examinatrice  
Directrice de thèse



## Remerciements

Cette thèse a été réalisée au laboratoire de génétique moléculaire des interactions *Leptosphaeria maculans* / colza de l'unité BIOGER-CPP de l'INRA Versailles-Grignon. Je tiens donc tout d'abord à remercier Mylène Balesdent et Thierry Rouxel de m'avoir accueilli au sein de leur équipe. Ces quatre années ont été extrêmement instructives pour moi, tant d'un point de vue scientifique qu'humain. Merci pour votre disponibilité, votre aide et l'opportunité offerte de travailler sur un sujet aussi passionnant ! Je remercie également Mr Marc Henri Lebrun, directeur de l'unité BIOGER-CPP de m'avoir accueilli dans l'unité.

Merci à Olivier Le Gall et au département SPE pour le soutien financier apporté au projet de thèse. Merci également à Jean-Noël Aubertot pour la mise en place de l'expérimentation et sa collaboration ainsi qu'à Julien Carpezat, Martine Leflon et Xavier Pinochet (CETIOM) pour nos échanges et le travail effectué autour de l'analyse HRM.

Je remercie messieurs Benois Moury et Didier Tharreau pour avoir accepté d'être les rapporteurs de mes travaux de thèse ainsi que mesdames Martine Leflon et Cecile Fairhead pour leur participation au jury de thèse. Merci également à Tatiana Giraud, Fabienne Malagnac, Bruno Lecam et Christophe Lemaire pour avoir participé à mes comités de thèse.

Je tiens également à remercier Aline Mahé et Sophie Nadot de m'avoir accueilli dans leurs équipes enseignantes à l'occasion de mon monitorat.

Je remercie tout particulièrement Michel Dron pour son travail au sein de l'école doctorale, sa disponibilité et son dévouement envers les doctorants. Par la même occasion, un grand « chapeau bas » au bureau des doctorants ! Merci également à Maryelle Fradin pour son travail et son aide ! Merci également à Mmes Jacqui Shykoff et Marianne Delarue pour avoir pris le relais à la tête de l'ED : bonne continuation !

Un grand merci aux membres de BIOGER que j'ai pu côtoyer. Un merci tout particulier à Lilian Gout pour son aide et sa disponibilité et à Angélique Gautier pour les réactions de séquences ! Je souhaite particulièrement remercier « les leptos », qu'ils soient « permanents » ou de « passage ». Tout d'abord Isabelle Fudal, pour son aide et ses conseils au cours du master et pendant la thèse, merci beaucoup ! Un merci tout particulier à Laurent Coudard pour son aide précieuse ! Merci à Bénédicte Ollivier pour son travail mais peut être encore plus pour avoir supporté et même souri à mes blagues... Merci également à Juliette (niveau « blagues pas drôles » tu as également été servie...), Jessica, Jonathan, Salim (désolé de ne pas avoir effectué toute la liste...), Alexandre, Azita (moi, je ne t'oublie pas ☺), Michel, Françoise, Jacques. Aux leptos de « passage » : Pascal, Nicolas, Thomas, Mennat, Marine... Bonne continuation à tous et à très bientôt ! Pour finir avec le labo, une dédicace spéciale à « mon » ficus : sois fort et survis !

Je remercie également ma famille et en particulier ma mère et mon frère, merci de votre patience, de votre soutien... Merci pour tout ! Une pensée toute particulière pour mes grands pères, j'aurai aimé vous avoir avec moi à la fin de cette thèse...

Pour finir, comment ne pas vous citer ? Sans ordre : Vivien, Guillaume, Thomas, Maud, Julie, Jérôme, Sébastien, Stéphanie, Elodie. Merci pour votre soutien, votre amitié, votre présence...



# Abréviations

aa.	acides aminés
ADN(c)	Acide DésoxyriboNucléique (complémentaire)
ARN(m)	Acide RiboNucléique (messenger)
AFLP	Amplified Fragment Length Polymorphism
BAC	Bacterial Artificial Chromosome
BET	bromure d'éthidium
CAPS	Cleaved Amplified Polymorphic Sequence
cM	centimorgan
dNTP	désoxynucléotide triphosphate
f. sp.	forma specialis
hrp	hypersensitive response and pathogenicity
HRM	High Resolution Melting
IR	Inverted Repeat
ITS	Internal Transcribed Spacer
kb	kilobase
LRR	Leucine Rich Repeats
LTR	Long Terminal Repeat
MAT	signe de compatibilité sexuelle
NBS	Nucleotide Binding Site
NLS	Nuclear Localisation Signal
ORF	Open Reading Frame
pb	paire de bases
PCR	Polymerase Chain Reaction
QTL	Quantitative Trait Loci
RACE-PCR	Rapid Amplification of cDNA Ends – PCR
RIP	Repeat Induced Point mutation
RT	Reverse Transcription (ou Reverse Transcriptase)
TAIL	Thermal Asymmetric Interlaced
TIR	Toll/Interleukin Receptor
UTR	Untranslated Transcribed Region
Cv	cultivar





# **Sommaire**



# SOMMAIRE

## INTRODUCTION GENERALE

Avant-propos.....	15
<b>A. Résistance des plantes aux microorganismes.....</b>	<b>16</b>
1. <u>Résistance non hôte</u> .....	17
2. <u>Résistance hôte</u> .....	18
a) Relation gène-à-gène.....	19
b) Gènes d'avirulence.....	21
c) Rôle intrinsèque des gènes d'avirulence.....	25
d) Gènes de résistance.....	27
e) Modèles de reconnaissance entre protéines de résistance et protéines d'avirulence.....	30
<b>B. Evolution des gènes d'avirulence vers la virulence.....</b>	<b>33</b>
a) Contournement de gènes de résistance.....	33
b) Mécanismes moléculaires à l'origine du contournement de la résistance.....	37
c) Coût de la virulence / déficit de fitness.....	41
d) Influence du coût de virulence.....	41
e) Durabilité d'une résistance.....	42
<b>C. Couple <i>Brassica napus</i> – <i>Leptosphaeria maculans</i>.....</b>	<b>45</b>
1. <u><i>Brassica napus</i></u> .....	45
2. <u><i>Leptosphaeria maculans</i></u> .....	45
a) Description / taxinomie / complexe d'espèce.....	45
b) Epidémiologie / Cycle de vie.....	49
c) Caractéristiques du génome de <i>L. maculans</i> .....	51
3. <u>Interaction <i>L. maculans</i> - <i>B. napus</i></u> .....	52
a) Résistance quantitative.....	52
b) Résistance spécifique.....	53
c) Gènes d'avirulence.....	53
d) Gènes de résistance.....	55
e) Contournement et durabilité des résistances à <i>L. maculans</i> .....	55
f) Couple <i>AvrLm7-Rlm7</i> .....	57
<b>D. Objectif de la thèse.....</b>	<b>58</b>

## CHAPITRE 1

### Double spécificité d'interaction du gène d'avirulence *AvrLm4-7*

Introduction.....	65
-------------------	----

### **Article 1**

*Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change

## **Chapitre 2**

### **Evolution moléculaire du gène d'avirulence *AvrLm4-7* sous pression de sélection**

Introduction.....87

#### **Article 2**

Experimental evolution of the *AvrLm4-7* gene of *Leptosphaeria maculans* under *Rlm7* resistance gene selection pressure in the field

## **Chapitre 3**

### **Interaction *AvrLm4-7* – *AvrLm3***

Introduction.....155

#### **Article 3**

A game of hide and seek: the complex interplay between the avirulence genes *AvrLm4-7* and *AvrLm3* in *Leptosphaeria maculans*

## **Discussion générale**

1 Aspect fonctionnel.....205  
2 Aspect évolutif.....215

## **Annexe**

Annexe 1 .....227

## **Références bibliographiques**

Références bibliographiques.....231

# **Introduction** **générale**



## Introduction

Dans leurs écosystèmes naturels ou au sein de milieux anthropiques, les plantes sont en constante et étroite relation avec de nombreux microorganismes : virus, bactéries, champignons, oomycètes, nématodes. Une grande gamme d'interactions peut être définie en fonction des bénéfiques qu'en retirent les différents partenaires : depuis la symbiose, lorsque une association à bénéfices réciproques s'établit, jusqu'au parasitisme, quand le microorganisme se développe aux dépens de la plante. Les organismes entrant dans cette dernière catégorie sont qualifiés d'agents pathogènes et sont souvent classés selon leur mode de nutrition. Deux grandes catégories sont ainsi établies : les organismes biotrophes, qui ont besoin de tissus vivant pour se nourrir, tel l'oomycete *Peronospora parasitica*, et les organismes nécrotrophes, tel le champignon *Botrytis cinerea*, qui se nourrissent des tissus nécrosés qu'ils viennent de dégrader. Certains organismes, comme *Leptosphaeria maculans* ou *Magnaporthe oryzae*, se comportent comme biotrophe ou nécrotrophe en fonction du stade de leur cycle de vie et sont qualifiés d'hémibiotrophes (Glazebrook, 2005).

L'invention de l'agriculture il y a environ 10000 ans a été une véritable révolution dans l'histoire de l'humanité avec comme principales conséquences la sédentarisation et un formidable essor des populations. En 2006, l'humanité comptait 6 milliards d'individus et devrait franchir la barre des 9 milliards d'ici à 2050. Nourrir une population aussi importante est l'un des enjeux majeurs de notre siècle et de grands défis agricoles devront être résolus pour pouvoir y répondre, en particulier concernant la protection des récoltes et la réduction des pertes dues aux agents pathogènes. Le développement de microorganismes pathogènes sur des plantes cultivées provoque en effet chaque année des pertes estimées entre 10% et 30% du potentiel agricole mondial (Agrios, 1997). Ajouté à l'impact économique qu'elles provoquent, ces pertes peuvent avoir des conséquences sociales considérables. L'exemple le plus parlant est sans doute la famine due à une épidémie de *Phytophthora infestans* sur les plantations de pomme de terre en Irlande au milieu du 19<sup>ème</sup> siècle, qui provoqua la mort de centaines de milliers de personnes et l'émigration de plus de 2 millions d'Irlandais vers les Etats-Unis d'Amérique (Donnelly, 2001).



Afin de répondre aux menaces posées par les agents pathogènes, plusieurs moyens de lutte, plus ou moins efficaces, ont été développés. Le plus utilisé d'entre eux se base sur l'emploi de pesticides, substances actives tuant ou empêchant le développement des agents pathogènes visés. Malgré une efficacité parfois remarquable de ces produits, de nombreux cas de résistance aux pesticides, auxquels s'ajoutent des problèmes de santé publique et d'impact environnemental, font des pesticides un moyen de lutte pour lequel des alternatives sont activement recherchées. Parmi ces alternatives, la lutte intégrée ou la lutte biologique présentent un intérêt environnemental notable mais restent encore relativement anecdotiques. L'alternative la plus développée à ce jour, la lutte génétique, exploite la présence dans la plante cultivée de gènes contrôlant les mécanismes de défense naturels de celle-ci contre le développement des agents pathogènes. Cette méthode est néanmoins limitée par le développement d'agents pathogènes pouvant outrepasser ces défenses. De nombreuses études visant à améliorer la compréhension de ces mécanismes de défense ont permis de rassembler de nombreuses données sur les interactions entre plantes et agents pathogènes. L'introduction que je propose ici a pour but d'aborder ces interactions ainsi que les mécanismes génétiques et moléculaires sous-jacents à cette lutte génétique.

## **A. Résistance des plantes aux microorganismes**

La résistance d'une plante à un agent pathogène se caractérise par une limitation totale ou partielle du développement de celui-ci dans la plante. La résistance a été décomposée en plusieurs catégories pouvant se recouper, selon son hérédité, son efficacité ou encore sa spécificité. Ainsi, une résistance va être décrite comme monogénique ou polygénique en fonction de son déterminisme génétique. Une résistance monogénique dépend d'un unique gène tandis qu'une résistance polygénique se base sur la fonction de plusieurs gènes contribuant à la résistance. La résistance peut également être qualitative (totale) ou quantitative (partielle). Dans le premier cas, l'agent pathogène est incapable de se développer sur la plante, la résistance est totale tandis que dans le second cas, la résistance est partielle, l'agent pathogène ne peut se développer dans la plante que de façon limitée. Enfin, selon le critère de spécificité d'hôte, qui sera celui choisi afin de

décrire la résistance des plantes dans cette introduction, on parlera de résistance « hôte » ou « non-hôte ».

En 1728, Duhamel De Monceau démontre qu'il existe, pour une catégorie d'agent pathogène donnée, des plantes hôtes et des plantes non-hôtes (Rapilly, 1998). En effet, pour une espèce végétale donnée, seules quelques espèces d'agents pathogènes sont assez adaptées pour pouvoir se développer et, réciproquement, une espèce d'agent pathogène possède le plus souvent un nombre relativement limité d'espèces de plantes qu'il pourra infecter (sa gamme d'hôte) (Knogge, 1996). Dans le cas d'une plante non-hôte, l'incompatibilité entre la plante et l'agent pathogène est due à une résistance « non-hôte ». Dans le cas d'une incompatibilité entre un agent pathogène et une plante hôte, on parle de résistance « hôte ». Les résistances « hôte » et « non-hôte » permettent d'expliquer que la très grande majorité des attaques d'agents pathogènes ne donnent pas lieu au développement d'une maladie, la maladie étant " l'exception plutôt que la règle " (Niks, 1987; Thordal-Christensen, 2003).

## 1. Résistance non-hôte

La résistance non-hôte est la plus répandue dans les interactions plante / agent pathogène et consiste en une immunité innée, non spécifique des génotypes des microorganismes en interaction, et partagée par tous les individus d'une même espèce végétale (Niks, 1987). Le développement d'une succession de systèmes de défense non spécifiques pour chaque étape de l'infection fait de la résistance non-hôte le principal obstacle au développement d'un agent pathogène (Heath, 2000; Thordal-Christensen, 2003; da Cunha *et al.*, 2006).

Dès la mise en contact avec la plante, le développement du parasite ne peut s'effectuer qu'après la reconnaissance d'un environnement favorable à son développement et après avoir déjoué des mécanismes de défense constitutifs mis en place préventivement par la plante, qu'ils soient physiques (cuticule, paroi cellulaire, etc.) ou moléculaires (présence d'enzymes antimicrobiennes, phytoanticipines, etc) (Thordal-Christensen, 2003; Mysore et Ryu, 2004). Ainsi, la formation d'appressorium (structure spécialisée dans la pénétration des tissus végétaux) chez *Blumeria graminis f. sp. hordei*, champignon pathogène de l'orge,

est dépendante de la présence et de la quantité d'un aldéhyde spécifique (C<sub>26</sub>) dans la cire présente sur les feuilles de la plante rencontrée. Dans le cas d'une concentration plus faible ou d'un aldéhyde différent de celui rencontré sur l'orge, la formation d'appressorium sera réduite voir nulle, empêchant alors le développement du champignon (Tsuba *et al.*, 2002).

L'agent pathogène s'expose ensuite à une série de mécanismes inductibles liés à la reconnaissance par la plante de signatures moléculaires associées aux agents pathogènes portés par des molécules élicitrices\* nommés PAMPs (Pathogen Associated Molecular Patterns) (Chisholm *et al.*, 2006). Ces PAMPs sont partagés par un grand nombre d'agents pathogènes et la reconnaissance de ces éliciteurs généraux permet de stopper le développement du parasite par la mise en place de réponses localisées (lignification de la paroi, production de phytoalexines, etc.) entrant en lice dans ce qui est communément appelé la PTI (PAMP Triggered Immunity) (Chisholm *et al.*, 2006). Le PAMP le plus étudié est la flagelline, protéine constitutive des flagelles indispensable à la motilité d'un grand nombre d'organismes procaryotes (Gomez-Gomez et Boller, 2002). La flagelline possède un domaine hautement conservé constitué de 22 acides aminés, nommé flg22, faisant office de PAMP (Felix *et al.*, 1999). La reconnaissance de flg22 s'effectue chez *Arabidopsis thaliana* par l'intermédiaire d'un récepteur moléculaire codé par le gène *FLS2* (flagellin-sensitive-2), qui une fois lié à flg22, est internalisé et déclenche des mécanismes de défense permettant de restreindre la croissance de la bactérie pathogène (Gomez-Gomez et Boller, 2000; Zipfel *et al.*, 2004). Chez les champignons, la chitine, polysaccharide présent dans la paroi cellulaire, constitue un des exemples de PAMP les plus connus (Numberger *et al.*, 2004).

## 2. Résistance hôte

Au fil de l'évolution, les agents pathogènes ont acquis différents outils moléculaires (cutinases, inhibiteur de protéases, etc.) ou organes spécialisés (appressorium, haustorium) qui leur donnent la possibilité de contourner les différents obstacles de la résistance « non-hôte » les empêchant de coloniser une plante (Knogge, 1996). L'infection est alors possible et la plante est dite « hôte ». Néanmoins, tous les agents pathogènes ayant acquis la capacité à surmonter la résistance « non-hôte » d'une plante ne sont pas capables d'infecter celle-ci. Cette

\* molécule élicitrice/éliciteur : molécule déclenchant les réactions de défense de la plante



incapacité est le fait de la résistance « hôte », due à des mécanismes déclenchés de façon spécifique à une espèce ou un petit nombre d'espèces végétales en interaction. Cette résistance « hôte » n'est efficace que face à une portion des génotypes de cet agent pathogène. Le succès ou l'échec de l'attaque dépend alors de la capacité du parasite à échapper à une reconnaissance reposant sur les principes de la « relation gène-à-gène » (Knogge, 1996; Thordal-Christensen, 2003).

#### a) Relation gène-à-gène

A la suite de ses travaux de génétique portant sur l'interaction entre le champignon filamenteux *Melampsora lini* et le lin (*Linum usitatissimum*), H. Flor développa en 1942 le concept de la relation gène-à-gène (Flor, 1971). Cette relation suppose qu'à tout gène de résistance dominant chez la plante correspond un gène d'avirulence dominant chez l'agent pathogène (figure 1). L'interaction entre un agent pathogène possédant un allèle avirulent (Avr) et une plante hôte possédant l'allèle résistant (R) correspondant est appelée interaction « incompatible » et déclenche le plus souvent une réaction hypersensible (RH) au niveau du site d'infection. Cette RH aura pour conséquence de stopper la progression du microorganisme et s'illustre le plus souvent par la mort de quelques cellules végétales au niveau du site d'infection ainsi que par l'expression de gènes de défense. On parle dans le cas d'une interaction incompatible, de plantes « résistantes » et d'agents pathogènes « avirulents ». Toute autre combinaison allélique n'empêchera pas le développement de la maladie, l'interaction sera alors « compatible », la plante dite « sensible » et l'agent pathogène « virulent » (Niks, 1987; Chisholm *et al.*, 2006).

Le modèle gène-à-gène tel que présenté par Flor n'a été rencontré que chez des agents pathogènes biotrophes ou hémibiotrophes (Glazebrook, 2005). En effet, la mort cellulaire provoquée par la reconnaissance de l'agent pathogène ne va pas avoir les mêmes conséquences en fonction du mode de nutrition de celui-ci : tandis que les microorganismes biotrophes ou hémibiotrophes, privés de leurs ressources nutritives, ne peuvent plus se développer, les microorganismes nécrotrophes vont au contraire être favorisés (Glazebrook, 2005; Friesen *et al.*, 2007). Néanmoins, un cas particulier du modèle gène-à-gène est à considérer pour certains organismes nécrotrophes de la classe des Dothidéomycetes. Tandis que la reconnaissance d'un

Génotype agent pathogène \ Génotype hôte	<i>R</i>	<i>r</i>
<i>Avr</i>	RH	maladie
<i>avr</i>	maladie	maladie

 Interaction incompatible  
 Interaction compatible

### **Figure 1 : relation gène à gène**

L'issue de l'interaction d'un agent pathogène et d'une plante est conditionnée par leurs génotypes. En présence de l'allèle dominant *R* d'un gène de résistance donné chez la plante et de l'allèle *Avr* dominant du gène d'avirulence correspondant chez l'agent pathogène, le développement de l'agent pathogène est bloqué par la mise en place d'une réaction hypersensible (RH). L'interaction est dite incompatible. Toute autre combinaison allélique donne lieu à une interaction compatible permettant le développement de la maladie. Adapté de Flor, (1971).

microorganisme biotrophe via le modèle gène-à-gène mène à la résistance de l'hôte, en partie via une mort cellulaire localisée, il existe un modèle miroir à celui présenté par Flor dans lequel les protéines d'avirulence sécrétées (ex : protéines HST) par l'agent pathogène agissent comme des protéines de virulence et les gènes de résistance présents chez l'hôte sont ici des gènes de sensibilité. La reconnaissance de l'agent pathogène va dans ce cas favoriser la maladie en promouvant la mort cellulaire des cellules végétales infectées (Friesen *et al.*, 2007 ; Oliver et Solomon, 2010). Ce type d'interaction moléculaire restant très marginal par rapport à la relation gène-à-gène typique, je considérerai dans ce manuscrit la relation gène-à-gène présentée par Flor comme le modèle de référence.

#### b) Gènes d'avirulence

De multiples exemples validant le modèle gène-à-gène ont, depuis sa présentation, été décrits chez les virus, bactéries, champignons, oomycètes et nématodes (Rouxel et Balesdent, 2010). Néanmoins, je ne m'intéresserai ici qu'aux gènes identifiés chez les bactéries, les champignons ainsi que chez les oomycètes.

##### i. Gènes d'avirulence bactériens

En 1984 a été réalisé le clonage du premier gène d'avirulence, AvrA, chez *Pseudomonas syringae* pv. *glycinera* (Staskawicz *et al.*, 1984). Depuis, plus de 40 autres gènes d'avirulence bactériens ont été clonés et permettent d'obtenir un aperçu des principales caractéristiques de ces gènes (Rouxel et Balesdent, 2010). La première de ces caractéristiques est leur mode de translocation au sein de la cellule végétale. Les barrières que constituent les membranes et les parois séparant les deux espaces cellulaires que forment la bactérie et la cellule végétale, sont franchies à l'aide d'un système protéique complexe fonctionnant comme une seringue nanométrique appelé système de sécrétion de type III (T3SS) codé par des gènes en cluster (Mudgett, 2005). Malgré quelques exceptions comme par exemple la protéine d'avirulence AvrXa21, chez *Xanthomonas oryzae* pv. *oryzae*, sécrétée dans le milieu apoplastique via l'utilisation d'un mécanisme dépourvu de pilus et reposant sur un système multiprotéique proche des transporteurs ABC (Adenosine triphosphate-Binding Cassette), nommé sécrétion de type I (Young et Holland, 1999; Lee *et al.*, 2006), la très grande majorité des protéines d'avirulence bactériennes répertoriées à ce jour sont injectées dans les cellules hôtes via le T3SS.

La seconde caractéristique des gènes d'avirulence bactériens est la très faible similarité de séquence que partagent ces gènes entre eux ainsi qu'avec les séquences de bases de données. De la même façon, très peu de domaines conservés ont été répertoriés à ce jour (Rouxel et Balesdent, 2010).

Les deux familles multigéniques *AvrBs3* et *AvrRxv/YopJ* présentes chez *Pseudomonas sp* et *Xanthomonas sp* constituent à ce jour les principales exceptions à ces caractéristiques (Lahaye et Bonas, 2001). Les gènes de la famille multigénique *AvrBs3* possèdent entre 90 et 97% d'identité entre eux et codent pour des protéines différant principalement par le nombre de répétitions d'un motif conservé de 34 acides aminés (AA) (Lahaye et Bonas, 2001). Ce motif joue un rôle structural plutôt qu'enzymatique et détermine la spécificité d'interaction entre la bactérie et la plante hôte. Ajouté à ce motif conservé, toutes les protéines de la famille *AvrBs3* possèdent des motifs NLS (Nuclear Localisation Signal) d'adressage au noyau et des motifs d'activation de la transcription AAD (Acidic transcriptional Activation Domain) (Lahaye et Bonas, 2001). Kay *et al.* (2007) ont montré qu'après dimérisation au niveau du cytoplasme de la cellule végétale, les protéines étaient importées dans le noyau de la cellule. Une fois en place, elles jouent alors le rôle de facteurs de transcription et provoqueraient une augmentation de la taille des cellules du mésophylle et de la production protéique de la cellule, donnant à la bactérie un environnement propice à son développement.

La famille multigénique *AvrRxv/YopJ* est quant à elle constituée de gènes codant pour des effecteurs retrouvés chez les agents pathogènes de plantes et de mammifères. Cette famille tient son nom des premiers effecteurs à avoir été isolés pour chacune de ces catégories, respectivement *AvRxv* chez l'agent pathogène des plantes *Xanthomonas campestris* pv. *vesicatoria* et *YopJ* chez l'agent pathogène des mammifères *Yersinia pseudotuberculosis* (Whalen *et al.*, 1993; Galyov *et al.*, 1994). Ces effecteurs partagent un domaine commun codant pour une protéase putative dont l'activité conduirait à la reconnaissance par la plante (Orth *et al.*, 2000).

## ii. Gènes d'avirulence fongiques et d'oomycètes

Sept ans après le clonage du premier gène d'avirulence bactérien, *Avr9* est, en 1991, le premier gène d'avirulence cloné chez un champignon, *Cladosporium fulvum* (van Kan *et al.*, 1991). Le clonage du premier gène d'avirulence d'un oomycète dut attendre 2004 avec le clonage de *Avr1b-1* chez *P. sojae* (Shan *et al.*,

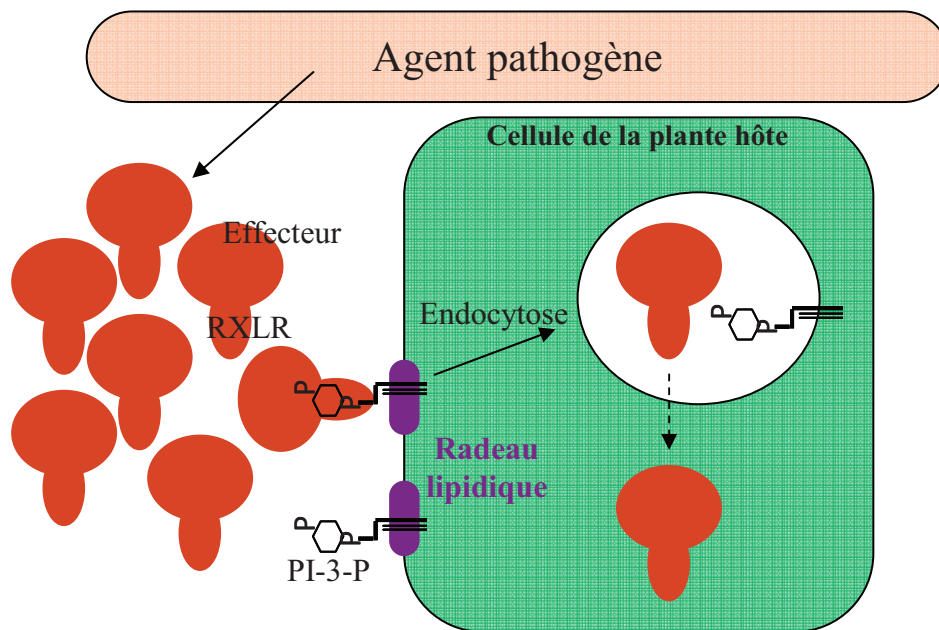
2004). Depuis, respectivement, 25 et 12 autres gènes d'avirulence ont été clonés chez les champignons et les oomycètes (Rouxel et Balesdent, 2010).

Comme chez les bactéries, les gènes d'avirulence de champignons ou d'oomycètes présentent très peu d'homologies de séquence, que ce soit entre gènes d'avirulence ou avec les bases de données (Rouxel et Balesdent, 2010). A ce jour et malgré des exceptions tel le gène d'avirulence ACE1 de *M. oryzae* dont la séquence code pour une protéine de 4034 acides aminés (Bohnert *et al.*, 2004), ces gènes codent pour des petites protéines (63-314 acides aminés) (Rouxel et Balesdent, 2010). Ces protéines possèdent dans la plupart des cas un peptide signal d'adressage extra-cellulaire les destinant à l'exportation dans les espaces intercellulaires de la plante (apoplaste, matrice extra-haustoriale...) (Rep, 2005; Dodds *et al.*, 2009). Il existe certaines protéines d'avirulence, telles AVR-a10 et AVR-k1, protéines d'avirulence de *B. graminis* pv. *hordei*, ne possédant pas de peptide signal, suggérant l'existence de voies de sécrétion alternatives (Ridout *et al.*, 2006). Le plus souvent, les protéines d'avirulence sont également riches en cystéine, acide aminé qui intervient dans la formation de ponts disulfures au sein de la protéine. Ces ponts permettraient de maintenir la structure tertiaire de la protéine, et donc a priori sa fonction. Ils confèreraient également à la protéine une relative stabilité face aux protéases végétales abondantes dans les espaces apoplastiques (Rep, 2005).

Une fois sécrétées, les protéines d'avirulence peuvent soit, comme les protéines d'avirulence de *C. fulvum* (Avr2, Avr4, Avr9, etc.) (Thomma *et al.*, 2005), rester dans le milieu intercellulaire, soit, comme les protéines d'avirulence de *M. lini* (AvrL567, AvrM, etc.), être internalisées dans la cellule végétale (Ellis *et al.*, 2007).

A la différence des bactéries, aucun système de sécrétion de type T3SS ou apparenté n'a été à ce jour décrit chez les champignons ou chez les oomycètes, laissant en suspens la question de l'internalisation des protéines. Les travaux portant sur le motif protéique conservé RXLR présent chez les oomycètes et les motifs RXLR-like trouvés chez certains champignons apportent les premières réponses à cette question. Le motif RXLR est apparenté au motif « Pexel », un signal d'adressage à l'hôte, présent chez les protistes du genre *Plasmodium*, responsable du paludisme chez l'Homme (Marti *et al.*, 2004; Morgan et Kamoun, 2007). De la même façon que le motif Pexel est nécessaire à la translocation de protéines de l'agent pathogène dans les cellules sanguines de l'hôte (Marti *et al.*, 2004), le motif RXLR est nécessaire à la translocation des protéines d'oomycètes portant ce motif





**Figure 2 : Modèle proposé pour l'entrée des effecteurs fongiques ou d'oomycètes au sein des cellules végétales**

La liaison des effecteurs par leur domaine RXLR au PI3P, supposé localisé au sein des radeaux lipidiques, mènerait à l'entrée de ce complexe dans la cellule via endocytose. Le mécanisme par lequel les effecteurs se libèreraient ensuite des endosomes n'est pas encore connu. Adapté de Kale *et al.* (2010).

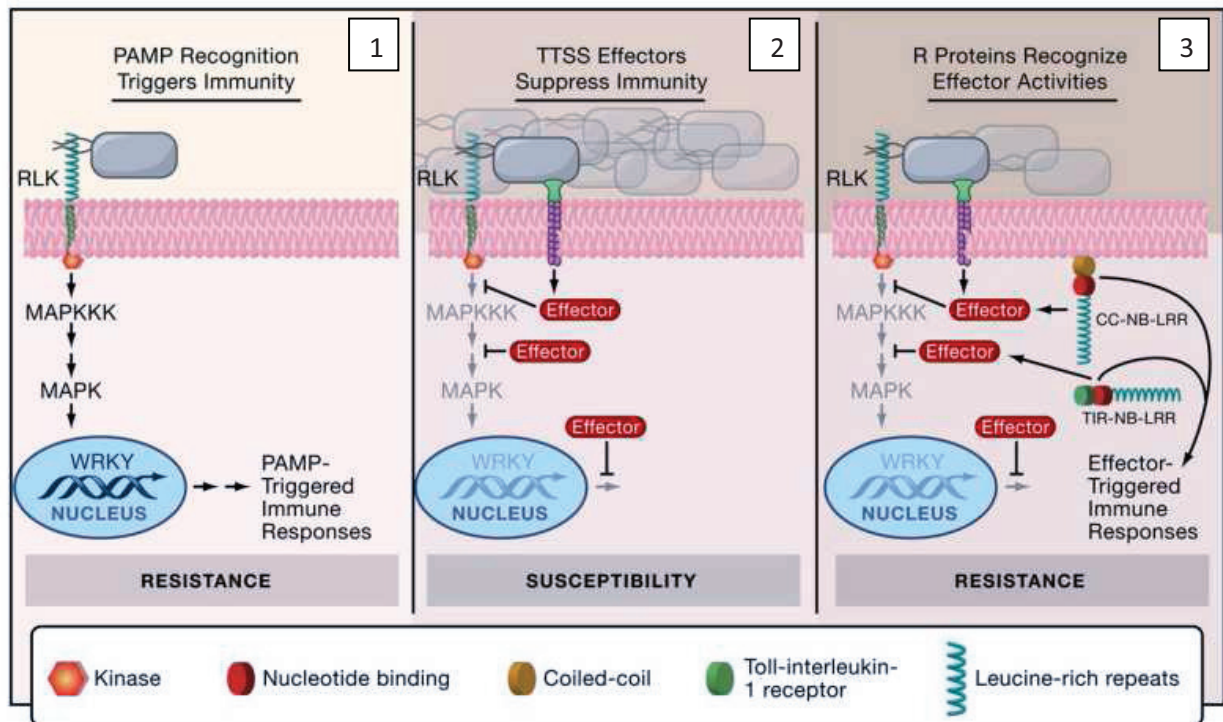
du milieu apoplastique au cytoplasme des cellules végétales (Bhattacharjee *et al.*, 2006). La mutation du motif RXLR d'une protéine suffit à empêcher son internalisation dans la cellule végétale. Ainsi, diverses mutations du motif RXLR de la protéine Avr3a chez *P. infestans* suffisent à stopper la translocation de celle-ci, provoquant l'accumulation d'Avr3a dans l'espace intercellulaire (Bhattacharjee *et al.*, 2006). Kale *et al.* (2010) ont montré que les motifs RXLR et RXLR-like sont retrouvés chez de nombreuses espèces fongiques telles *M. lini*, *F. oxysporum*, *L. maculans*, *M. Oryzae*, *Laccaria bicolor* ou encore *Aspergillus fumigatus* et que ces motifs se lient au phosphatidylinositol-3-phosphate (PI3P), un phospholipide abondant à la surface des membranes plasmidiques des cellules végétales et de certaines cellules animales. Dans le cas où la liaison avec le PI3P ne peut s'effectuer, l'internalisation de la protéine n'a pas lieu (Figure 2).

Une autre caractéristique fréquemment rencontrée chez les gènes d'avrulence fongiques ou d'oomycète, concerne leur expression lors de l'infection. Ces gènes sont souvent surexprimés, voire spécifiquement exprimés lors de l'infection, ou lors d'une phase spécifique de l'infection. Ainsi, la famille de gènes d'avrulence *AvrL567*, chez *M. lini*, est spécifiquement exprimée lors de l'infection au niveau des haustoria, structures spécialisées impliquées dans la nutrition de l'agent pathogène (Ellis *et al.*, 2007).

### c) Rôle intrinsèque des gènes d'avrulence

Si on se base sur le modèle émis par Flor, les gènes d'avrulence n'ont pas d'autre rôle biologique que le déclenchement des réactions de défense de la plante suite à leur reconnaissance. Une telle définition n'implique pour l'agent pathogène qu'un rôle délétère au gène d'avrulence et on peut dès lors se poser la question du maintien de ces gènes chez l'agent pathogène (Knogge, 1996).

La caractérisation de la fonction de gènes d'avrulence permet en partie de dénouer ce paradoxe. Comme décrit précédemment, chez *Xanthomonas sp* les gènes de la famille *AvrBs3* jouent le rôle de facteurs de transcription impliqués dans l'augmentation de taille des cellules du mésophylle de la plante hôte (Kay *et al.*, 2007). Chez *R. secalis*, le gène *NIP1* (Necrosis Inducing Peptide 1) code pour une phytotoxine qui induit une nécrose chez la plante hôte et permet ainsi au champignon un meilleur accès aux ressources nutritives de la cellule (Rohe *et al.*, 1995). Chez *C. fulvum*, *AVR4* code pour une protéine fixant la chitine et protégerait



**Figure 3 : Modèle d'évolution des interactions plantes / agents pathogènes (exemple chez les bactéries)**

De gauche à droite : **1** : la reconnaissance des PAMPs (ex : flagelline) par la plante déclenche la résistance de la plante (PTI), via une cascade de signalisation (cascades MAP kinase). **2** : la sécrétion d'effecteurs via le système de sécrétion de type III de la bactérie permet de supprimer la PTI de la plante en ciblant de multiples protéines hôtes impliquées dans les voies de signalisation ou les réactions de défense. Les bactéries peuvent dans ce cas se développer. **3** : Les protéines de résistances de la plante (ici, CC-NB-LRR et TIR-NB-LRR) détectent l'activité ou la présence des effecteurs et déclenchent les réactions de défenses de la plante (ETI). La plante est alors à nouveau résistante. Adapté de Chisholm *et al.* (2006).

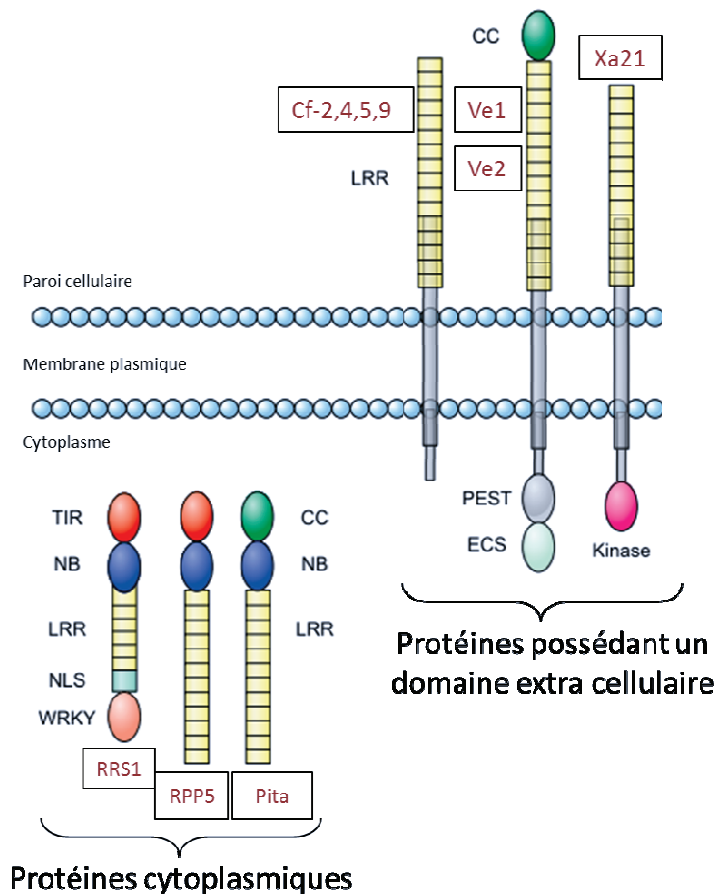
ainsi le champignon des chitinases de la plante (van den Burg *et al.*, 2003). Chez *M. lini*, *AvrP123* code pour un inhibiteur de protéase putatif et protégerait le microorganisme des protéases déployées pour défendre les cellules hôte (Catanzariti *et al.*, 2006). De la même façon, *Avr2*, chez *C. fulvum*, supprime chez la plante l'activité de la protéase *Rcr3* qui est nécessaire à l'induction des réactions de défense liées au gène de résistance *Cf2* (Rooney *et al.*, 2005).

Dans de nombreux cas, la fonction intrinsèque du gène d'avirulence n'a toutefois pas encore été établie mais les conséquences biologiques de la perte de cette fonction pour les agents pathogènes ont pu être évaluées. L'importance de la présence du gène d'avirulence pour la colonisation des plantes ou l'intensité des symptômes a ainsi été mise en évidence pour les gènes *AvrRpm1* de *Pseudomonas syringae pv. maculicola* (Ritter et Dangl, 1995), *AvrPto* de *P. syringae pv. tomato* (Shan *et al.*, 2000), ou *Six1* chez *Fusarium oxysporum* (Rep *et al.*, 2005). C'est également le cas, que nous discuterons plus loin, du gène *AvrLm4* chez *L. maculans* et dans une moindre mesure d'*AvrLm1* (Huang *et al.*, 2010).

Ces données, mises en parallèle avec les caractéristiques des protéines d'avirulence décrites plus haut (sécrétion/translocation, expression lors de l'infection, etc.), montrent que le rôle des produits des gènes d'avirulence ne se limite pas à la possibilité d'être reconnu par la plante hôte mais qu'ils interviennent dans la capacité de l'agent pathogène à infecter la plante. On parle désormais de gènes codant pour des effecteurs, c'est à dire des molécules nécessaires ou favorisant le développement d'un agent pathogène en interaction avec un hôte mais dont la fonction n'est pas requise pour le développement du champignon en dehors de la plante (Hogenhout *et al.*, 2009). Au cours de la co-évolution plante – agent pathogène, certains de ces effecteurs auraient été pris pour cible par des protéines de résistance spécifiques de la plante hôte, endossant la double casquette d'effecteur et de facteur d'avirulence. Les réactions de défense de la plante déclenchées lors de la reconnaissance d'un effecteur par le produit du gène de résistance correspondant sont regroupées sous le terme d'effector – triggered immunity (ETI) (Chisholm *et al.*, 2006; Jones et Dangl, 2006) (figure 3).

#### d) Gènes de résistance

Les gènes de résistance des plantes constituent le deuxième partenaire de la relation gène-à-gène. Leur identification et leur étude permettent, en association



**Figure 4 : Représentation schématique des différents domaines prédits chez des protéines conférant une résistance spécifique.**

Les gènes de résistance sont différenciés selon leur localisation cytoplasmique ou membranaire et en fonction des domaines protéiques qu'ils possèdent. Ici sont présentés les principales catégories de gènes de résistance accompagnés d'un ou de plusieurs exemples. RRS1, RPP5 : protéines de résistance d'*Arabidopsis thaliana* ; Cf-2, Cf-4, Cf-5, Cf-9, Ve1, Ve2 : protéines de résistance de la tomate ; Xa21, Pita : protéines de résistance du riz. Domaines protéiques: TIR : Toll interleukin receptor ; NBS : Nucleotide Binding Site ; LRR : Leucine Rich Repeat ; NLS : Nuclear Localization Sequence ; CC : Coiled-Coil ; PEST : motif Pro-Glu-Ser-Thr ; ECS : signal d'endocytose ; Kinase : domaine kinase. Adapté de Hammond-Kosack et Parker, (2003)

avec les données issues des recherches effectuées sur les gènes d'avirulence, de mieux appréhender les bases moléculaires de la relation gène-à-gène.

La majorité des gènes de résistance code pour des protéines structurellement proches, possédant une architecture modulaire faisant intervenir différents domaines fonctionnels (Martin *et al.*, 2003) (figure 4). La très grande majorité des gènes de résistance code pour des protéines cytoplasmiques, possédant un domaine NBS (Nucléotide Binding Site) impliqué dans la fixation ou l'hydrolyse de nucléotides et un domaine LRR (Leucline Rich Repet) composé d'une répétition de motifs riches en leucine et impliqué dans les interactions protéines – protéines. Cette catégorie peut être sub-divisée en deux sous-classes selon l'ajout en N-terminal de la protéine, d'un motif TIR (Toll Interleukin Receptor) impliqué dans la cascade de signalisation (ex : RPP5 d'*A. thaliana*) ou d'un motif CC (Coiled-Coil), engagé dans les interactions protéines – protéines, la liaison aux acides nucléiques et la transduction de signaux (ex : Pita chez le riz) (Buschges *et al.*, 1997; Parker *et al.*, 1997; Orbach *et al.*, 2000; Martin *et al.*, 2003) (figure 4). Certains récepteurs NBS-LRR peuvent également posséder d'autres domaines spécifiques. Ainsi, RRS1-R, protéine de résistance de *Ralstonia solanacearum* de type TIR-NBS-LRR possède en C-terminal de la protéine un domaine WRKY qui permettrait sa liaison à l'ADN (Deslandes *et al.*, 2002; Martin *et al.*, 2003) (figure 4).

La deuxième grande classe de gènes de résistance rencontrée correspond aux protéines constituées d'un domaine LRR extracellulaire (eLRR) et d'un court domaine transmembranaire (TM) (ex : gènes de *C. fulvum* Cf-2, Cf-4, Cf-5, Cf-9) (Joosten et de Wit, 1999). Comme pour les gènes de résistance intracellulaire, certains domaines supplémentaires peuvent être trouvés sur ces eLRR. La protéine Xa21 trouvée chez le riz est ainsi une eLRR à laquelle un domaine kinase est attaché en C-terminal cytoplasmique (Martin *et al.*, 2003) tandis que les protéines Ve1 et Ve2 de la tomate possèdent respectivement, un domaine CC ou un motif PEST (Pro-Glu-Ser-Thr) impliqué dans les processus d'ubiquitination (Martin *et al.*, 2003) (figure 4).

Il existe enfin un certain nombre de protéines de résistance, classées actuellement à part, du fait de leurs structures spécifiques. Hm1, protéine de résistance présente chez le Maïs confère la résistance à *Cochliobolus carbonum* et contrôle l'expression d'une protéine de détoxification (Martin *et al.*, 2003). Rpg1, protéine de résistance de l'orge contre *Puccinia graminis f. sp. tritici* possède 2

domaines kinase disposés en tandem et un domaine transmembranaire (Brueggeman *et al.*, 2002). Enfin, Xa27, présent dans le riz et conférant la résistance à *X. oryzae* pv. *oryzae* ne possède pas d'homologie avec les autres protéines de résistance (Gu *et al.*, 2005; Chisholm *et al.*, 2006). Ces protéines atypiques classées comme protéines de résistance doivent être considérées avec attention. En effet, d'autres gènes ont tout d'abord été considérés comme gènes de résistance mais la découverte de nouveaux acteurs moléculaires nécessaires à la reconnaissance de l'effecteur, a conduit à une redéfinition de leur fonction. L'exemple de Pto illustre tout particulièrement ce point. Pto est une protéine kinase présente chez la tomate qui a été initialement considérée comme la protéine de résistance correspondant à la protéine d'avorulence AvrPto sécrété par *P. syringae* pv. *tomato*. Il a par la suite été montré qu'une troisième protéine, Prf, était nécessaire au déclenchement des réactions de défense. Pto est en fait, une protéine prenant part à la reconnaissance d'AvrPto orchestrée par la véritable protéine de résistance, Prf, qui est elle, une protéine NBS-LRR typique (van der Biezen et Jones, 1998; Mucyn *et al.*, 2006).

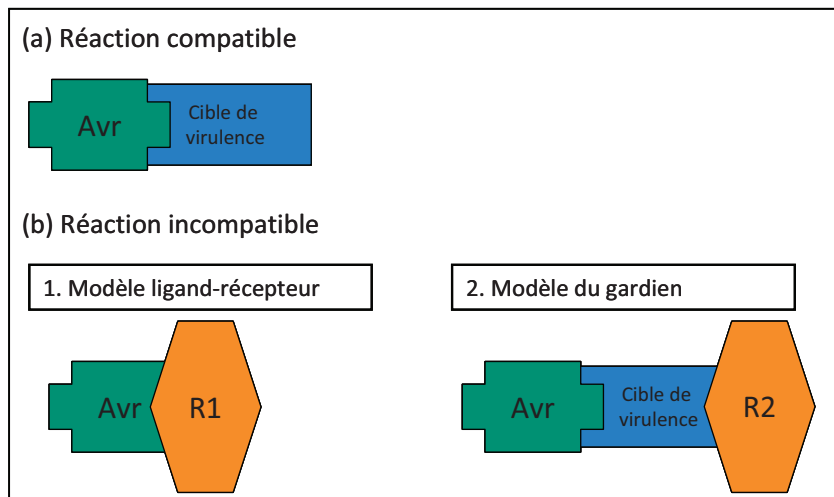
#### e) Modèles de reconnaissance entre protéines de résistance et protéines d'avorulence

Le premier et plus simple modèle développé pour expliquer la reconnaissance spécifique entre produits de gènes d'avorulence et produits de gènes de résistance propose une interaction directe entre les deux protéines, on parle alors de modèle « ligand – récepteur » (Keen, 1982). Ce modèle a été validé par exemple pour l'interaction *Avr-Pita/Pita* dans le système *M. oryzae*/riz (Jia *et al.*, 2000). Des analyses en « doubles hybrides » chez la levure et des tests de liaison *in vitro* ont montré une interaction physique (directe) entre la protéine Avr-Pita et le produit du gène de résistance *Pi-ta* trouvé chez le riz. La disruption de cette interaction physique par la mutation d'un acide aminé chez Avr-Pita suffit à empêcher la mise en place des réactions de défense (Jia *et al.*, 2000). On peut également citer l'interaction directe entre les produits des gènes issus du locus *AvrL567* de *M. lini* et les produits des gènes de résistance correspondants L5 et L6 présents chez le lin (Dodds *et al.*, 2006) ou celle entre RRS1-R d'*A. thaliana* et la protéine d'avorulence *PopP2* de *R. solanacearum* (Deslandes *et al.*, 2003).

Néanmoins, ce type d'interaction moléculaire a rarement été mis en évidence et ne permet pas de décrire la plupart des interactions protéines de résistance – protéine d'avirulence, telle l'interaction entre Cf-9 et AVR9 chez le système *C. fulvum/tomate* pour laquelle aucune liaison physique n'a pu être mise en évidence (Luderer *et al.*, 2001). Un autre modèle nommé « modèle du gardien », supposant la perception indirecte de l'effecteur via un troisième partenaire moléculaire a alors été proposé (van der Biezen et Jones, 1998). La protéine de résistance ne servirait pas à détecter la présence d'un effecteur mais plutôt à surveiller l'intégrité moléculaire de la protéine cible de cet effecteur, devenant de ce fait le « gardien » de cette protéine (Dangl et Jones, 2001) (figure 5). Chez *A. thaliana*, la protéine de résistance RPM1 est ainsi capable de détecter la présence d'AvrRPM1 et d'AvrB, deux effecteurs de *P. syringae* via une troisième protéine, RIN4. Cette protéine membranaire est hyperphosphorylée sous l'action d'AvrRPM1 ou d'AvrB, ce qui provoque une modification structurelle de RIN4, laquelle va alors être perçue par RPM1 et va déclencher l'activation des défenses de la plante (Mackey *et al.*, 2002). De la même façon, la dégradation de RIN4 par l'effecteur AvrRpt2 est perçue par la protéine de résistance RPS2 (Résistance to Pseudomonas Syringae 2) (Axtell et Staskawicz, 2003; Mackey *et al.*, 2003) (figure 6).

On voit ici que, à la différence du récepteur du modèle ligand – récepteur qui ne peut reconnaître qu'un effecteur ou un ensemble limité d'effecteurs structurellement proches, la protéine de résistance dans le rôle du « gardien » est capable de reconnaître un nombre plus important d'effecteurs non apparentés interagissant avec la cible « gardée ». Selon le modèle ligand – récepteur, la reconnaissance de tous les effecteurs chez un agent pathogène, du fait de leur grand nombre ainsi que de leur forte diversité allélique, contraindrait la plante hôte à posséder un nombre similaire et donc important de gènes de résistance. La possibilité offerte par le modèle du gardien de détecter la présence de plusieurs effecteurs en ne surveillant qu'une seule cible permet de limiter le nombre de gènes de résistance nécessaires pour couvrir l'ensemble des effecteurs des agents pathogènes. Cette stratégie serait donc moins « coûteuse » pour la plante hôte et permettrait une plus grande adaptabilité à la diversité des effecteurs rencontrés (Jones et Dangl, 2006). Jones et Dangl (2006) suggèrent que ce modèle permettrait de décrire la majorité des interactions effecteurs–protéines de résistance décrites à ce jour.





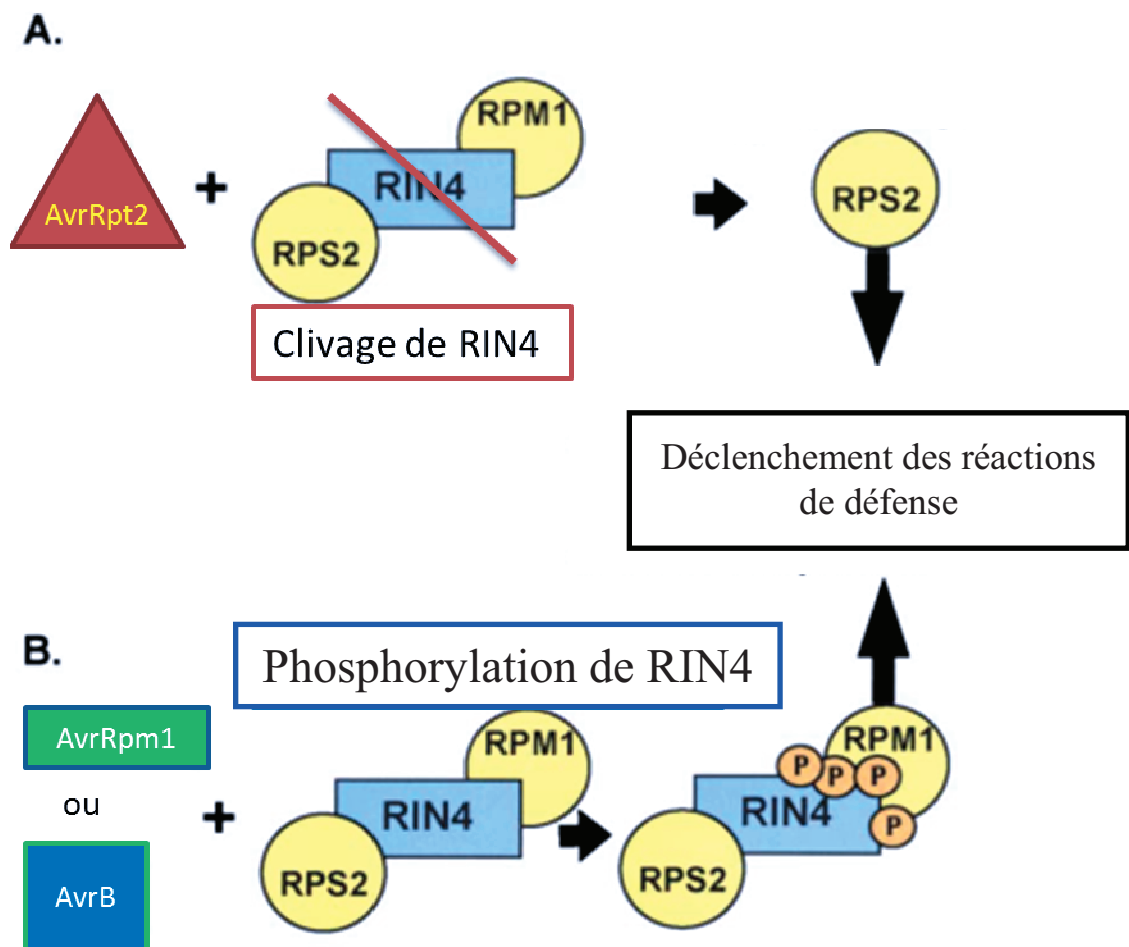
**Figure 5 : représentation schématique de la reconnaissance d'une protéine d'avirulence par une protéine de résistance selon le modèle ligand-récepteur ou selon le modèle du gardien.**

(a) Chez une plante sensible (absence de protéine de résistance), la protéine d'avirulence peut interagir avec sa cible (cible de virulence). Il n'y a pas ici de reconnaissance de l'agent pathogène par la plante et donc une réaction compatible. (b) Chez une plante résistante, la présence d'une protéine de résistance spécifique va permettre la détection de la protéine d'avirulence. Cette reconnaissance peut se faire de deux façons : (1) de façon directe entre la protéine de résistance et la protéine d'avirulence dans le cas du modèle ligand –récepteur ou (2) de façon indirecte via la surveillance de l'intégrité moléculaire de la protéine cible de la protéine d'avirulence. Adapté de Hammond-Kosack et Parker, (2003)

Toutefois, plusieurs interactions entre protéines d'avirulence et protéines de résistance ne correspondant pas au modèle « ligand – récepteur » ne peuvent également être totalement expliquées par le modèle du gardien. Selon le modèle du gardien, c'est l'interaction avec la protéine gardée, en l'absence de protéine de résistance, qui permet aux effecteurs d'avoir une fonction dans le pouvoir pathogène de la souche. Il a pourtant été montré dans certains cas que des effecteurs apportaient tout de même une amélioration du pouvoir pathogène de la souche malgré l'absence des protéines gardées correspondantes (van der Hoorn et Kamoun, 2008). Le modèle de garde pose également un problème évolutif dû à la position instable des protéines cibles partagées entre deux forces évolutives opposées. La première, en l'absence du gène de résistance correspondant, tend à faire baisser l'affinité entre la protéine cible et l'effecteur afin d'échapper à l'agent pathogène. Au contraire, en présence du gène de résistance, les plantes possédant les protéines gardées ayant la plus grande affinité avec l'effecteur seront sélectionnées du fait d'une meilleure perception de l'agent pathogène (van der Hoorn et Kamoun, 2008) (figure 7). Suite à ce constat, un nouveau modèle permettant de libérer la protéine cible de ses contraintes évolutives a été proposé : le « modèle leurre » (« guard – decoy ») (figure 8). Celui-ci s'articule autour de la duplication des cibles d'effecteurs ou de l'évolution indépendante d'une cible leurre qui serait uniquement impliquée dans la perception des effecteurs par la protéine de résistance et non dans l'induction de la maladie. Le leurre imiterait ainsi la cible de l'effecteur et piègerait celui-ci dans un système de reconnaissance où le changement de conformation du leurre dû à l'action de l'effecteur serait détecté par la protéine de résistance qui déclencherait alors les défenses de la plante hôte (van der Hoorn et Kamoun, 2008). La protéine de résistance RCR3 présente chez la tomate serait ainsi une protéine leurre pour l'effecteur AVR2 de *C. fulvum*, RCR3 prenant la place de la véritable cible PIP1, protéase sécrétée lors des réactions de défense (Shabab *et al.*, 2008).

## **B. Evolution des gènes d'avirulence vers la virulence**

### a) Contournement de gènes de résistance

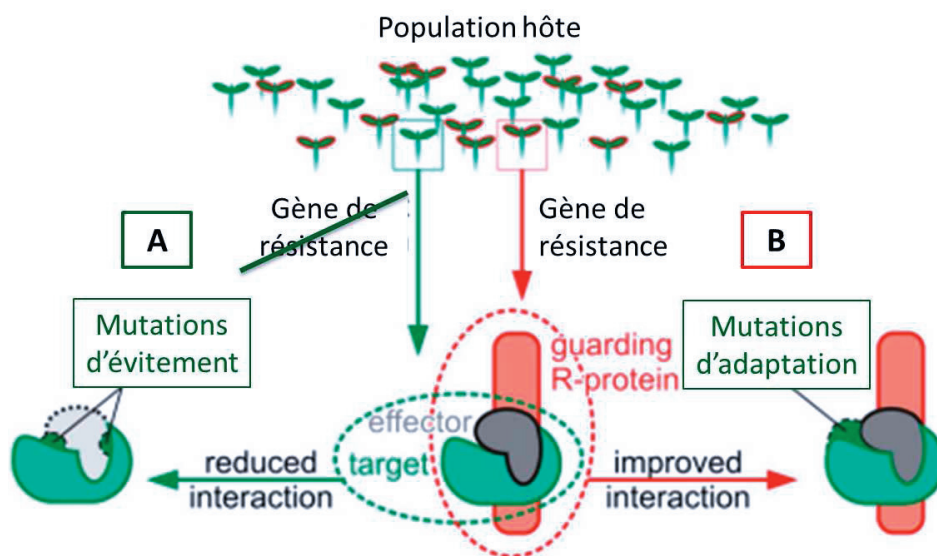


**Figure 6 : Modèle de reconnaissance indirecte d'effecteurs de *Pseudomonas syringae* par les gènes de résistance RPS2 et RPM1**

RIN4, protéine d'*Arabidopsis thaliana*, est l'intermédiaire de la reconnaissance indirecte de plusieurs protéines d'avorulence de *P. syringae* par deux protéines de résistance. **A.** L'effecteur sécrété AvrRpt2 clive la protéine RIN4 en deux. Ce clivage est reconnu par la protéine RPS2, préalablement liée à RIN4, ce qui déclenche alors les réactions de défenses de la plante. L'activité de la protéine de résistance RPM1 est quant à elle bloquée du fait de la suppression de RIN4. **B.** Les effecteurs AvrRpm1 et AvrB phosphorylent tous les deux la protéine RIN4. Cette phosphorylation est reconnue par la protéine de résistance RPM1, préalablement liée à RIN4, qui en conséquence déclenche les réactions de défense de la plante. Adapté de Axtell *et al.* (2003).

Les plantes et les agents pathogènes ont développé différents systèmes de défense et d'attaque, de reconnaissance et de camouflage, etc. Cette diversité est le résultat d'une longue co-évolution entre agents pathogènes et plantes hôtes, assimilable à une "course aux armements" (« arms race ») (Anderson *et al.*, 2010). Le but des deux partenaires étant opposé, colonisation vs protection, l'acquisition par l'agent pathogène d'une nouvelle « arme » permettant d'augmenter son efficacité vis-à-vis de la plante va favoriser sa multiplication, et donc la sélection de cette nouvelle arme, mais aussi exercer une forte pression de sélection sur la population hôte. Cette pression aura pour conséquence de sélectionner et donc de favoriser la multiplication des plantes capables de se développer malgré l'atout sélectif obtenu par l'agent pathogène. La pression de sélection se décalera une nouvelle fois sur la population d'agent pathogène, qui rentrera alors dans une nouvelle boucle de cette course aux armements. L'alternance entre PTI et effecteurs, effecteurs et ETI et enfin ETI et nouveaux effecteurs illustre cette course aux armements que Jones et Dangl, (2006) ont représentée par un modèle en ZigZag (figure 9), chaque avantage acquis par l'un ou l'autre des partenaires provoquant la sélection chez l'autre partenaire d'un mécanisme compensatoire.

Une telle course aux armements est applicable pour toutes les interactions hôte – agent pathogène mais trouve un écho tout particulier dans les interactions prenant place en milieu agricole. En effet, ici, l'hôte est un cultivar sélectionné par l'homme pour ses propriétés agronomiques parmi lesquelles la résistance aux agents pathogènes, et est, dans le cas des plantes de grandes cultures des pays « développés », cultivé en monoculture sur de grandes surfaces. Cette homogénéité exerce une très forte pression de sélection sur les populations pathogènes et provoque, dans le cas d'une résistance monogénique, la sélection de souches virulentes vis-à-vis de ce gène de résistance. La multiplication et la dispersion de ces souches virulentes seront par conséquent importantes et très rapides. Le gène de résistance n'est alors plus efficace face à la population pathogène, on parle du « contournement du gène de résistance » par l'agent pathogène (« resistance breakdown ») (McDonald et Linde, 2002). Dès lors qu'une résistance spécifique est déployée à grande échelle, son contournement constitue une forte menace et a pu être observé à de nombreuses reprises comme par exemple le contournement du gène de résistance *Vf* du pommier par certaines souches de *Venturia inaequalis* (Parisi *et al.*, 1993) ou encore le contournement successif par *Puccinia sp.* des



**Figure 7 : Forces de sélection opposées s'exerçant sur des protéines « gardées » dans le contexte d'une population hôte polymorphe pour le gène de résistance considéré**

L'application du modèle de garde dans une population hôte polymorphe pour le gène de résistance considéré (plante possédant ou dépourvue du gène de résistance considéré) impose deux forces évolutives opposées. **A.** En l'absence du gène de résistance (flèches vertes), la sélection aura tendance à favoriser les souches présentant une interaction protéine cible – effecteur faible (« target-effector ») **B.** En présence du gène de résistance, la sélection favorisera au contraire les souches possédant une plus forte interaction cible-effecteur. Adapté de Van der Hoorn et Kamoun, (2008).

gènes de résistance utilisés sur blé (*Lr1*, *Lr2*, *Lr22b*, etc.) (McIntosh et Brown, 1997). Néanmoins, à la différence d'interactions en milieu naturel, la pression de sélection sur la population hôte, induite par l'apparition de souches virulentes, est annulée par l'utilisation de cultivars n'ayant pas été soumis à cette pression. On assiste ainsi à une course aux armements non pas entre les populations d'agents pathogènes et leurs hôtes mais entre les populations d'agents pathogènes et l'homme à travers une répétition de cycles de déploiement de cultivars résistants (« the boom ») et d'adaptation des populations pathogènes à ces résistances (« the bust »), aussi appelés « boom and bust cycles » (McDonald et Linde, 2002). De tels cycles aboutissent à un gaspillage irréversible de ressources génétiques parfois rares.

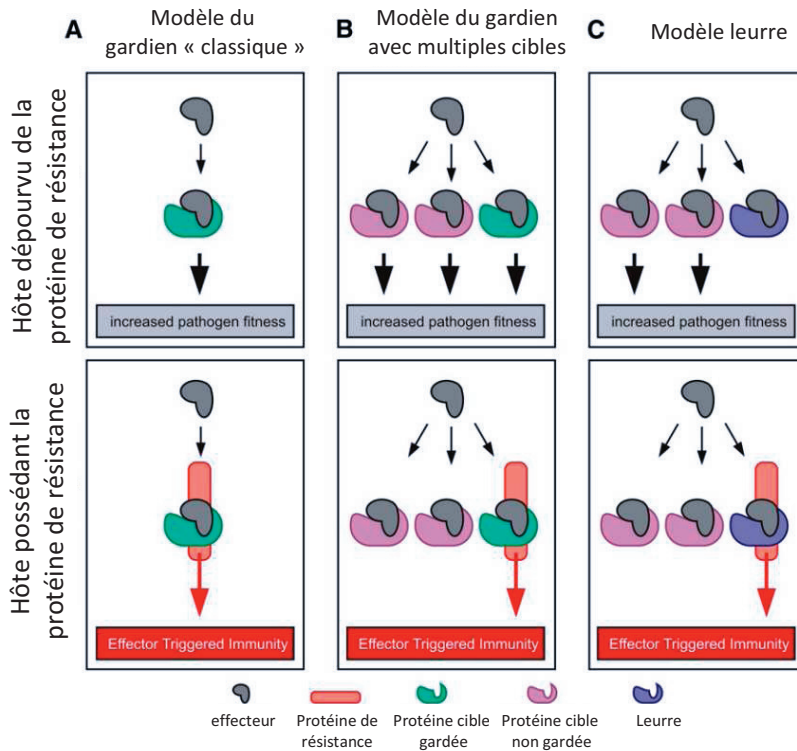
#### b) Mécanismes moléculaires à l'origine du contournement de la résistance

Quelques études portant sur différents contournements de gènes de résistance ont été effectuées et ont permis de répertorier plusieurs mécanismes moléculaires menant à la virulence de l'agent pathogène.

Le premier de ces mécanismes correspond à la génération de mutations ponctuelles dans la séquence du gène d'avirulence considéré. Ces mutations sont regroupées selon trois catégories : substitution, insertion ou délétion d'un ou plusieurs nucléotides dans la séquence codante.

La substitution d'un nucléotide par un autre dans le cas d'une mutation non synonyme (mutation visible au niveau protéique) peut altérer la fonctionnalité de la protéine, sa reconnaissance ou encore sa stabilité (structure tertiaire). La reconnaissance d'AVR4 chez *C. fulvum* par la protéine de résistance Cf-4 est ainsi altérée par la déstabilisation de la protéine AVR4 provoquée par le remplacement de certaines cystéines de la protéine (van den Burg *et al.*, 2003).

L'insertion ou la délétion d'un ou de plusieurs nucléotides au sein de la séquence codante d'un gène provoque un décalage du cadre de lecture qui aura pour conséquence la synthèse d'une protéine non fonctionnelle qui, de ce fait, ne pourra pas être reconnue par le gène de résistance correspondant. La virulence d'une souche de *C. fulvum* vis-à-vis de *Cf2* a ainsi été obtenue suite à la délétion d'un nucléotide de la séquence du gène *AVR2*. Cette mutation a provoqué, du fait du décalage du cadre de lecture, l'apparition d'un codon stop dans la séquence codante du gène d'avirulence, donnant lieu à la synthèse d'une protéine tronquée (Luderer *et al.*, 2002).



**Figure 8 : comparaison du modèle du gardien et du modèle leurre**

**A.** Modèle du gardien « classique » : un effecteur agit sur une cible dont l'intégrité moléculaire est surveillée par une protéine de résistance. **B.** Un effecteur peut avoir différentes cibles au sein de l'hôte. Dans ce modèle, une des différentes cibles de l'effecteur est gardée par une protéine de résistance. **C.** Dans le modèle leurre, la protéine gardée n'a pas d'activité favorable pour l'agent pathogène en l'absence de protéine de résistance, elle est uniquement impliquée dans la perception des effecteurs par la protéine de résistance et non dans l'induction de la maladie. Adapté de Van der Hoorn et Kamoun, (2008).

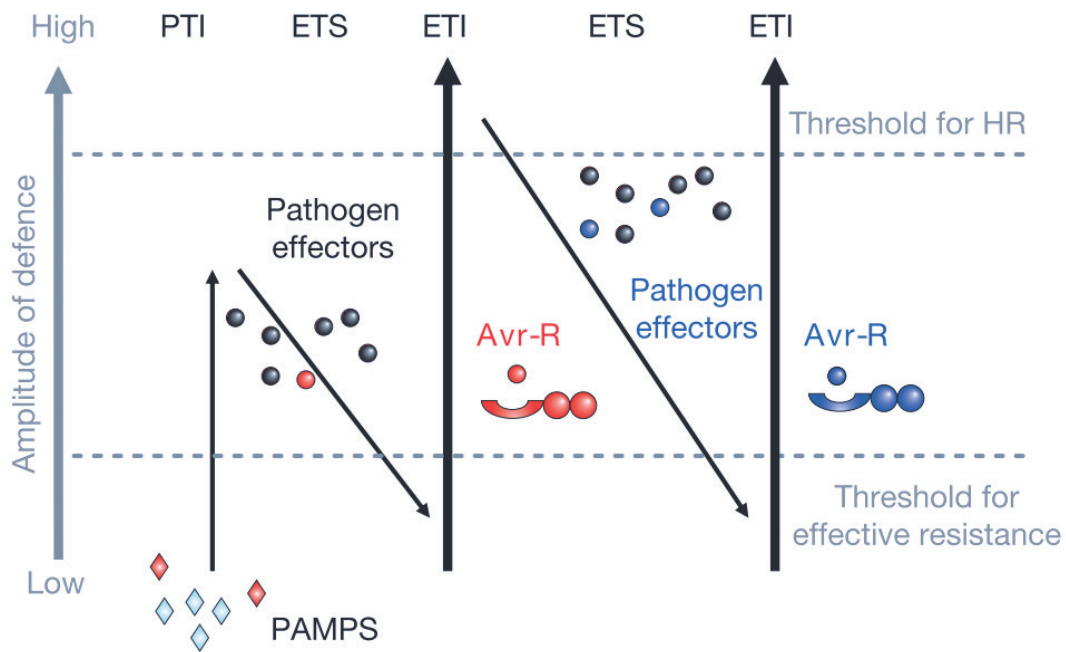
Le deuxième mécanisme moléculaire rencontré menant au contournement d'un gène de résistance, consiste en la perte du gène d'avirulence correspondant par délétion. L'ampleur de la région génomique déléetée est assez variable. De faible, avec des délétions de 100pb affectant *Avr-Pita* chez *M. oryzae* (Orbach *et al.*, 2000) ou encore relativement faible avec la délétion de 12kb associée à la perte du gène *SIX1* chez *F. oxysporum* (Rep *et al.*, 2004), ces délétions peuvent concerner de très grandes régions génomiques, telle la délétion de 260 kb comprenant le gène *AvrLm1* chez *L. maculans* (Gout *et al.*, 2007).

Un troisième évènement moléculaire menant à la virulence correspond à l'inactivation du gène d'avirulence via l'insertion d'un élément transposable dans la séquence. Cet évènement a été retrouvé par exemple dans le cas de la perte d'avirulence due à ACE1 chez *M. oryzae* (Fudal *et al.*, 2005) ou encore dans le contournement du gène de résistance *Pi-ta* chez *M. oryzae* par l'insertion du transposon Pot3 dans le promoteur d'*Avr-Pita* (Kang *et al.*, 2001 ; Zhou *et al.*, 2007).

Le positionnement des gènes d'avirulence en région génomique instable semble être un facteur potentiellement important pour le contournement des gènes de résistance. Les délétions d'*AvrLm1* ou d'*AvrLm6* chez *L. maculans* semblent ainsi liées à leur localisation dans un environnement supposé instable car peuplé d'éléments transposables répétés pouvant provoquer mésappariement et délétion à la méiose (Gout *et al.*, 2007; Fudal *et al.*, 2009). De la même façon, la délétion d'*AvrPphB* chez *P. syringae* pv. *phaseolicola* s'expliquerait par sa position au sein d'un îlot de pathogénie (Pathogenicity Island, PAI), région génomique elle aussi particulièrement instable (Pitman *et al.*, 2005). La localisation en région télomérique de certains gènes tel *Avr-Pita* chez *M. oryzae*, semble également favoriser la délétion des gènes (Farman, 2007).

A l'évolution des protéines d'avirulence via différents mécanismes moléculaires, s'ajoute l'illustration directe du modèle zigzag avec la mise en place de nouveaux effecteurs par l'agent pathogène. Dans le cas de *F. oxysporum*, ce cas de figure est observé pour le contournement des gènes de résistance *I-2* et *I-3*, du à la présence d'un troisième effecteur nommé *Avr1*. *I-2* et *I-3* permettent à la plante hôte, dans le cas de souches avirulentes, de détecter, selon le modèle gène-à-gène, la présence des deux effecteurs correspondant, *Avr2* et *Avr3* et ainsi de déclencher les réactions de défense de la plante. Dans le cas d'une souche possédant *Avr1*, du fait





**Figure 9 : Modèle en Z (zigzag)**

Chaque avantage acquis par l'un ou l'autre des partenaires d'une interaction hôte – agent pathogène provoque la sélection chez l'autre partenaire d'un mécanisme compensatoire. Les PAMPS déclenchent la PTI qui est potentiellement supprimée par certains effecteurs. Ces mêmes effecteurs ou leurs actions sont reconnus par des protéines de résistance qui déclencheront les réactions de défense de la plante. L'action de ces effecteurs est elle-même annulée par d'autres effecteurs qui seront à leur tour potentiellement cibles d'autres gènes de résistance. PTI: PAMP Triggered Immunity; ETS: Effector Triggered Sensibility; ETI: Effector Triggered Sensibility. Adapté de Jones et Dangl, (2006).

d'un mécanisme encore inconnu, la reconnaissance *Avr2-12* et *Avr3-13* ne s'effectuent plus et l'interaction devient compatible (Houterman *et al.*, 2008).

#### c) Coût de la virulence / déficit de fitness

La perte de la fonction d'avirulence d'un gène se traduit par une diminution du pouvoir pathogène liée à la fonction d'effecteur postulée pour tous les gènes d'avirulence. Cette fonction, quoique dispensable à la survie du microorganisme, lui donne un avantage sélectif sur les souches ne le possédant pas. La perte de ce gène a donc un « coût » pour l'agent pathogène qui va influencer sa capacité à survivre et à se reproduire (« Fitness »). On parle ici du « cout de la virulence » (Leach *et al.*, 2001). Ce coût de la virulence ou, « déficit de fitness », est un concept très fréquemment utilisé dans le domaine de la recherche agronomique mais qui n'a été que rarement démontré, principalement du fait de l'importance plus ou moins élevée de ce coût en fonction du rôle ou de la redondance de l'effecteur considéré (Leach *et al.*, 2001). L'étude de Huang *et al.* (2006) sur le gène d'avirulence *AvrLm4* de *L. maculans* constitue un bon exemple du coût de virulence associé à la perte d'un gène d'avirulence. En conditions expérimentales sur cultivars sensibles, les souches possédant l'allèle *AvrLm4* induisent des lésions en plus grand nombre et de diamètre plus important, une plus grande production d'inoculum secondaire ainsi qu'une meilleure colonisation des tissus végétaux que les souches possédant l'allèle virulent. L'allèle avirulent confère donc un avantage sélectif à la souche considérée et son absence se traduit par un déficit de fitness.

#### d) Influence du coût de virulence

Le coût de virulence va avoir une répercussion directe sur le maintien des souches virulentes dans la population considérée (Leach *et al.*, 2001). Dans le cas d'un environnement où la population hôte exposée aux agents pathogènes n'est que rarement résistante, un fort coût de virulence aura tendance à empêcher la dispersion de l'allèle virulent dans la population pathogène, gardant ainsi le taux de souches virulentes à un faible niveau. A l'inverse, dans un environnement où les hôtes sont majoritairement, voir exclusivement résistants, comme ils peuvent l'être dans les parcelles cultivées, seules les souches virulentes pourront se développer et se disperser malgré un coût de virulence élevé. Néanmoins, les souches avirulentes seront avantagées sur les hôtes sensibles présents et pourront ainsi se maintenir

dans la population pathogène à un taux dépendant de la proportion d'hôtes sensibles. Dans le cas d'une augmentation de la proportion d'hôtes sensibles, les souches avirulentes pourront profiter de leur avantage sélectif et se disperser de nouveau. Dans le cas d'*AvrLm4* chez *L. maculans*, l'utilisation en France pendant près de 40 ans de cultivars possédant la résistance *Rlm4* a provoqué l'émergence de souches virulentes mais le fort coût de virulence associé à l'allèle virulent ainsi qu'une utilisation faible mais continue de cultivars sensibles ont permis à l'allèle avirulent de se maintenir dans environ 9% des souches françaises de *L. maculans* (Balesdent *et al.*, 2006).

#### e) Durabilité d'une résistance

A partir du moment où une résistance se trouve contournée, la durabilité de celle-ci peut alors être évaluée. Il s'agit de la période durant laquelle une résistance est restée efficace au champ malgré une utilisation intensive de cultivars portant cette résistance dans des conditions favorables aux attaques des agents pathogènes correspondants (Johnson, 1984). La durabilité d'une résistance est une notion qui, de par sa définition, ne peut être établie qu'a posteriori et dont la prédiction est rendue difficile du fait d'un important nombre de paramètres à prendre en compte.

Un des paramètres influençant fortement la durabilité d'une résistance est la nature de cette résistance. En effet les résistances polygéniques ou quantitatives mettent en jeu plusieurs gènes impliqués dans la résistance de la plante et sont de ce fait supposées plus durables que les résistances spécifiques. La perte d'efficacité d'une résistance polygénique requiert le contournement de plusieurs gènes de résistance et serait donc moins probable du fait du plus grand nombre de mutations nécessaire que dans le cas du contournement d'un unique gène de résistance.

De la même façon, le coût de virulence associé au contournement d'une résistance est un facteur primordial, plus ce coût sera important moins les souches virulentes seront aptes à se développer et donc à se propager : plus le coût de virulence est élevé, plus la résistance sera durable (Leach *et al.*, 2001). Dans le cas d'une résistance polygénique, le coût de virulence est d'autant plus important qu'il est le résultat de l'addition du coût de virulence associé à chaque gène contourné.

La pression de sélection exercée sur une population pathogène est un autre point important à prendre en compte dans la probabilité de voir une résistance contournée. Cette pression est en grande partie dépendante de l'importance du

déploiement des cultivars possédant la résistance et du nombre d'années d'utilisation de ces cultivars. Les résistances quantitatives imposeraient également une pression de sélection moindre qu'une résistance spécifique.

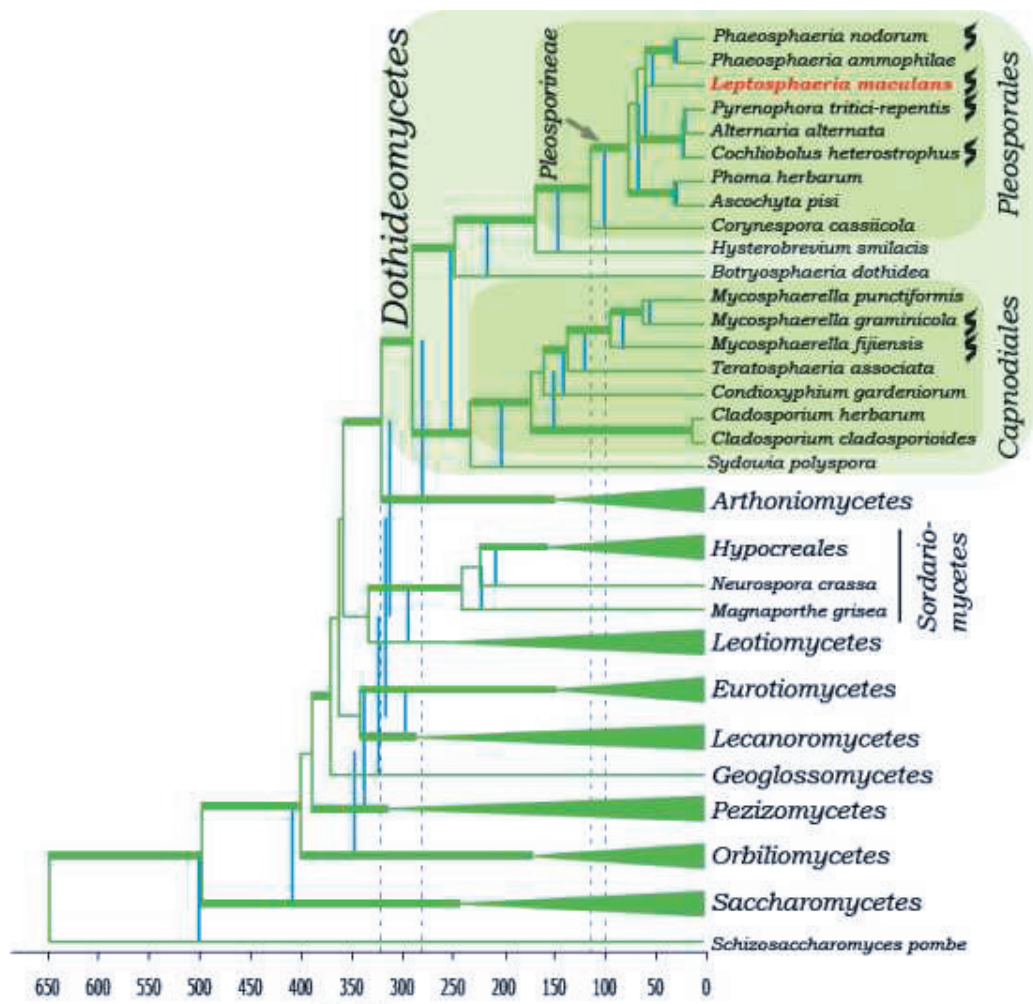
Enfin, la durabilité d'une résistance est également hautement variable en fonction du pathosystème considéré. En effet, en fonction de la capacité de dispersion de l'agent pathogène et de son potentiel évolutif, principalement lié au mode de reproduction de celui-ci, le risque que l'agent pathogène contourne la résistance empêchant son développement va être plus ou moins important (McDonald et Linde, 2002).

Aujourd'hui, l'un des principaux objectifs dans la gestion des résistances variétales est de pouvoir éviter le plus longtemps possible le contournement des gènes de résistance spécifiques utilisés et donc d'augmenter la durabilité de ces résistances. Pour se faire, différentes stratégies jouant sur les critères influençant la durabilité d'une résistance existent.

Une des solutions possible est de pratiquer l'alternance de cultivars portant différentes résistances dans le temps mais aussi dans l'espace afin de ne pas soumettre les populations à une unique pression de sélection constante dans le temps. La culture de mélanges variétaux est une autre solution possible. L'agent pathogène ne pouvant se développer que sur les cultivars sensibles, la perte d'inoculum liée à la rencontre d'hôtes résistant permet de ralentir les épidémies (Mundt, 2002).

Une autre stratégie possible consiste à utiliser des cultivars possédant plusieurs gènes de résistance majeurs (i.e. pyramidage de gène de résistance). Par cette méthode, la probabilité que tous les gènes de résistance soient contournés est d'autant plus faible que le nombre de gènes de résistance est élevé. Enfin, une autre méthode permettant en théorie de garder une résistance fonctionnelle consiste en l'association de résistances spécifiques à un cultivar possédant un fort niveau de résistance général (Brun *et al.*, 2010).

Afin de comprendre l'évolution des résistances dans le temps et de garder le cultivar le plus longtemps possible résistant, la modélisation de l'évolution du pathosystème étudié est devenue de plus en plus importante (Bergelson *et al.*, 2001). Cette modélisation a pour but de déterminer la durabilité d'une résistance et donc l'intérêt de son utilisation mais également l'évolution de cette durabilité en fonction des stratégies de culture potentiellement utilisées. Il s'agit ainsi de



**Figure 10 : Arbre phylogénétique et chronogramme concernant les ascomycètes avec une attention particulière portée aux dothidéomycètes**

Le sigle ADN visible devant certaines espèces spécifie celles dont les données génomiques sont disponibles. L'échelle de temps est indiquée en million d'années. Adapté de Rouxel *et al.* (2011).

déterminer l'intérêt d'une résistance et la meilleure stratégie à utiliser afin de conserver son efficacité.

## **C. Couple *Brassica napus* – *Leptosphaeria maculans***

### **1. *Brassica napus***

Le colza, *Brassica napus* var. *oleifera* est un allotétraploïde ( $2n=38$ , AACCC) de la famille des Brassicaceae résultant de l'allopolyploidisation naturelle de deux espèces diploïdes de Brassica : *B. oleracea* (chou,  $2n=18$ , CC) et *B. rapa* (navette,  $2n=20$ , AA) (U, 1935). Le colza est la principale espèce de Brassica cultivée dans le monde. Sa culture se situe en zone tempérée avec comme principaux producteurs, la Chine, l'Europe, le Canada et l'Australie. Pour l'année 2008-2009, les surfaces cultivées en France représentaient 1 471 000 hectares pour 5,6 millions de tonnes produites (données CETIOM).

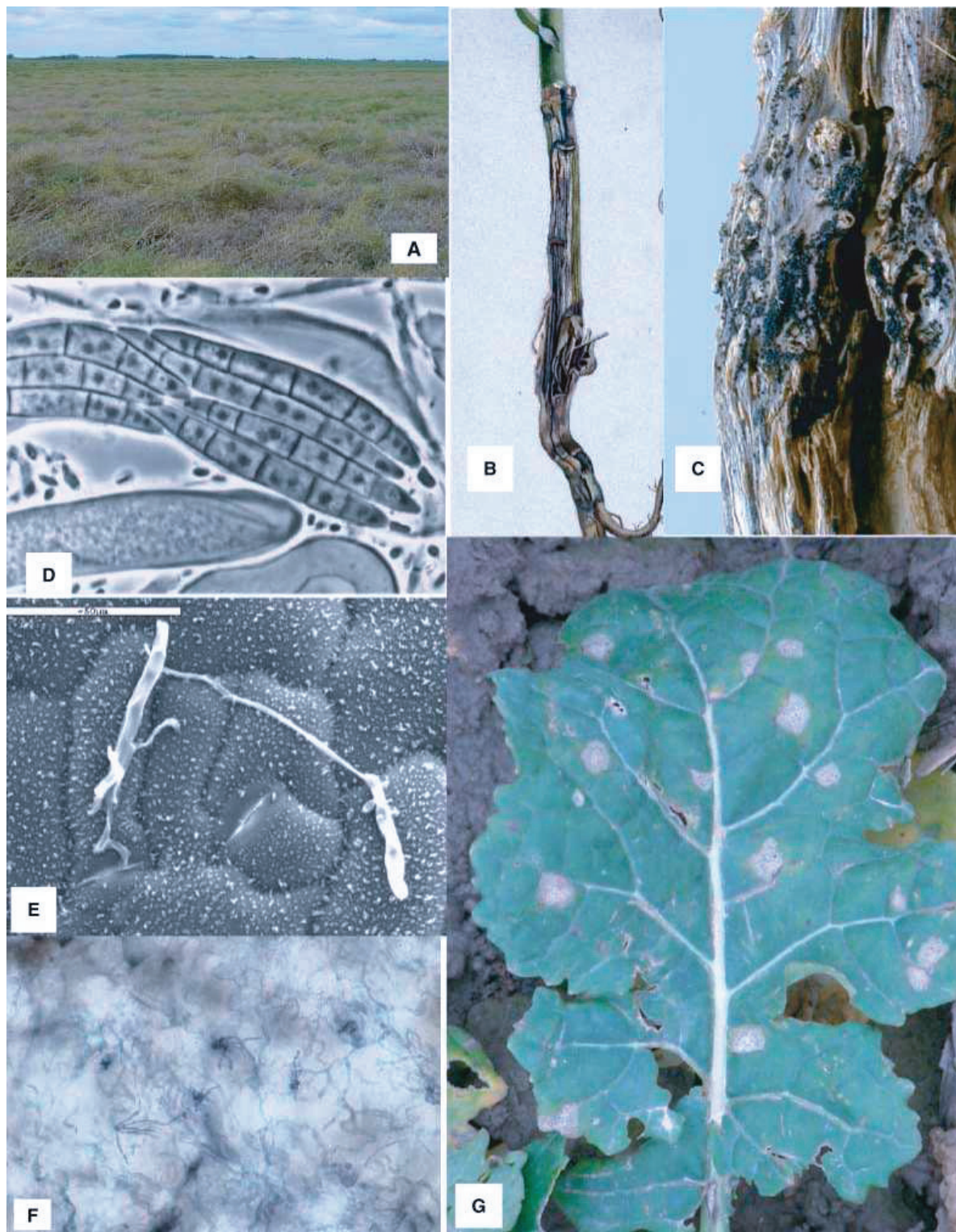
Les cultures de colza se destinent principalement à la production d'huile à usage alimentaire ou industrielle ainsi qu'à l'alimentation animale de par la production de tourteaux à forte teneur protéique. Depuis quelques années, les cultures de colza sont également destinées à la production de biocarburants. L'huile produite à cet usage est utilisée, soit directement comme carburant (relativement anecdotique), soit, après transformation en ester méthylique d'huile végétale (Diester), comme adjuvant du gazole (données CETIOM).

Les cultures de colza sont susceptibles de subir chaque année des attaques d'insectes et de microorganismes. Deux espèces fongiques sont particulièrement importantes et sous surveillance : *Sclerotinia sclerotiorum*, agent de la pourriture blanche, et *L. maculans*, agent de la nécrose du collet des crucifères. *L. maculans* représente, après les insectes, le plus gros risque biotique pour les cultures de colza et provoque des dégâts se traduisant par des pertes de rendement estimées entre 5 et 20% de la production nationale (données CETIOM).

### **2. *Leptosphaeria maculans***

#### **a. Description / taxinomie / complexe d'espèce**

*L. maculans*, forme parfaite de *Phoma lingam*, est un champignon filamenteux ascomycète de la classe des Dothidéomycetes et de l'ordre des Pléosporales, sous ordre des Pleosporineae (figure 10). Le sous ordre des Pléosporinae, différencié il y



**Figure 11 : Symptômes et phases de la maladie due à *L. maculans***

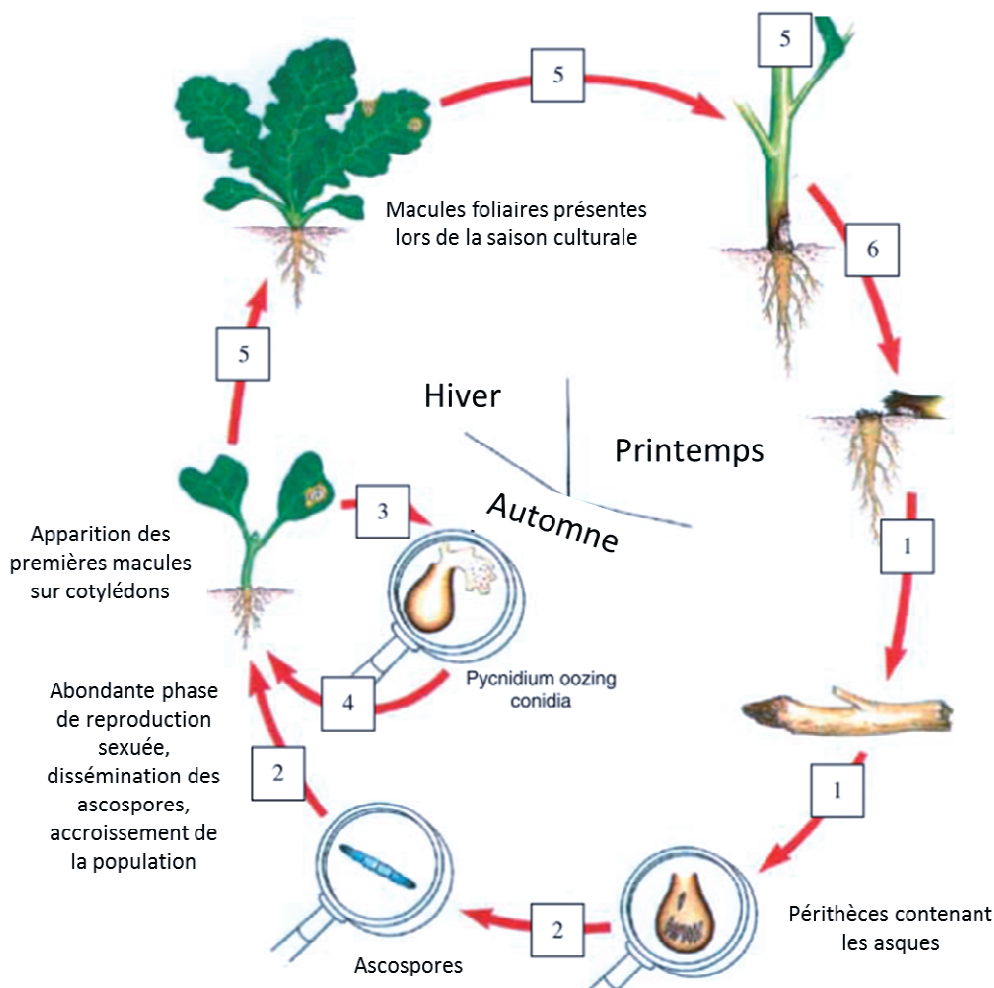
**A.** Verse parasitaire d'une culture de colza due à une attaque de *L. maculans* ; **B-C.** Nécrose des tissus du collet de plantes infectées, développement de périthèces sur des résidus de cultures 9 mois après la récolte ; **D.** Observation au microscope d'asques typiques contenant huit ascospores pluricellulaires ; **E.** Observation au microscope électronique à balayage d'ascospores germant sur des tissus de *Brassica napus* ; **F.** Colonisation stricte des espaces intercellulaires du mésophylle par du mycélium de *L. maculans* (coloration au bleu de trypan) ; **G.** Macules foliaires sur feuilles de *B. napus* résultant d'infections primaires naturelles au champ. Adapté de Rouxel et Balesdent, (2005).

a environ 100-120 millions d'années regroupe la plupart des agents phytopathogènes de l'ordre des pléosporales importants comme par exemple *Phaeosphaeria* (*Stagonospora*) *nodorum*, *Cochliobolus heterostrophus*, *Alternaria alternata*, ou encore *Pyrenophora tritici-repentis* (Schoch *et al.*, 2009). *L. maculans* est responsable de la nécrose du collet des crucifères, également appelé Phoma du colza (« Stem canker » ou « Blackleg disease » en anglais) qui provoque en fin de saison culturale la verse parasitaire de la plante (figure 11). *L. maculans* est un agent pathogène se développant sur les plantes de la famille des Brassicaceae (ex crucifère) et plus particulièrement sur les plantes du genre *Brassica* (Rouxel et Balesdent, 2005). Parmi les plantes crucifères sur lesquelles peut se développer *L. maculans*, Bohman *et al.* (2004) ont montré que la plante modèle *A. thaliana* pouvait être un hôte potentiel en condition de laboratoire.

Le nom *L. maculans* représente en fait un complexe d'espèces partagé entre deux espèces selon différents critères biologiques, morphologiques, biochimiques, ou moléculaires : *L. maculans* et *L. biglobosa* (Williams et Fitt, 1999; Shoemaker et Brun, 2001). *L. maculans*, est subdivisée en 2 sous espèces (« brassicae » et « lepidii ») tandis que *L. biglobosa* est quant à elle composée de 6 sous espèces («brassicae, canadensis, erysimii, australensis, occiaustralensis et thlaspii») (Mendes-Pereira *et al.*, 2003; Vincenot *et al.*, 2008).

*L. maculans* et *L. biglobosa* possèdent le même cycle de vie, se développent sur les mêmes hôtes et peuvent être retrouvées simultanément sur la même plante. Elles sont toutes les deux capables de provoquer des lésions nécrotiques et de coloniser les tissus des plantes hôtes mais leurs impacts sur les cultures ne sont pas similaires, seul *L. maculans* est susceptible de provoquer la verse parasitaire et donc d'être dommageable pour les cultures (Williams et Fitt, 1999; West *et al.*, 2002; Rouxel *et al.*, 2004). De nombreuses interactions spécifiques (interactions gène-à-gène) ont été montrées entre *L. maculans* et ses différentes plantes hôtes tandis qu'aucune interaction de ce type n'a encore été décrite chez *L. biglobosa* (Vincenot *et al.*, 2008). Les 2 espèces sont le plus souvent co-existantes dans les pays producteurs de colza mais seule la Chine et certains pays de l'Europe de l'est, touchés par *L. biglobosa*, semblent pour l'instant épargnés par *L. maculans* (Fitt *et al.*, 2008). L'ouverture des marchés à l'international ainsi que la possible propagation de la maladie du fait du réchauffement climatique (amélioration des conditions de





**Figure 12 : Cycle biologique de *L. maculans* en Europe**

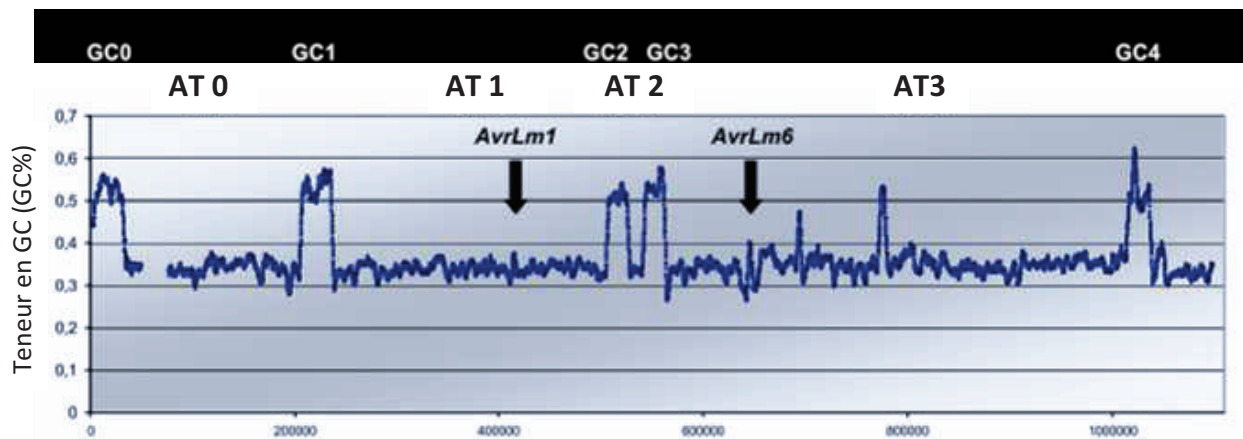
1. Développement saprophyte sur des résidus de cultures permettant la survie du champignon entre deux cultures de colza ; 2. Reproduction sexuée, génération de l'inoculum primaire ; 3-4. Courte phase nécrotrophe faisant suite à l'infection primaire due aux ascospores. Etablissement de macules foliaires, premiers symptômes de l'infection. Au cours de cette phase s'effectue la reproduction asexuée, source d'inoculum secondaire ; 5. Phase endophyte, colonisation des tissus de la tige et du collet ; 6. Phase nécrotrophe à l'origine de la verse parasitaire. Adapté de Rouxel et Balesdent, (2005).

croissance du champignon) sont pour ces pays source de craintes quant au futur développement de *L. maculans* sur leurs territoires (Fitt *et al.*, 2008).

#### b. Epidémiologie / Cycle de vie

*L. maculans* est un agent phytopathogène hémibiotrophe possédant un cycle de vie complexe au cours duquel se succèdent plusieurs modes de nutrition (figure 11 et 12). Il est ainsi successivement saprophyte, nécrotrophe, biotrophe et enfin à nouveau nécrotrophe (Rouxel et Balesdent, 2005).

Son cycle de vie (West *et al.*, 2001) débute tout d'abord par un développement **saprophytique (étape 1 du cycle)** sur les résidus de tiges de colza sur lesquelles il peut survivre 2 à 3 années. Cette phase est particulièrement importante pour le champignon puisqu'elle permet la reproduction sexuée du champignon. Celle-ci est primordiale car elle conduit à la production d'ascospores, inoculum primaire, principal moyen de dispersion de *L. maculans* (West *et al.*, 2001). En Europe, les ascospores sont libérées par le champignon à partir de la fin de l'été, période correspondant à la période de levée des pieds de colza d'hiver. Une fois en contact avec un cotylédon ou une feuille de colza les ascospores germent et pénètrent dans les tissus via les stomates ou des blessures. Après quelques jours de colonisation des tissus du mésophylle sans expression visible de symptômes, le champignon devient **nécrotrophes (étape 2 du cycle)**. Le développement du champignon à ce stade provoque la mort des cellules infectées, formant alors une macule foliaire. Ces macules sont le lieu de formation de pycnides, organes contenant des pycnidiospores (conidies) issues de la reproduction asexuée et source d'inoculum secondaire. Les infections secondaires se réalisent sur de courtes distances par la dispersion de ces conidies due à la pluie (rain splash) (Travadon *et al.*, 2007). En Europe, la formation de macules n'est que rarement préjudiciable pour la plante, qui peut poursuivre sa croissance, mais en Australie, le stade initial du cycle peut provoquer la mort de la plante et être ainsi à l'origine d'une fonte des semis (« damping off ») potentiellement importante (West *et al.*, 2001). A ce stade, *L. maculans* devient **endophyte (étape 3 du cycle)** et colonise les tissus du pétiole puis de la partie inférieure de la tige sans causer de dégâts apparents. En fin de période culturale (fin du printemps, début de l'été), *L. maculans* redevient nécrotrophe (**étape 4 du cycle**) et provoque la nécrose des tissus du collet. C'est



**Figure 13 : exemple de l'organisation en isochores du génome de *L. maculans***

L'ensemble du génome de *L. maculans* est segmenté en isochores définis selon leur composition en nucléotide. La région génomique de 1,1 Mb présentée ici présente cinq isochores équilibrés en G+C : GC0 à GC4 et quatre isochores pauvres en G+C : AT0 à AT3. Les gènes *AvrLm1* et *AvrLm6* sont isolés au milieu des isochores AT1 et AT3, respectivement. D'après Gout *et al.* (2006) et Fudal *et al.* (2007).

cette nécrose, qui en fragilisant la base de la plante, est susceptible de provoquer la verse de la plante.

### c. Caractéristiques du génome de *L. maculans*

Le génome de *L. maculans* vient d'être séquencé et les premières analyses effectuées sur celui-ci ont montré quelques caractéristiques singulières à ce champignon (Rouxel *et al.*, 2011).

*L. maculans* possède un génome de 45.12 Mb subdivisé en 17 ou 18 chromosomes comportant 12469 gènes prédits. Ce génome est plus grand que ceux des autres dothidéomycètes séquencés à ce jour et comporte une grande proportion d'éléments transposables qui semblent avoir envahi le génome. Ces éléments répétés représentent environ 1/3 du génome, contre 7% chez *P. nodorum*.

On observe au sein du génome de *L. maculans*, une structure mise en évidence au sein des génomes de mammifères et de divers autres eucaryotes (Gautier, 2000) qui consiste en une alternance de longs segments d'ADN de composition nucléotidique homogène, appelés isochores (Eyre-Walker et Hurst, 2001) (figure 13). *L. maculans* est le premier champignon pour lequel la présence d'isochores a été décrite. Deux catégories d'isochores sont rencontrées dans ce génome, des isochores possédant un taux de guanine et cytosine (taux de GC) équilibré (51% de GC) appelés isochores GC et des isochores possédant un taux de GC relativement faible (33,9% de GC) appelés isochores AT. Le génome de *L. maculans* est une mosaïque composée d'une alternance de ces deux types d'isochores avec 399 isochores GC et 413 isochores AT (Rouxel *et al.*, 2011).

Les isochores GC regroupent 95% des gènes prédits et ne comprennent que très peu d'éléments répétés tandis que les isochores AT, qui représentent 36% du génome ne regroupent que 5% des gènes prédits, le reste des isochores AT étant composé quasiment exclusivement de mosaïques d'éléments répétés tronqués et affectés par le RIP (Repeat Induced Point mutation) (Rouxel *et al.*, 2011). Le RIP est un phénomène préméiotique spécifique des champignons affectant les régions génomiques répétées de plus de 400 pb. Le RIP induit dans les séquences affectées, des transitions de G vers A et C vers T conduisant à l'inactivation des gènes touchés, le plus souvent par génération dans la séquence codante de codons STOP (Selker, 1990; Idnurm et Howlett, 2003). Parmi les gènes présents dans les isochores AT, on constate un enrichissement en gènes codant pour des petites

protéines sécrétées putatives, protéines présentant des caractéristiques d'effecteur : sur-expression lors de l'infection, pas de domaine conservé connu ou d'homologues chez d'autres champignons, richesse de la protéine en cystéines. Ces isochores AT semblent être des « niches » à effecteurs ayant un rôle spécifique dans l'expression et le contrôle de ces effecteurs (Parlange, 2006; Fudal *et al.*, 2007).

### 3. Interaction *L. maculans* - *B. napus*

Afin de protéger les cultures contre les attaques de *L. maculans*, différentes méthodes de lutte sont disponibles.

La lutte chimique par traitement fongicide contre *L. maculans* est efficace mais très peu utilisée à ce jour en France du fait du coût qui lui est associé, du manque d'efficacité qu'elle assure contre les infections par les ascospores et également du fait de la volonté croissante d'abaisser le plus possible la quantité d'intrant dans les cultures.

Un deuxième moyen de lutte contre *L. maculans* consiste en une gestion agronomique réfléchie des parcelles de culture (lutte intégrée) basée sur une optimisation des pratiques culturales reposant sur les connaissances acquises sur la biologie de *L. maculans*. Néanmoins, les pratiques associées à la lutte intégrée sont en opposition avec l'évolution récente des pratiques culturales. Ainsi, la pratique du labour qui permettrait une meilleure gestion des résidus de culture par la limitation de la dissémination des ascospores et donc l'exposition des plantes aux infections (West *et al.*, 2001) fait place de plus en plus souvent à des techniques de travail simplifié du sol. Seul le respect d'un calendrier précis des dates de semis permettant théoriquement de limiter la probabilité d'avoir un maximum d'ascospores rejeté au stade le plus sensible de la culture (avant le stade 6 feuilles) est actuellement respecté (West *et al.*, 2001; Aubertot *et al.*, 2004).

La méthode de lutte la plus efficace et la plus utilisée face à *L. maculans* consiste en l'utilisation de cultivars de *B. napus* sélectionnés pour leurs gènes de résistance spécifiques ainsi que pour leur niveau de résistance générale.

#### a) Résistance quantitative

La résistance générale ou quantitative d'un cultivar ne permet pas à la plante d'empêcher totalement le développement de *L. maculans* mais conduit à une

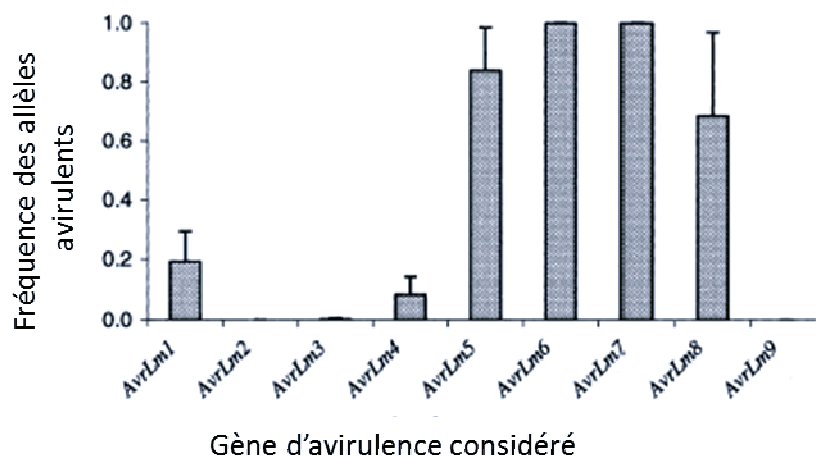
diminution ou disparition des symptômes au collet. Le comportement au champ du cultivar est alors considéré comme plus ou moins bon en fonction de l'importance des symptômes à la récolte. Cette résistance est relativement durable du fait du peu de pression de sélection imposée sur les populations de *L. maculans* et s'illustre par l'utilisation de la variété Jet Neuf en Europe de 1977 à 1983 sans que ce cultivar ne connaisse de baisse de son niveau de résistance (Pinochet *et al.*, 2003). La résistance quantitative présente donc un grand intérêt pour le contrôle à long terme de la maladie mais du fait d'un déterminisme génétique complexe faisant intervenir plusieurs QTL (Quantitative Trait Loci), les mécanismes d'action et de contrôle de cette résistance ne sont pas encore bien appréhendés et rendent complexe la sélection et l'amélioration de cultivars pour la résistance générale (Delourme *et al.*, 2008; Poland *et al.*, 2009).

#### b) Résistance spécifique

La résistance spécifique dans le couple *L. maculans* / *B. napus* suit le modèle gène-à-gène (Ansan-Melayah *et al.*, 1995; Ansan-Melayah *et al.*, 1998; Balesdent *et al.*, 2002) et implique la présence de gènes de résistance spécifiques *Rlm* (Résistance à L. maculans) chez le colza et de gènes d'avirulence correspondant *AvrLm* (Avirulence de L. maculans) chez le champignon. La résistance spécifique s'exprime au premier stade de l'infection lors de la colonisation des organes aériens de la plante et procure un haut niveau de protection vis-à-vis de *L. maculans* lorsque les allèles d'avirulence correspondants sont prédominants dans la population pathogène (Ansan-Melayah *et al.*, 1997). La facilité de sélection de cultivars de *B. napus* possédant des gènes de résistance spécifiques, en comparaison de la difficulté à sélectionner des cultivars ayant une bonne résistance quantitative, font de la résistance spécifique, la résistance la plus communément utilisée en sélection variétale.

#### c) Gènes d'avirulence

9 gènes d'avirulence ont été génétiquement déterminés à ce jour chez *L. maculans* : *AvrLm1* à *AvrLm9*. Des travaux de génétique ont montré que sept de ces gènes étaient organisés en clusters génétiques : *AvrLm1-AvrLm2-AvrLm6* d'une part, et *AvrLm3-AvrLm4-AvrLm7-AvrLm9* d'autre part. *AvrLm5* et *AvrLm8* sont quant à eux indépendants (Balesdent *et al.*, 2002). Deux de ces gènes ont été clonés, *AvrLm1* et



**Figure 14 : Fréquence des allèles avirulents des gènes *AvrLm1* à *AvrLm9* en France en 2001/2002**

1787 isolats naturels de *L. maculans* ont été prélevés sur 20 sites à travers la France. Leur analyse phénotypique a permis d'établir une estimation de la proportion de la fréquence des allèles avirulents de chacun des gènes d'avirulence génétiquement déterminés à ce jour. Adapté de Balesdent *et al.* (2006).

*AvrLm6*. Ces deux gènes sont situés au sein d'un isochore riche en AT et codent tous les deux pour des petites protéines potentiellement sécrétées typiques de la majorité des gènes d'avirulence fongiques (Gout *et al.*, 2006; Fudal *et al.*, 2007).

Une analyse de la fréquence des allèles des 9 gènes d'avirulence *AvrLm1* à *AvrLm9* dans les populations, a été menée en France en 2000/2001 et en Europe en 2002/2003 (Balesdent *et al.*, 2006; Stachowiak *et al.*, 2006). Ces deux analyses ont globalement fait apparaître les mêmes résultats. Les allèles avirulents *AvrLm6* et *AvrLm7* sont présents à plus de 99% dans les populations étudiées. A l'inverse, plus de 99% des souches testées sont virulentes pour *AvrLm2*, *AvrLm3* et *AvrLm9*, les allèles virulents *avrLm2* et *avrLm9* étant même fixés (100% des souches testées). Les allèles virulents pour les gènes *AvrLm1* et *AvrLm4* sont majoritairement présents (80.4% et 91.4%) tandis que les allèles avirulents des gènes *AvrLm5* et *AvrLm8* sont majoritairement présents (83.8% et 68.4%) (Figure 14).

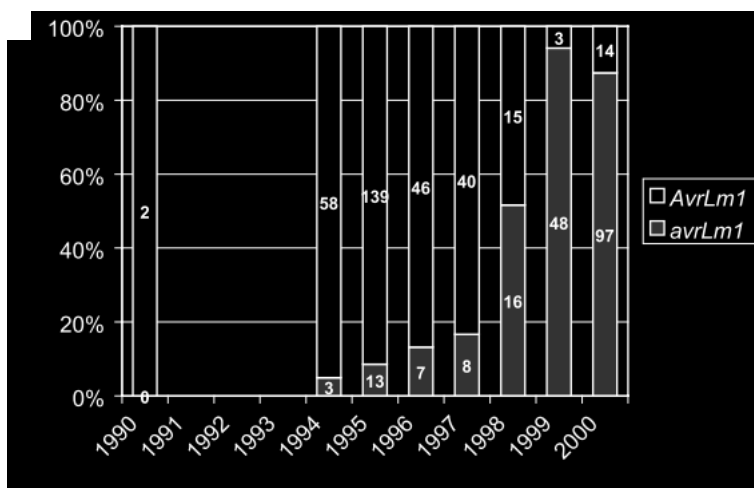
#### d) Gènes de résistance

A ces 9 gènes d'avirulence, correspondent 9 gènes de résistance conférant une résistance spécifique vis-à-vis de *L. maculans*. 6 de ces gènes ont été identifiés chez *B. napus* (*Rlm1*, *Rlm2*, *Rlm3*, *Rlm4*, *Rlm7* et *Rlm9*) et 3 au sein d'espèces voisines de *Brassicaceae* : deux chez *B. juncea* (*Rlm5* et *Rlm6*), et un chez *B. rapa* (*Rlm8*) (Balesdent *et al.*, 2002). Trois autres gènes de résistance spécifique ont également été caractérisés chez *B. rapa* : *LepR1*, *LepR2* et *LepR3* (Rimmer, 2006), sans que l'on puisse à ce jour estimer une éventuelle redondance entre ces trois gènes et les gènes *Rlm1* à *Rlm9*.

#### e) Contournement et durabilité des résistances à *L. maculans*

Depuis 1970, trois gènes de résistance spécifiques ont été successivement utilisés à des fins commerciales et largement employés dans les cultures françaises (*Rlm2*, *Rlm4* et *Rlm1*) (Rouxel *et al.*, 2003). La pression de sélection exercée par leur utilisation a favorisé l'émergence de souches virulentes pour chacun de ces gènes et a rendu leur utilisation rapidement inefficace. La très grande prédominance des allèles virulents des gènes d'avirulence correspondant à ces gènes (*AvrLm1*, *AvrLm2*, *AvrLm4*) dans les populations françaises actuelles illustre bien l'historique des résistances utilisées et contournées (Balesdent *et al.*, 2006).





**Figure 15 : Contournement de la résistance *Rlm1* en France.**

Fréquences des allèles virulents (*avrLm1*) et avirulent (*AvrLm1*) du gène *AvrLm1* entre 1990 et 2000 des isolats naturels de *L. maculans* collectés sur cultivars dépourvus du gène de résistance correspondant *Rlm1*. Le nombre d'isolats utilisés pour calculer les fréquences est indiqué dans chacune des colonnes. Adapté de Rouxel *et al.* (2003).

La proportion des allèles d'*avrLm2* et d'*avrLm9* avant l'utilisation au champ de leurs gènes de résistance correspondant n'est pas connue puisque aucune collecte massive de souches de *L. maculans* n'a été effectuée en France avant la fin des années 90. Néanmoins, on peut supposer que la très grande proportion de souches virulentes observée en France en 2000 pour ces deux gènes d'avirulence (Balesdent *et al.*, 2006) fait suite au contournement de leur gène de résistance respectif, comme observé lors du contournement de *Rlm1*. En effet, l'utilisation commerciale de cultivars possédant *Rlm1* a provoqué en seulement 3 années culturales (1998-2001), le passage d'une population très majoritairement avirulente à une population largement virulente (87% en 2000) (Rouxel *et al.*, 2003) (Figure 15). De la même manière, l'utilisation de *Rlm4* depuis les années 70 a provoqué l'apparition d'une population de souches virulentes *avrLm4* mais à la différence des gènes *AvrLm2* et *AvrLm9*, près de 9% des souches avirulentes *AvrLm4* sont encore présentes dans les populations françaises de *L. maculans* (Balesdent *et al.*, 2006). Ce maintien de souches avirulentes pourrait s'expliquer par le déficit de fitness que présentent les souches *avrLm4* (Huang *et al.*, 2006), comparé au moindre coût mesuré pour *AvrLm1* (Huang *et al.*, 2010).

Comme l'illustre l'écart de temps mis par les populations de *L. maculans* pour contourner *Rlm1* et *Rlm4*, la durabilité des résistances monogéniques employées est variable. Les résistances spécifiques utilisées à ce jour chez le colza se sont révélées globalement peu durables et le contournement rapide des résistances monogéniques utilisées au champ comme *Rlm1* constitue un problème majeur dans la lutte contre *L. maculans*. La combinaison de résistances spécifiques et résistances quantitatives semble être une solution possible pour augmenter la durabilité de ces résistances. Ainsi, la résistance contrôlée par le gène de résistance *Rlm6* combinée à la résistance globale du cultivar Darmor reste efficace au champ au minimum deux ans de plus que dans un cultivar possédant un fond génétique sensible (Brun *et al.*, 2010).

#### f) Couple *AvrLm7-Rlm7*

Malgré le contournement rapide des gènes de résistance spécifiques en situation de culture intensive, le développement et l'utilisation de cultivars de *B. napus* possédant de nouveaux gènes de résistance restent prioritaires du fait de leur efficacité immédiate, de leur sélection relativement aisée et du manque d'alternatives

à leur utilisation. Ainsi, *Rlm7* est un « nouveau » gène de résistance introduit commercialement pour la première fois en France en 2003 avec la variété Roxet (Pinochet, communication personnelle), mais surtout à partir de 2005 avec les variétés hybrides Exagone, Exocet, Extend, etc (Pinochet communication personnelle). Il correspond à *AvrLm7*, gène d'avirulence présent en 2000/2001 dans les populations française à plus de 99.9% sous sa forme allélique avirulente (Balesdent *et al.*, 2006). Comme nous le verrons dans le chapitre 1, *AvrLm7* est, après *AvrLm1* et *AvrLm6*, le troisième gène d'avirulence à avoir été cloné chez *L. maculans* et de la même façon que ces deux autres gènes, code pour une petite protéine sécrétée putative dont le gène est isolé au sein d'un isochoire AT. Cette protéine présente une double spécificité d'interaction vis-à-vis des gènes de résistance *Rlm4* et *Rlm7*. *AvrLm4* et *AvrLm7* ne sont en fait qu'un unique gène dénommé *AvrLm4-7* (Chapitre1).

## **D. Objectif de la thèse**

La protection des cultures de colza face à *L. maculans* repose principalement sur l'utilisation de gènes de résistance spécifiques ciblant des gènes d'avirulence présents dans la population d'agents pathogènes. Plusieurs de ces résistances ont déjà été massivement déployées en France et dans le monde, connaissant dans un premier temps un fort succès, tant pour la protection fournie, que commercial, suivi d'une perte d'efficacité très rapide due à l'apparition de nouvelles populations virulentes. Actuellement, un des principaux objectifs dans la lutte contre les agents pathogènes, et contre *L. maculans* en particulier, est de mettre en place une gestion des résistances spécifiques permettant d'augmenter la durabilité des résistances exploitées. Pour cela, la compréhension des processus évolutifs et moléculaires agissant sur les gènes d'avirulence et à l'origine du contournement des gènes de résistance est primordiale. Néanmoins, seules quelques études au champ portant sur l'évolution moléculaire des gènes d'avirulence en milieu naturel sous pression de sélection ont été effectuées (Gassmann *et al.*, 2000; Schurch *et al.*, 2004). Chez *L. maculans*, une étude portant sur les mécanismes moléculaires impliqués dans le gain de la virulence vis-à-vis de variétés possédant *Rlm1* a été effectuée mais correspond à une explication rétrospective du phénomène qui a conduit au contournement du gène de résistance, plusieurs années après que celui-ci a eu lieu

(Gout *et al.*, 2007). Cette étude ne permet pas l'accès à tous les événements ayant conduit au contournement ni à la succession d'épisodes qui fait qu'un événement a largement pris le dessus sur toutes les autres possibilités. Des éléments de réponses pourraient être apportés par l'étude de l'interaction *AvrLm6-Rlm6* pour laquelle plusieurs événements conduisant à la virulence ont été identifiés (Fudal *et al.*, 2009), mais pour laquelle la résistance correspondante n'est pas actuellement utilisée au champ en Europe.

Le nombre d'études au champ des processus évolutifs et moléculaires impliqués dans le contournement d'un gène de résistance est donc très limité, en particulier chez les champignons. Le clonage d'*AvrLm4-7* et l'utilisation récente de *Rlm7* en pratique agricole fournissent un contexte de choix pour étudier en temps réel l'évolution moléculaire d'*AvrLm4-7* sous pression de sélection et font de ce gène d'avirulence un candidat idéal pour une étude précoce et détaillée des mécanismes moléculaires à l'origine du contournement d'une résistance.

Suite au clonage du gène *AvrLm4-7* par F. Parlange et à la validation de la fonctionnalité du résidu R<sup>120</sup> dans la perte de reconnaissance par *Rlm4* à laquelle j'ai contribué durant mon stage de Master 2 (Chapitre 1), j'ai donc entrepris une analyse du contournement du gène de résistance *Rlm7* en situation de forte pression de sélection et ai caractérisé les événements moléculaires responsables de ce contournement. Un tel contournement étant accompagné d'une résurgence de la spécificité *AvrLm3*, absente des populations françaises au début de ce travail, j'ai enfin initié une analyse visant au clonage d'*AvrLm3* et à la caractérisation de ses relations moléculaires avec *AvrLm7*.

L'ensemble des résultats obtenus durant ma thèse est présenté dans ce manuscrit selon trois chapitres :

Dans la 1<sup>ère</sup> partie de ce manuscrit je présenterai ma participation au clonage et à la caractérisation fonctionnelle d'*AvrLm4-7* (Parlange, 2006). Ma contribution à cette étude a consisté à valider le rôle de l'acide aminé G<sup>120</sup> d'*AvrLm4-7* dans la spécificité d'interaction *AvrLm4-Rlm4* par mutagenèse dirigée, participant ainsi à la mise en évidence de la double spécificité du gène et de son déterminisme moléculaire. Ce travail a fait l'objet d'une publication dans « *Molecular Microbiology* » en 2009.

Je présenterai ensuite dans une 2<sup>nd</sup>e partie, l'analyse dynamique des populations de *L. maculans* sous pression de sélection (*AvrLm7-Rlm7*) sur les sites de Grignon et de Versailles. Il s'agit ici d'une part d'évaluer l'influence des systèmes de culture sur le contournement du gène de résistance *Rlm7* et d'autre part d'identifier les tous premiers événements conduisant à la virulence *avrLm7* en situation de pression de sélection au champ puis de suivre l'évolution dans le temps (3 cycles culturaux soit 3 cycles sexuels) de la virulence *avrLm7*. Il s'agit de la première étude de ce type chez un champignon. Cette étude a permis de montrer que les systèmes de culture influençaient grandement la vitesse de contournement de *Rlm7* et, par l'établissement et l'analyse d'une grande collection de souches virulentes de *L. maculans* vis-à-vis de *Rlm7*, de démontrer que cette virulence est due à de multiples événements de mutation ayant lieu chaque année probablement au sein même de la parcelle.

Finalement, je rapporterai le travail effectué sur le lien existant entre les deux gènes d'avirulence de *L. maculans* *AvrLm3* et *AvrLm4-7*. En effet, la perte de fonction d'*AvrLm7* pour une souche se traduit le plus souvent pour celle-ci par la résurgence du phénotype avirulent *AvrLm3*. Le lien *AvrLm3-AvrLm7* a été confirmé par la caractérisation phénotypique des souches naturelles répertoriées sur Grignon et l'interaction fonctionnelle entre ces deux gènes démontrée par des expériences de complémentation et d'extinction d'expression par ARNi. Je décrirais également dans cette partie, les premières étapes qui devraient mener au clonage d'*AvrLm3*.





# **Chapitre 1**

## **Double spécificité d'interaction du** **gène d'avirulence *AvrLm4-7***





## Introduction

Les deux premiers gènes d'avirulence de *L. maculans* à avoir été clonés, *AvrLm1* et *AvrLm6*, font partie du même cluster génétique et présentent de nombreuses caractéristiques communes. Ces deux gènes codent en effet pour des petites protéines potentiellement sécrétées et ne possèdent ni domaines conservés ni fonction connue. Leur expression *in planta* est également similaire, tout comme leur localisation au sein d'un isochoire AT riche en éléments transposables dégénérés (Gout *et al.*, 2007; Fudal *et al.*, 2009). Le clonage d'un nouveau gène d'avirulence génétiquement indépendant du cluster *AvrLm1-2-6* était alors envisagé afin de savoir si les spécificités retrouvées chez *AvrLm1* et *AvrLm6* ainsi que leur environnement génomique particulier étaient extrapolables, ou non, à d'autres gènes, ou cluster de gènes d'avirulence, de *L. maculans*. Ce clonage devait également permettre de mieux comprendre la fonction de ces gènes d'avirulence ainsi que la pertinence de la notion de cluster génétique. Le clonage d'*AvrLm7*, gène d'avirulence appartenant au cluster de gènes d'avirulence *AvrLm3-4-7-9* et correspondant à un gène de résistance qui n'était pas encore utilisé au champ (Balesdent *et al.*, 2002; Balesdent *et al.*, 2005) a alors été entrepris par Francis Parlange au cours de sa thèse (Parlange, 2006).

Le clonage d'*AvrLm7* ainsi que les premières analyses moléculaires effectuées sur celui-ci ont permis de montrer qu'il partageait avec les deux autres gènes d'avirulence clonés chez *L. maculans*, les mêmes caractéristiques structurales (taille, peptide signal, absence de domaines connus...), une localisation génomique semblable ainsi qu'une cinétique d'expression *in planta* analogue (Parlange, 2006). *AvrLm6* et *AvrLm7* possèdent un nombre important de cystéines (respectivement 6 et 8), à la différence d'*AvrLm1* qui n'en possède qu'une. Ces similitudes sont aujourd'hui considérées comme représentatives des gènes d'avirulence et sont devenues les critères de base servant à une recherche systématique de gènes d'avirulence et d'effecteurs dans le génome de *L. maculans*.

*AvrLm7* se distingue néanmoins d'*AvrLm1* et d'*AvrLm6* par sa double spécificité d'interaction vis-à-vis de *Rlm4* et de *Rlm7*. En effet, tandis que *AvrLm1* et *AvrLm6* sont deux gènes distincts, *AvrLm4* et *AvrLm7* s'avèrent n'être qu'un unique gène désormais rebaptisé *AvrLm4-7*. Une première analyse comparative de souches avirulentes et virulentes vis-à-vis de *Rlm4* et de *Rlm7* suggérait une possible distinction des souches virulentes et avirulentes vis-à-vis de *Rlm4* selon la présence d'une guanine ou d'une cytosine en position 358 du gène. La validation du rôle de ce nucléotide pour la spécificité d'interaction *AvrLm4-Rlm4* m'a été confiée et m'a permis de confirmer par mutagenèse dirigée le rôle primordial de ce nucléotide dont la mutation affecte la reconnaissance d'*AvrLm4* par *Rlm4* sans toutefois altérer la reconnaissance d'*AvrLm7* par *Rlm7*.

*AvrLm7* est désormais du fait de ses caractéristiques propres ou partagées, ainsi que du fait de l'utilisation récente de *Rlm7* au champ, le modèle d'étude privilégié du laboratoire en termes d'évolution moléculaire (Chapitre 2) et d'analyse fonctionnelle (Chapitre 1 et Chapitre 3).

Le travail de F. Parlange sur le clonage et la caractérisation moléculaire d'*AvrLm4-7* ainsi que ma contribution à l'étude du rôle du nucléotide 358 d'*AvrLm4-7* dans la spécificité d'interaction *AvrLm4-Rlm4* sont ici présentés sous la forme de l'article publié en 2009 dans le journal « *Molecular Microbiology* » et intitulé « *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change ».

# *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change

Francis Parlange,<sup>1,2†</sup> Guillaume Daverdin,<sup>2</sup>  
Isabelle Fudal,<sup>2</sup> Marie-Line Kuhn,<sup>2</sup>  
Marie-Hélène Balesdent,<sup>2</sup> Françoise Blaise,<sup>2</sup>  
Bruno Grezes-Bessel<sup>1</sup> and Thierry Rouxel<sup>2\*</sup>  
<sup>1</sup>BIOGEMMA, Laboratoire de Biologie Cellulaire et  
Moléculaire, Domaine de Sandrean, F-31700  
Mondonville, France.  
<sup>2</sup>INRA, UMR1290-Biogor, Route de Saint Cyr, F-78026  
Versailles Cedex, France.

## Summary

*Leptosphaeria maculans* is the ascomycete responsible for one of the most damaging diseases of oilseed rape (*Brassica napus*), stem canker of crucifers. Both avirulence (*AvrLm*) genes in the fungus and resistance (*Rlm*) genes in the plant are genetically clustered. Using a map-based cloning strategy, we delineated a 238 kb region containing the *AvrLm7* locus. Structural features of the region were reminiscent of those previously found on another chromosome for genomic regions encompassing *AvrLm1* and *AvrLm6*, i.e. GC-equilibrated, gene-rich isochores alternating with AT-rich, recombination-deficient, gene-poor isochores. These latter corresponded to mosaics of degenerated and truncated transposable elements. *AvrLm7* is the only gene located within a 60 kb AT-rich isochore. It induced resistance responses in plants harbouring either *Rlm7* or *Rlm4*, and was thus renamed *AvrLm4-7*. It encodes a 143-amino-acid cysteine-rich protein, predicted to be secreted, and strongly induced during early stages of plant infection. Sequencing and

restriction analyses of *AvrLm4-AvrLm7* or *avrLm4-AvrLm7* alleles in *L. maculans* field isolates, and targeted point mutagenesis strongly suggested that one single base mutation, leading to the change of a glycine to an arginine residue, was responsible for the loss of *AvrLm4* specificity whereas *AvrLm7* recognition was unaltered.

## Introduction

During interactions between plants and microbial pathogens, plant active disease resistance often relies on specific recognition of invading pathogens. This specificity complies with the gene-for-gene model (Flor, 1971), which predicts that a plant resistance (*R*) gene product specifically detects the product of a pathogen avirulence (*Avr*) gene. Perception of the *Avr* protein, or of the effect of the *Avr* protein on a host protein target, by the plant elicits a large array of defence responses that usually results in a localized host cell death called hypersensitive response (Greenberg and Yao, 2004). To date, *Avr* genes have been cloned for only seven fungal species: the ascomycetes *Cladosporium fulvum* (Rivas and Thomas, 2005), *Rhynchosporium secalis* (Rohe *et al.*, 1995), *Magnaporthe grisea* (Orbach *et al.*, 2000; Farman *et al.*, 2002; Böhnert *et al.*, 2004), *Fusarium oxysporum* (Rep *et al.*, 2004; Houterman *et al.*, 2008), *Leptosphaeria maculans* (Gout *et al.*, 2006; Fudal *et al.*, 2007) and *Blumeria graminis* (Ridout *et al.*, 2006), and the basidiomycete *Melampsora lini* (Dodds *et al.*, 2004; Catanzariti *et al.*, 2006). Ascomycete *Avr* genes often, but not exclusively, encode small cysteine-rich proteins, without homology with each other or with proteins in the database, and are often predicted to be secreted by fungal cells during plant infection (Rep, 2005). However, their intrinsic function is difficult to predict owing to the lack of known conserved domains.

*Leptosphaeria maculans* is an ascomycete fungus causing stem canker of crucifers (Rouxel and Balesdent, 2005). Its interaction with oilseed rape (*Brassica napus*) usually corresponds to a typical gene-for-gene relationship

Accepted 10 November, 2008. \*For correspondence. E-mail rouxel@versailles.inra.fr; Tel. (+33) 1 30 83 32 29; Fax (+33) 1 30 83 31 95.  
†Present address: University of Zurich, Institute of Plant Biology, Zollikerstrasse 107, CH-8008 Zurich, Switzerland. Data deposition: The reported sequences have been submitted in the DDBJ/EMBL/GenBank databases under accession numbers AM998638 (*AvrLm4-7* gene), AM998635 (BAC 1B4), AM998636 (BAC 2B2) and AM998637 (BAC 28G8).

and, to date, nine avirulence (*AvrLm*) genes specifying the recognition of *L. maculans* by Brassica species have been genetically identified. Seven of them are clustered in two unlinked, genetic clusters (*AvrLm1-2-6* and *AvrLm3-4-7-9*), which is unusual for an ascomycete fungus (Balesdent *et al.*, 2002). A map-based cloning strategy and the sequencing of a 1.1 Mb genomic region encompassing the *AvrLm1-2-6* cluster permitted the cloning of *AvrLm1* (Gout *et al.*, 2006) and *AvrLm6* (Fudal *et al.*, 2007). *AvrLm1* and *AvrLm6* proteins share common features with most of the currently known fungal avirulence proteins: they are small proteins possessing an N-terminal signal peptide, without any homology with sequences available in databases, no conserved domain and no similarity between one another (Gout *et al.*, 2006; Fudal *et al.*, 2007). Expression of both genes is strongly induced during primary leaf infection, with a low level of expression *in vitro* (Gout *et al.*, 2006; Fudal *et al.*, 2007). One of the main features of *AvrLm1* and *AvrLm6* regards their occurrence in gene-poor isochores (Gout *et al.*, 2006; Fudal *et al.*, 2007). Isochores have been described in plant and mammalian genomes (Bernardi, 2000) but, to our knowledge, they have never been described in ascomycete fungi. They consist of long DNA stretches homogenous in nucleotide composition (GC content), with abrupt changes from one segment to the other, and a correlation between GC content and gene density, recombination frequency, and presence of transposable elements (Bernardi, 2000; Oliver *et al.*, 2001). In the *AvrLm1-2-6* region, two types of isochores were observed: first, regions showing a low GC content (30–40%), consisting of mosaics of intermingled copies of diverse complete or truncated degenerate retrotransposons, mostly devoid of coding sequences except *AvrLm1* and *AvrLm6*, and resulting in heterochromatin-like structures; and second, GC-equilibrated (~52%), gene-rich regions, devoid of transposable elements (Gout *et al.*, 2006; Fudal *et al.*, 2007). In contrast to common occurrence of coding sequences in GC-equilibrated isochores, *AvrLm1* and *AvrLm6* occurred as 'lost in the middle of nowhere' solo genes in large stretches of AT-rich non-coding sequences, respectively, 269 and 133 kb in length (Gout *et al.*, 2006; Fudal *et al.*, 2007). The genomic context of both *AvrLm1* and *AvrLm6* genes questioned whether this unusual genomic environment was also the rule for other *L. maculans* avirulence genes belonging to other linkage groups of the genetic map and the influence of such an environment on gene expression and/or evolution.

Thus, the objectives of this study were, first, to investigate the genomic environment surrounding the second avirulence gene cluster known in *L. maculans*, the *AvrLm3-4-7-9* cluster, and second, to develop a map-based cloning strategy to identify *AvrLm7*. Following construction of genetic and physical maps, a ~380 kb BAC (Bacterial Artificial Chromosome) contig was assembled,

and a 238 kb interval spanning the *AvrLm7* locus was delineated and entirely sequenced. As with *AvrLm1* and *AvrLm6*, *AvrLm7* candidate was shown to be isolated in a 60 kb gene-poor AT-rich region. Complementation assays with the *AvrLm7* candidate reveals that the gene has the unusual characteristic of specifying the *AvrLm7* and *AvrLm4* phenotypes, and it was thus renamed *AvrLm4-7*. Moreover, following extensive sequence analyses in natural isolates and targeted point mutation experiments, one single nucleotide polymorphism resulting in a Gly to Arg change in the protein was shown here to result in loss of *AvrLm4* specificity without affecting *AvrLm7* specificity.

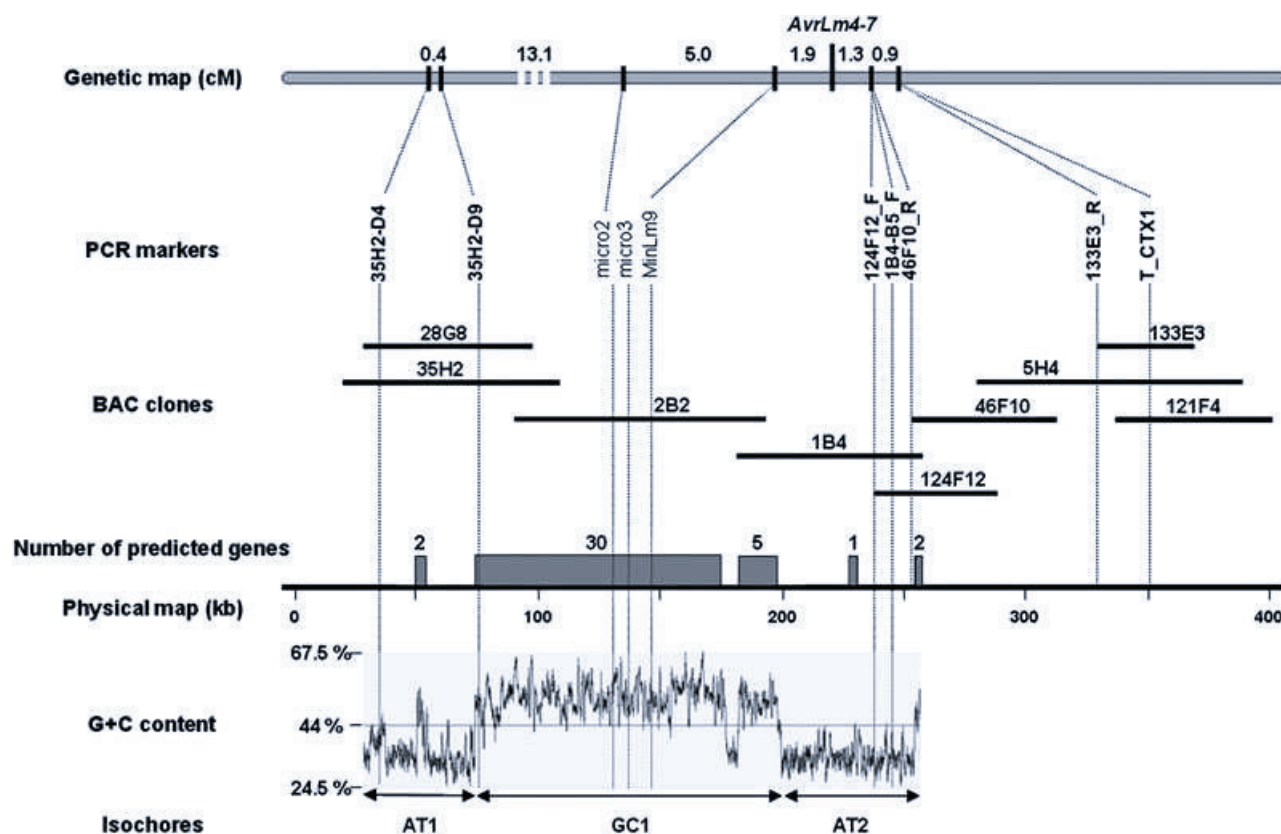
## Results

### Genetic and physical mapping of the *AvrLm7* locus

Analysis of 230 progeny from cross #33, v23.1.11 (*AvrLm4-AvrLm7*) × Nzt-4 (*avrLm4-avrLm7*), revealed a complete lack of recombination between *AvrLm4* and *AvrLm7* (Balesdent *et al.*, 2002). This result permitted us to take advantage of the genetic map derived from cross #23, v11.3.6 (*avrLm4-AvrLm7*) × 21.3.1 (*AvrLm4-AvrLm7*) (Kuhn *et al.*, 2006). The closest AFLP marker, T\_CTX1, located 5.0 cM from *AvrLm4* in cross #23 (Kuhn *et al.*, 2006), was cloned, converted into a SCAR marker, genetically mapped 2.2 cM from *AvrLm4* and *AvrLm7* in cross #33, and used to initiate a bidirectional chromosome walk to *AvrLm7* (Fig. 1). Following sequencing of BAC- and BAC-subclone-ends, four polymorphic SNP-derived CAPS markers and two polymorphic polymerase chain reaction (PCR) markers were identified, allowing us to orient the chromosome walk and eventually to border the *AvrLm7* locus on both sides (Fig. 1). A strong heterogeneity between physical and genetic distances was observed, with ratios ranging from 7.6 kb cM<sup>-1</sup> to 90–150 kb cM<sup>-1</sup>, indicating an uneven distribution of recombination events in the *AvrLm4-AvrLm7* region. Markers 124F12\_F and 35H2-D9 delineated a genetic interval containing *AvrLm4* and *AvrLm7*, and three overlapping BAC clones, 1B4, 2B2 and 28G8, covering a minimum tilling path between the two markers, were sequenced (Fig. 1).

### Structure of the *AvrLm4-AvrLm7* region

The final assembled sequence covered 238 kb and showed an isochore structure with two regions showing a low GC content (30–41% GC), AT1 (47.5 kb) and AT2 (60 kb), and a 130.5 kb central GC-equilibrated region (52% GC), GC1 (Fig. 1). Multiple transposable elements were identified in the contig, mostly within AT1 and AT2. They mainly corresponded to four previously characterized long-terminal repeat (LTR) retrotransposons *Pholy*,



**Fig. 1.** Genetic and physical maps of the *AvrLm4-7* genomic region in v23.1.3 (a3A4A7a9) isolate. The genetic map corresponds to cross #33 [v23.1.11 (*AvrLm4-AvrLm7*) × Nzt-4 (*avrLm4-avrLm7*)]. Dashed lines connect the physical location of markers to their location on the genetic map (micro3 is not polymorphic in cross #33). Except T\_CTX1, markers in bold were developed from BAC(sub)clone-end sequences. Number and location of predicted genes on the physical map are indicated. GC content was analysed (sliding window of 400 bp) in the 238 kb sequenced region composed of BACs 28G8, 2B2 and 1B4. Isochores are defined on the basis of their homogeneous GC content.

*Olly*, *Polly* and *Rolly* (Gout *et al.*, 2006), along with a novel *Gypsy* retroelement termed *Dolly* and a DNA transposon, termed *IR49*. The repeats, which accounted for 85% of the AT-rich isochores, were organized as mosaics of complete or truncated elements. The elements were highly degenerate following Repeat Induced Point mutation as has been previously reported for *Pholy* (Attard *et al.*, 2005). The complete copy of a second novel degenerated LTR retrotransposon, termed *Zolly*, assigned to the *Copia* family was also identified. Contrasting with all currently known retrotransposons of *L. maculans*, *Zolly* was inserted in the GC-rich isochore GC1, between two coding sequences, as shown by the abrupt decrease in GC content within GC1 (Fig. 1).

Forty putative genes were predicted within the contig (Fig. 1). Based on homology to gene or protein sequences deposited in databases, three showed homologies with hypothetical fungal proteins of unknown function, 28 could be assigned putative functions and corresponded mostly to genes involved in essential cellular processes, whereas nine showed no significant homology (Supplementary Table S1). Thirty-five of the 40

putative genes were in the GC1 isochore, and the remaining five predicted genes were isolated in AT-rich isochores: two in AT1, and three in AT2 (Fig. 1).

Hybridization of isolate v23.1.3 electrokaryotypes with a GC-equilibrated region from BAC 46F10, along with previous localization of the nitrate reductase gene in GC1 (Table S1) (Kuhn *et al.*, 2006), showed that *AvrLm7* was carried on a chromosome whose size ranged between 3.0 and 3.2 Mb, thus distinct from the 2.8 Mb chromosome carrying the *AvrLm1-2-6* cluster (Kuhn *et al.*, 2006).

#### Identification of *AvrLm7*

Based on the contig sequence, three novel molecular markers were developed from Simple Sequence Repeats, and one of these, *MinLm9* restricted the interval comprising *AvrLm7* to 1.9 cM, corresponding to an 85 kb region (Fig. 1). Eleven genes were predicted within this interval: 10 in GC1, and an isolated gene embedded in AT2 (Fig. 1). Occurrence of a solo gene in a large AT-rich non-coding region was reminiscent of the gene organization for *AvrLm1* and *AvrLm6*, and this gene was thus the

**Table 1.** Characterization of *L. maculans* transformants (T). Molecular identification of presence of the construct and pathogenicity behaviour on oilseed rape cultivars carrying diverse resistance genes.

	PCR <sup>a</sup>	Pathogenicity test on <sup>b</sup>					
		Rlm1 +					
		Westar	Rlm7	Rlm7	Rlm4	Rlm3	Rlm9
v23.1.3 (a3-A4-A7-a9) <sup>c</sup>	+	V	A	A	A	V	V
v23.1.2 (a3-a4-A7-a9)	+	V	A	A	V	V	V
Nzt-4 (a3-a4-a7-a9)	-	V	V	V	V	V	V
30 Nzt-4 transformants with the v23.1.3 allele A4A7 <sup>d</sup>							
T1-T5 and T7-T30	+	V	A	A	A	V	V
T6	-	V	V	V	V	V	V
12 Nzt-4 transformants with the v23.1.2 allele a4A7 <sup>d</sup>							
T1-T11	+	V	A	nd	V	V	V
T12	-	V	V	nd	V	V	V
19 Nzt-4 transformants with the mutagenized allele a4*A7 <sup>e</sup>							
T1-T9 and T11-T19	+	V	A	nd	A	V	V
T10	-	V	V	nd	V	V	V
13 v23.1.2 transformants with the mutagenized allele a4*A7 <sup>e</sup>							
T1-T13	+	V	A	nd	A	V	V

a. Presence of the AvrLm4-7 gene as indicated by PCR amplification.

b. Isolates were inoculated onto reference cultivars carrying diverse resistance genes (except Westar devoid of known resistance gene). Rlm7: 01.23.2.1, Exagone, Hearty, Roxet, 00.4.1.1; Rlm1 + Rlm7: 99.149.1.1; Rlm4: Jet Neuf, Eleonore, Madrigal, Pollen, Pixel; Rlm3: 00.22.1.1; Rlm9: Darmor. (A) avirulent phenotype; (V) virulent phenotype; nd, not done.

c. A indicates that the isolate is avirulent at a given locus, a indicates it is virulent at a given locus, the a4\*A7 allele corresponds to the virulent avrLm4-AvrLm7 allele with a single targeted replacement of C<sup>358</sup> by G<sup>358</sup>.

d. Transformants not tested on Pixel.

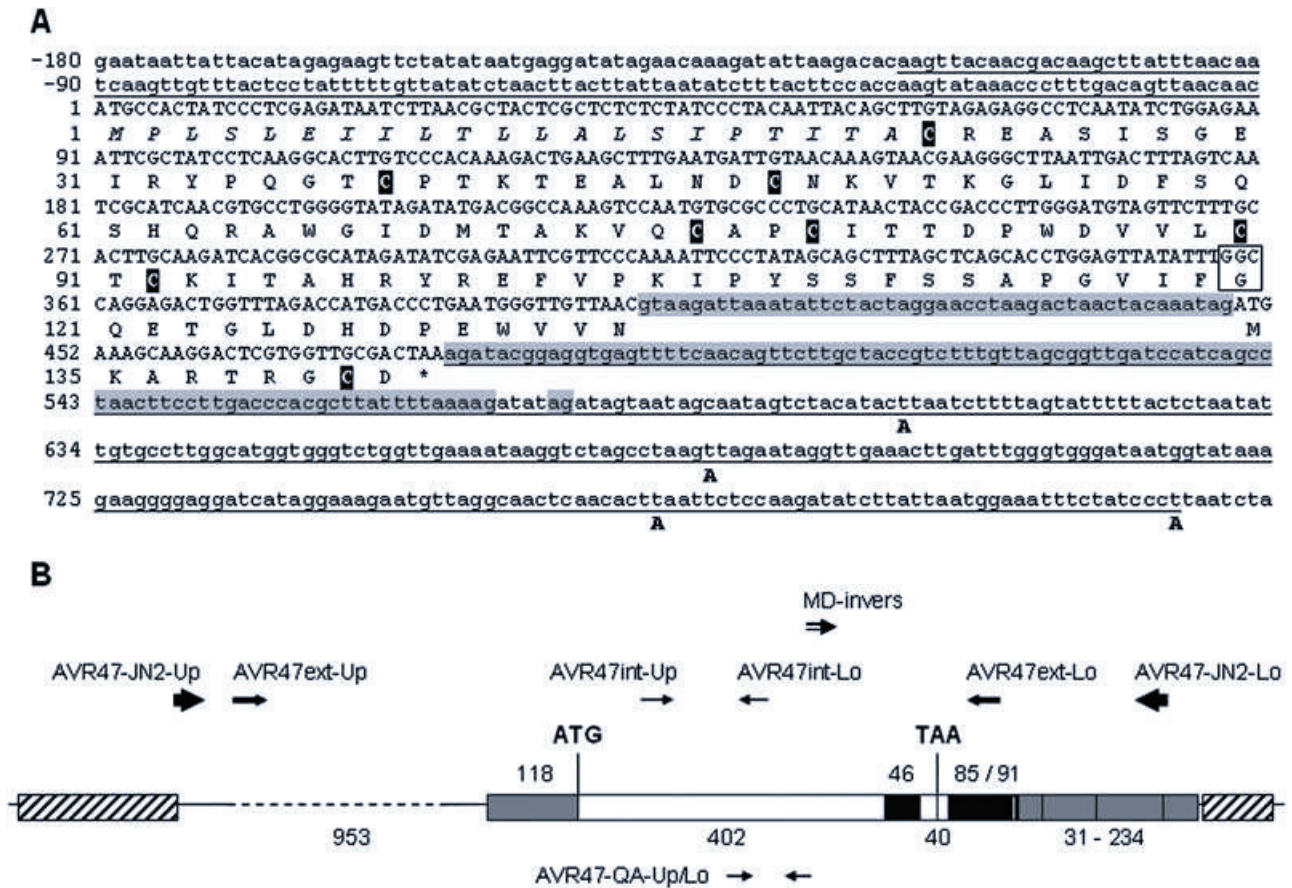
e. Transformants not tested on Eleonore, Madrigal and Pollen.

first candidate investigated. After reverse transcription polymerase chain reaction (RT-PCR) confirmation that the candidate gene was expressed in vitro and in infected leaves (data not shown), a 2353 bp genomic fragment corresponding to the coding sequence and putative promoter and terminator regions of the gene was isolated by enzymatic digestion of clone 1B4, cloned and transformed into the double virulent isolate Nzt-4 (termed a4a7) by *Agrobacterium tumefaciens* mediated transformation. In this latter isolate, the candidate gene was absent as observed following PCR amplification and Southern blotting experiments (data not shown). Thirty independent transformants were obtained, and PCR amplification indicated presence of the candidate gene for all but one of them, T6 (Table 1). All transformants maintained their virulence towards plants harbouring Rlm3, Rlm9 or no Rlm gene (Table 1). All transformants except T6 showed a restored avirulent phenotype towards Rlm7 genotypes, but also unexpectedly towards a range of cultivars harbouring Rlm4. To further investigate the putative double specificity expressed by AvrLm7, the same genomic region was cloned from isolate v23.1.2 (virulent allele for AvrLm4, and avirulent for AvrLm7: a4A7) and transformed into Nzt-4. Twelve transformants were obtained, of which one, T12, did not amplify AvrLm7. All transformants retained their virulence towards plants harbouring Rlm3, Rlm9, Rlm4 or no Rlm gene. All of them but T12 restored

the avirulent phenotype towards Rlm7 (Table 1). These data indicated that the same sequence, containing only one predicted open reading frame (ORF), can specify either the AvrLm7 specificity alone or both AvrLm4 and AvrLm7 specificities. The solo gene was thus renamed AvrLm4-7.

Gene annotation, sequence analysis and expression of AvrLm4-7

The restriction fragment used for functional complementation contained a single predicted ORF, located in an 1887 bp interval bordered by two truncated retrotransposons (Fig. 2). Precise annotation of the coding sequence was performed by 5' and 3' RACE-PCRs on cDNA synthesized from v23.1.3. Four different gene organizations were obtained, that shared a common 5' Untranslated Transcribed Region (UTR) of 118 nt, a first exon of 402 nt, a first intron of 46 nt, and a 40 nt second exon containing a stop codon (Fig. 2). The four transcripts obtained differed in their 3' region, after the stop codon by harbouring either a second 85 or 91 nt intron, and different polyadenylation sites (Fig. 2). In all cases, the AvrLm4-7 coding sequence was 432 nt long, with a transcript size between 591 and 794 nt. AvrLm4-7 encodes a predicted 143-amino-acid (aa) preprotein including a 21 aa cleavable signal peptide (P = 0.990). Eight cysteine



**Fig. 2.** Organization of the *AvrLm4-7* gene and its surrounding region.

A. Nucleotide sequence of the 995 nucleotide (nt) region encompassing *AvrLm4-7* and amino acid sequence of the *AvrLm4-7* predicted protein. 5' and 3' UTRs are underlined. Introns are indicated by gray shading (first intron: 46 nt; second intron: 85 or 91 bp, second acceptor site is indicated in gray). The four possible polyadenylation sites are represented by a 'A' in bold below the nucleotide sequence. The predicted 21 aa-signal peptide is in italics, and the eight cysteine residues are indicated by black shading. The mutated codon ( $G^{358} \rightarrow C^{358}$ ) leading to a  $G^{120} \rightarrow R^{120}$  change and affecting *Rlm4* recognition is indicated with a box.

B. Schematic representation of the *AvrLm4-7* gene. 5' and 3' UTRs are indicated by gray boxes. Bars in the 3' UTR represent the different polyadenylation sites. Exons are indicated by open boxes. Introns are indicated by black boxes, the white bar inside the second intron represents acceptor site of 85 nt intron. Hatched boxes represent truncated *Pholy* and *Rolly* retrotransposons. Locations of primers are indicated.

residues were present, suggesting the possibility to build disulphide bonds (Fig. 2).

No significant homology was detected with sequences deposited in public databases nor with other fungal avirulence genes including *AvrLm1* and *AvrLm6*. In spite of microsynteny observed between isochore GC1 and genomic regions of the closely related species *Stagonospora nodorum* (Table S1), no significant matches were obtained between *AvrLm4-7* and corresponding sequences of *S. nodorum*. A very low similarity (25% identity, 47% similarity) was observed between *AvrLm4-7* and the Pig-14 protein of *Pyrenophora graminea* (CAB90218), a 126 aa protein expressed during infection of barley. Interestingly, the two proteins shared related cysteine residue spacing:  $CX_{13/15}CX_{10/9}CX_{27}CX_2CX_{6/10}CXCX_{40/49}C$ .

*AvrLm4-7* expression was investigated during infection and *in vitro* growth by quantitative RT-PCR (qRT-PCR)

(Table 2 and Fig. 3). Whereas *AvrLm4-7* expression was very low in germinating conidia, it was detected at a higher level in plants 3 days post infection (dpi), and then increased during primary leaf infection to reach a peak at 7 dpi. At this time point, *AvrLm4-7* was expressed twice more than  $\beta$ -tubulin (Fig. 3). Expression slowly decreased after 7 dpi but always remained at a higher level than during *in vitro* mycelial growth, even during late stem necrosis. (Table 2 and Fig. 3). *AvrLm4-7* expression kinetics during infection was thus very similar to that of *AvrLm1* and *AvrLm6* (Fudal *et al.*, 2007) (Table 2).

#### Polymorphism of *AvrLm4-7* in field isolates

Three hundred field isolates were investigated for occurrence of *AvrLm4* and *AvrLm7* alleles (Table S2 and data not shown): 44 A4A7, 223 a4A7 and 33 a4a7 isolates. No



**Table 2.** Characteristics of *L. maculans* avirulence genes.<sup>a</sup>

	<i>AvrLm1</i>	<i>AvrLm6</i>	<i>AvrLm4-7</i>
Transcript size (bp)	1053	817	591-794
GC content (%)	39.0	43.0	39.4-40.9
Protein size (aa)	205	144	143
Signal peptide (aa)	22	20	21
Cysteine residues	1	6	8
Expression level <sup>b</sup> in			
Germinating conidia	0.03%	0.02%	0.02%
Mycelial growth	0.49%	0.02%	0.88%
Infected cotyledons 3dpi	0.11%	0.04%	0.34%
Infected cotyledons 5dpi	1.20%	1.67%	37.11%
Infected cotyledons 7 dpi	58.26%	222.23%	274.07%
Infected cotyledons 12dpi	5.09%	27.43%	29.51%
Stem necrosis 60 dpi	0.92%	1.78%	3.56%

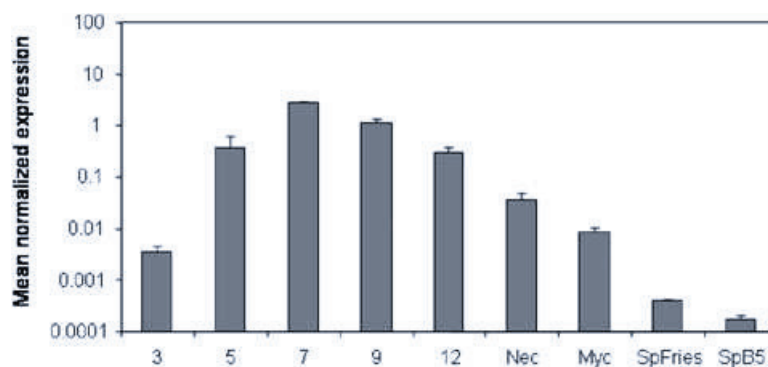
a. Data for *AvrLm1* and *AvrLm6* from Fudal *et al.* (2007).

b. Relative expression as a percentage of  $\beta$ -tubulin expression.

isolate avirulent towards *Rlm4* only (*A4a7*) was present in our collection. Following PCR amplifications with external or internal primers of *AvrLm4-7* (Table S3), 25 of 33 *a4a7* isolates showed a lack of amplification with one or both combination of primers, suggesting that complete or partial deletion of the gene is mainly responsible for the double virulent phenotype. This was confirmed by lack of Southern hybridization for nine out of 10 of these isolates (data not shown). In contrast, all *A4A7* and *a4A7* isolates showed a PCR product of the expected size. Sequencing of *AvrLm4-7* was performed in 11 *A4A7* and 33 *a4A7* field isolates (Table S2). Overall, the *AvrLm4-7* coding sequence showed a very low level of polymorphism, with only five polymorphic sites, of which one was only found in one *AvrLm4* isolate, and one in two *avrLm4* isolates (bases 123 and 134 corresponding, respectively, to aa

positions 41 and 45 in Table 3). These polymorphisms resulted in six haplotypes of which two were found in one isolate each. *A4A7* isolates showed a very conserved gene sequence with one haplotype represented in 10 isolates and identical to that of the reference v23.1.3 isolate, and the 11th being characterized by four additional point mutations of which only one was in the coding sequence ( $G^{123} \rightarrow C^{123}$  mutation leading to aa change  $K^{41} \rightarrow N^{41}$  in Table 3). *a4A7* alleles sequences defined four haplotypes: the most common was identical to the reference v23.1.2 isolate, and was characterized by two non-synonymous mutations on bases 239 ( $T^{239} \rightarrow C^{239}$  corresponding to aa change  $I^{80} \rightarrow T^{80}$  in Table 3) and 358 ( $G^{358} \rightarrow C^{358}$  corresponding to aa change  $G^{120} \rightarrow R^{120}$  in Table 3) as compared with the v23.1.3 sequence. The other three haplotypes were mostly combinations of non-synonymous mutations at bases 239, 256 and 358 (Table 3). The two reference v23.1.2 and v23.1.3 isolates used as donors in our transformation experiments differed only on two locations, corresponding to the  $I^{80} \rightarrow T^{80}$  and  $G^{120} \rightarrow R^{120}$  aa changes. The  $I^{80} \rightarrow T^{80}$  change was present in 60% of *avrLm4* isolates and thus cannot be correlated with the *avrLm4* phenotype, whereas the  $G^{120} \rightarrow R^{120}$  mutation was the only common trait shared by all the *avrLm4* sequences (Table 3). This mutation corresponds to replacement of a small (57.05 Da) non-polar aa by a threefold larger (156.19 Da) and positive aa, which was then postulated to be responsible for a large modification of the protein structure and for acquisition of the *avrLm4* phenotype.

Correspondence between the presence of  $R^{120}$  and an *avrLm4* phenotype was further examined by restriction analysis in 223 field isolates (190 *a4A7* and 33 *A4A7*).



**Fig. 3.** *AvrLm4-7* expression during *L. maculans* oilseed rape infection and *in vitro* growth. Expression was analysed in the avirulent isolate v23.1.3 by quantitative RT-PCR (qRT-PCR). Primer efficiency was calculated for *AvrLm4-7*,  $\beta$ -tubulin and *actin*, and found to range between 0.992 and 1. Similar expression levels of *actin* and  $\beta$ -tubulin were observed among *in planta* and *in vitro* conditions, suggesting that both genes are adequate constitutive reference genes (data not shown), and gene expression level of *AvrLm4-7* was measured relatively to  $\beta$ -tubulin. Lanes 3-12: RT-PCR product obtained from RNA isolated from oilseed rape cotyledons (susceptible cultivar Westar) 3-12 days post infection; Nec: RT-PCR product obtained from RNA isolated from stem necrosis (2 months after infection); Myc: RT-PCR product obtained from RNA isolated from mycelium grown *in vitro*; Sp: RT-PCR product obtained from RNA isolated from conidia germinating in rich medium (Fries) or under nitrogen starvation (B5). RNA extracted from uninfected cotyledons and water were used as negative controls. Each data point is the average of two biological repeats (two extractions from different biological material) and two technical repeats (two RT-PCR). Standard error of mean normalized expression level is indicated by error bars.

**Table 3.** Polymorphic sites in the AvrLm4-7 protein deduced from sequencing of AvrLm4-7 gene from 11 AvrLm4-AvrLm7 and 33 avrLm4-AvrLm7 natural isolates.

Phenotype	Amino acid position in the protein <sup>a</sup>					Number of isolates	Reference isolate
	41	45	80	86	120		
AvrLm4-AvrLm7	K	L	I	D	G	10	v23.1.3
	N	L	I	D	G	1	
avrLm4-AvrLm7	K	L	T	D	R	20	v23.1.2
	K	L	I	N	R	10	
	K	L	I	Y	R	1	
	K	S	I	D	R	2	

a. K<sup>41</sup> > N<sup>41</sup> aa change owing to G > C mutation at base 123; L<sup>45</sup> > S<sup>45</sup>; T > C mutation at base 134; I<sup>80</sup> > T<sup>80</sup>; T > C mutation at base 239; D<sup>86</sup> > N<sup>86</sup>; G > A mutation at base 256; D<sup>86</sup> > Y<sup>86</sup>; G > T mutation at base 256; G<sup>120</sup> > R<sup>120</sup>; G > C mutation at base 358. Direction of mutation given with v23.1.3 haplotype as reference.

Base 358 of the gene is part of an MscI restriction site (TGGCCA) postulated to be conserved only in AvrLm4 alleles. According to restriction analysis, all A4A7 isolates possessed a G at position 358, whereas 189 of 190 avrLm4 isolates did not show digestion of the PCR product, indicating a mutation of the restriction site (data not shown). Sequencing of AvrLm4-7 in the sole avrLm4 isolate whose PCR product was digested by MscI revealed ambiguous bases at positions 256 and 358, suggesting it corresponded to a mixture of isolates, which was not detected in the pathogenicity test.

#### Functional validation of G/R<sup>120</sup> involvement in AvrLm4 specificity

Site-directed mutagenesis of the cloned a4A7 allele of v23.1.2 allowed us to generate the a4\*A7 allele (C<sup>358</sup> → G<sup>358</sup> mutation corresponding to a R<sup>120</sup> → G<sup>120</sup> change in the predicted protein). The following transformants were obtained: 13 v23.1.2 with the a4\*A7 allele, and 19 Nzt-4 with the a4\*A7 allele. All transformants maintained their virulence towards plants harbouring Rlm3, Rlm9 or no Rlm gene (Table 1). Eighteen of the 19 Nzt-4\_a4\*A7 gained avirulence towards the Rlm7 line, suggesting site-directed mutagenesis maintained functionality of AvrLm7 avirulence (Table 1). Similarly, and except for one untransformed Nzt-4\_a4\*A7 isolate, all Nzt-4\_a4\*A7 and v23.1.2\_a4\*A7 transformants showed a restored AvrLm4 phenotype towards Rlm4 genotypes (Table 1). Change of the single C<sup>358</sup> → G<sup>358</sup>, resulting in a R<sup>120</sup> → G<sup>120</sup> change in the mature protein, thus was enough to recover the lost Rlm4-mediated recognition specificity.

## Discussion

In this paper, we report on the map-based cloning of AvrLm4-7, the third avirulence gene cloned from *L. maculans*. Whereas the gene showed structural and expression features similar to those of other ascomycete

avirulence genes, AvrLm4-7 shares with the previously characterized AvrLm1 and AvrLm6 a very peculiar genome location within large stretches of non-coding regions. In addition, AvrLm4-7 is unusual in that it is responsible for a double interaction specificity and that mutation of a single nucleotide is enough to escape Rlm4-mediated 'recognition' of AvrLm4-7.

#### The third 'lost in the middle of nowhere' avirulence gene of *L. maculans*

Location of most of the currently known *L. maculans* Avr genes in two unlinked genetic clusters (Balesdent *et al.*, 2002), shown here to actually correspond also to distinct chromosomes, represented an unusual feature in fungi. Together with previously demonstrated clustering of the corresponding R genes in the *B. napus* genetic map (Rlm1, 3, 4, 7 and 9) (Delourme *et al.*, 2004), this represents an intriguing aspect of coevolutionary trends and the 'arms race' leading to this cluster-to-cluster relationship.

Cloning of AvrLm4-7 highlighted on a new genomic region the remarkable (for an ascomycete) genome organization of *L. maculans* and the peculiar location of Avr genes in this genome. Both AvrLm1-AvrLm6 and AvrLm4-7 environments show an extremely well-resolved isochore structure, with abrupt changes between GC-equilibrated, gene-rich, recombination-rich regions and AT-rich, gene-poor, recombination-deficient regions (Gout *et al.*, 2006; Fudal *et al.*, 2007). Interestingly, AT-rich isochores constituted of mosaics of transposable elements seem to be linked in *L. maculans* with hosting of Avr genes or effectors. Transposable elements are known to have potentially mutagenic effects, directly through insertion into a gene or aberrant transposition (Kang *et al.*, 2001; Hua-Van *et al.*, 2002; Daboussi and Capy, 2003; Ganko *et al.*, 2003), or indirectly through ectopic recombination between dispersed elements (Nitta *et al.*, 1997; Kim *et al.*, 1998; Hua-Van *et al.*, 2000; Davière *et al.*, 2001; Daboussi and Capy, 2003). They are thus considered as major actors

of eukaryotic genes and chromosomes structural dynamics (Wang *et al.*, 1999; Eichler and Sankoff, 2003), and contribute to increase evolutive ability (Ganko *et al.*, 2003; Kazazian, 2004). Even though usually restricted to small genomic regions, colocalization of transposable elements with *Avr* or pathogenicity genes has already been reported in many bacterial (Kim *et al.*, 1998) or a few fungal/oomycete (Orbach *et al.*, 2000; Rehmany *et al.*, 2003; Rep *et al.*, 2004; Houterman *et al.*, 2008) pathogens, and some 'beneficial' rearrangements have been described, such as gain of virulence through insertion or deletion events of repeated elements (Kang *et al.*, 2001; Farman *et al.*, 2002). In *L. maculans*, rapid adaptation to *Rlm1*-resistance selection pressure was found to be the consequence of large deletion events of the region encompassing *AvrLm1*, favoured by such a genome environment (Gout *et al.*, 2007). Our study also provides preliminary data suggesting that deletion of *AvrLm4-7* might be the major event leading to acquisition of virulence towards *Rlm7*-mediated resistance. It is tempting to hypothesize that the particular location of *AvrLm* genes could have been selected to increase genetic flexibility of pathogenicity factors and that adaptation of *Avr* genes to specific genome environments provides extensive potential for protection from meiotic recombination and rapid adaptation to novel constraints (Fudal *et al.*, 2007; Gout *et al.*, 2007). In addition, repeated sequences are commonly implicated in formation of heterochromatin-like structures less accessible to DNA binding or transcription factors (Grewal and Moazed, 2003). The similar pattern of expression observed for all three *AvrLm* genes during *in planta* infection, although no conservation of regulatory motifs in their promoters was observed (I. Fudal, unpubl. data), also rises the question of incidence of this environment in epigenetic regulation of *AvrLm* genes expression, as shown for telomere-associated secondary metabolite clusters in *Aspergilli* (Bok *et al.*, 2005).

#### *Features of AvrLm4-7: so typical and so atypical at the same time*

Recently van der Does and Rep (2007) reviewed fungal host-determining 'virulence factors' termed effectors, with a focus on small secreted proteins involved in evolutionary arms races between plants and pathogens. *AvrLm1*, *AvrLm6* and *AvrLm4-7* are typical of this family of 'virulence factors' recognized as avirulence factors during 'arms race' and very similar to many fungal *Avr* genes (van der Does and Rep, 2007) (Table 2). However, GC content of *AvrLm1*, *AvrLm6* and *AvrLm4-7* presents a very specific feature. Coding regions in the nuclear genome of *L. maculans* show a *ca.* 52% GC content, as evaluated by analysis of ESTs and gene-rich isochores (Blaise *et al.*, 2007; this study). Compared with these data and with

most fungal *Avr* genes known to date, *AvrLm* genes show an unusually low GC content (39-43%) close to that of their AT-rich genome environment (Table 2). Unusual GC content or codon usage as well as genome clustering was cited by van der Does and Rep (2007) as indicative of possible acquisition via horizontal gene transfer, an hypothesis that will have to be evaluated with whole genome sequence investigation and comparative genomics approaches.

The most striking feature of *AvrLm4-7* is that the gene determines interaction specificities linked to resistance genes *Rlm4* and *Rlm7*. At the moment, it is unclear whether *Rlm4* and *Rlm7* represent allelic forms of the same resistance gene as these genes, along with *Rlm3* and *Rlm9*, are clustered on the same linkage group (Delourme *et al.*, 2004). In addition, *Rlm3* and *Rlm4*, or *Rlm7* and *Rlm3* are never found together in the same pure lines of *B. napus*, whereas they can be found in hybrid genotypes (M.H. Balesdent, unpubl. data). Lack of isolates with *AvrLm4* specificity alone (this study) renders impossible to evaluate the presence of *Rlm4* in *B. napus* genotypes expressing *Rlm7*. Then, *Rlm3*, *Rlm4*, *Rlm7* and *Rlm9* could correspond to a cluster of tightly linked genes, to a unique gene with different alleles, or to a combination of these two hypotheses (Delourme *et al.*, 2004).

*AvrLm4-7* is an extremely monomorphic protein with only five polymorphic locations out of 122 aa in the *AvrLm4-7* mature protein. The common adaptation of pathogens following recognition of one effector as an avirulence determinant either lies in generation of numerous highly divergent alleles, as illustrated by the *AvrL567* locus in the Basidiomycete *M. lini* (Ellis *et al.*, 2007; Wang *et al.*, 2007) or the *ATR1<sup>NdWSB</sup>* loci of the Oomycete *Peronospora parasitica* (Rehmany *et al.*, 2005), or in loss of gene function following partial or complete deletion, insertion of transposable element or premature generation of stop codons following point mutations. Occurrence of *AvrLm* genes in a genome environment strongly inhibiting meiotic recombination seems to be detrimental for diversification of alleles and complete deletion was the favoured mode of evolution towards virulence in *AvrLm1* (Gout *et al.*, 2007). However, in the case where the effector represents a strong fitness advantage for the pathogen, as previously shown for *AvrLm4* (Balesdent *et al.*, 2006; Huang *et al.*, 2006), the most favourable events will be targeted mutation suppressing recognition but maintaining effector function. Our data are consistent with this, as 222 of 256 *avrLm4* isolates have undergone a targeted mutation of base 358 or of any base of the *MscI* restriction site, and only 25 showed a whole or partial deletion of the gene. Thus, *AvrLm4-7* shows a double recognition specificity and the targeted single-base mutation only suppresses one of the recognition specificities. A somewhat similar example was observed in the *Avr3a* gene of the Oomycete *Phytophthora*

*infestans* (Armstrong *et al.*, 2005; Bos *et al.*, 2006) showing non-synonymous mutations on three residues of which two, located in the C-terminal region, are implicated in avirulence towards plant genotypes carrying the *R3a* resistance gene, but also in virulence, by suppressing the *INF1*-mediated plant cell death (Bos *et al.*, 2006). In natural strains of *C. fulvum*, *Avr4*, found to possess a virulence function by protecting fungus against plant chitinases, was shown to circumvent *Cf-4*-mediated resistance through point mutations involved in substitution of cysteine residues and resulting in production of unstable proteins (Joosten *et al.*, 1997; van den Burg *et al.*, 2003). Even though actual virulence role of *AvrLm4-7* is unknown, there are consistent indications that the gene has strong effects on *L. maculans* fitness in the field and in controlled conditions plant inoculations (Balesdent *et al.*, 2006; Huang *et al.*, 2006). The data presented here thus substantiate the fitness advantage linked with maintenance of a functional *AvrLm4-7* gene in absence of *Rlm7* selection pressure, with a diversifying selection targeted on one residue and suppressing only *Rlm4* recognition. The fact that reduced fitness is associated with *AvrLm4-7* alleles characterized by the G<sup>120</sup> → R<sup>120</sup> change suggests that intrinsic function of *AvrLm4-7* as an effector is affected by the single aa mutation. Further research will focus on deciphering whether residue 120 is part of a functional site of the *AvrLm4-7* protein, whose activity could be involved in fitness of *L. maculans*, and/or whether it is part of a (direct or indirect) recognition site for the R gene product. New targeted mutations of base 358 and surrounding ones will also be undertaken to investigate whether other mutations than the G<sup>120</sup> → R<sup>120</sup> can generate virulence towards *Rlm4*. Maintenance of effector while escaping recognition is usually postulated to be linked with direct recognition of effector protein by resistance gene product, whereas indirect recognition will rather favour gene inactivation or deletion (Bent and Mackey, 2007). Our preliminary data suggest that while point mutation seems to be the major event leading to avoidance of *Rlm4*-mediated resistance, deletion of *AvrLm4-7* may be the major event leading to acquisition of virulence towards *Rlm7* resistance, with 25 of 33 *avrLm7* isolates showing whole or partial deletion of *AvrLm4-7*. However, in absence of known *Rlm7*-selection pressure for these isolates, the hypothesis of a difference in recognition mechanism (direct for *Rlm4*, indirect for *Rlm7*) cannot yet be conclusively addressed. Recent use of *Rlm7* resistance in France will provide us with unprecedented opportunities to evaluate evolution of *AvrLm7* specificity. In particular, it will be investigated whether other point mutations can be generated to maintain *AvrLm4-7* putative effector function while avoiding *Rlm7*-mediated recognition, and thus highlight other putative recognition/active sites of the protein to clarify its intrinsic function and recognition mechanism.

## Experimental procedures

### Fungal isolates

v23.1.3 (*a3A4A7a9*) and v23.1.2 (*a3a4A7a9*) isolates originated from in vitro cross #23 (Balesdent *et al.*, 2001). Nzt-4 (*a3a4a7a9*) is a field isolate from New Zealand. Pathogenicity and polymorphism of *AvrLm4-7* was investigated on 300 *L. maculans* field isolates: 219 French isolates of whom 201 originated from a population collected in 2000 and 2001 (Balesdent *et al.*, 2006) and 14 from populations collected in 2005 and 2006 (M.H. Balesdent and G. Daverdin, unpublished); 49 Australian isolates collected in 2005 (M.H. Balesdent and M. Barbetti, unpublished); 23 isolates originating from the International Blackleg of Crucifers Network (IBCN) collection (Balesdent *et al.*, 2005); 7 other Australian isolates (WACxxx and PHW1223); and 2 isolates collected in New Zealand (Table S2 and data not shown). All fungal cultures and highly sporulating cultures were obtained as described (Ansan-Melayah *et al.*, 1995).

### Plant inoculations

Interaction phenotypes were assessed following inoculation with 10 µl droplet of 10<sup>7</sup> ml<sup>-1</sup> conidia suspension on cotyledons of 15-day-old plantlets of *B. napus* cultivar Westar (susceptible control) and cultivars or lines harbouring the *Rlm4* (Jet Neuf, Eleonore, Madrigal, Pixel and Pollen) or the *Rlm7* (Exagone, Hearty, Roxet, 01.23.2.1, 00.4.1.1 and 99.149.1.1) resistance gene. They were also inoculated on line 00.22.1.1 carrying *Rlm3*, and cultivar Darmor carrying *Rlm9*. These two resistance genes specify recognition of the avirulence genes *AvrLm3* and *AvrLm9* genetically clustered with *AvrLm4-7* (Balesdent *et al.*, 2002; 2005). Symptoms were scored on 10–12 plants per interaction, 14–21 days after inoculation using the IMAScore rating scale comprising six infection classes (IC), where IC1 to IC3 correspond to avirulent isolates (*AvrLm*) and IC4 to IC6 to virulent isolates (*avrLm*) (Balesdent *et al.*, 2006).

Time course studies of in planta *AvrLm4-7* gene expression were essentially as described by Fudal *et al.* (2007): cotyledons of plantlets of Westar were inoculated by a conidia suspension of v23.1.3, and sampling was performed at 3, 5, 7, 9 or 12 dpi. For analysis of expression during stem lesion development, junctions of the second leaf petiole and stem of plantlets of Westar were punctured three times and inoculated with a 10 µl droplet of v23.1.3 conidial suspension (10<sup>7</sup> ml<sup>-1</sup>). Plants were incubated in a growth chamber similarly to inoculated cotyledons, and stem necrosis tissues were collected 2 months after inoculation (Fudal *et al.*, 2007).

### Genetic mapping

In order to generate a high-density genetic map of the *AvrLm3-4-7* locus, 89 progeny of cross #23 [v11.3.6 (*a3a4A7*) × 21.3.1 (*a3A4A7*)], and 230 progeny of cross #33 [v23.1.11 (*a3A4A7*) × Nzt-4 (*a3a4a7*)], were used. Linkage analyses among loci were performed using the MapMaker/EXP version 3.0b software (available online; Whitehead/MIT Center for Genome Research, Cambridge, MA, USA) with a

LOD score of 3.0 and a maximum recombination frequency of 30 centimorgans (Kuhn *et al.*, 2006).

#### DNA and RNA manipulations

For PCR, genomic DNA was extracted from conidia with the DNeasy 96 Plant Kit (Qiagen S.A., Courtaboeuf, France) as already described (Attard *et al.*, 2002). Primer sequences used for the chromosome walk and sequence analysis of *AvrLm4-7* in natural isolates are displayed in Table S3. PCR reactions were done as previously described (Fudal *et al.*, 2007; Remy *et al.*, 2008). Sequencing was performed using a Beckman Coulter CEQ 8000 automated sequencer (Beckman Coulter, Fullerton, CA, USA) according to the manufacturer's instructions. Procedures for gel electrophoresis and Southern blotting have been reported (Attard *et al.*, 2005). Total RNA extraction, treatment and single-strand cDNA synthesis were performed as described in Fudal *et al.* (2007).

#### Quantitative RT-PCR

qRT-PCR was performed using 7700 real-time PCR equipment (Applied Biosystems, Foster City, CA, USA) and Absolute SYBR Green ROX dUTP Mix (ABgene, Courtaboeuf, France) as described by Fudal *et al.* (2007). For each condition tested, two RNA extractions from different biological samples and two reverse transcriptions for each biological repeat were performed. RNA extracted from uninfected leaves and water were used as negative controls. A standard curve of amplification efficiency for  *$\beta$ -tubulin*, *actin* and *AvrLm4-7* primers (Supplementary Table S3) was generated from purified RT-PCR products corresponding to 12-dpi-infected plant tissues across a three-orders-of-magnitude dilution series ( $10^{-1}$ - $10^{-3}$  dilutions). Samples were analysed from a 10-fold dilution of the original RT products. Diluted RT products (5  $\mu$ l) were added to 12.5  $\mu$ l of PCR mix and 7.5  $\mu$ l of primers to a final concentration of 300 nM. The cycling profile consisted of an initial cycle of 50°C for 2 min, a second cycle of 95°C for 15 min, followed by 40 cycles of 95°C for 15 second and 60°C for 1 min. The amplified product was detected every cycle at the end of the 60°C step. After cycling, dissociation curve analysis confirmed absence of non-specific products in the reaction (temperature increases from 60°C to 95°C in 20 min). Fluorescence threshold (Ct) value was determined at 0.1 of fluorescence intensity. Ct values were analysed as described by Muller *et al.* (2002).  *$\beta$ -tubulin* was used as a constitutive reference gene.

#### Gene annotation and polymorphism analysis

Annotations of UTRs, transcription start and stop sites and intron positions were performed following PCR amplification and sequencing of 3' and 5' ends of cDNA using the BD SMART RACE cDNA Amplification Kit (Clontech, Palo Alto, CA, USA) according to the manufacturer's recommendations and using primers AVR47int-Up/-Lo (Table S3). For sequencing of *AvrLm4-7*, the gene was amplified with primers AVR47ext-Up/-Lo, and the purified PCR product was sequenced using primers AVR47ext-Up,-Lo,-Up2,-Up3, and AVR47int-Up,-Lo (Table S3).

#### BAC libraries and chromosome walk

Construction and screening of the v23.1.3 isolate genomic DNA HindIII and NdeII BAC libraries, have been described previously (Attard *et al.*, 2002; Gout *et al.*, 2006). End sequences of BAC clones or subclones were obtained as described (Attard *et al.*, 2005) and used to identify SNPs or specific junctions spanning two truncated repeats to generate PCR or CAPS markers. Relative position on the physical map and overlap between BAC clones were estimated with PCR-amplified markers derived from BAC-(sub)clone end sequences and by analysing the restriction pattern of the BAC clones following digestion with restriction enzymes such as HindIII, BamHI, EcoRI or EcoRV.

#### BAC sequencing and annotation

Sequencing, finishing and assembling of the three BAC clones were done at Genoscope (Evry, France) as previously described (Gout *et al.*, 2006). Sequences were analysed using BLAST programs against public databases of nucleotide and protein sequences, including the National Center for Biotechnology Information non-redundant databases, and against available sequence releases of fungal genomes (Broad Institute, Cambridge, MA, USA). Sequence analyses were also performed using the SignalP 3.0 and TargetP 1.1 programs (<http://www.cbs.dtu.dk/services/SignalP/>); the MEME program (<http://meme.nbcr.net/meme/intro.html>) for comparison with *AvrLm1* and *AvrLm6* sequences; and by searches against the Pfam protein families database (<http://pfam.jouy.inra.fr/>). Gene prediction was performed with FGENESH (<http://www.softberry.com>) set with *A. nidulans*, *F. graminearum*, *M. grisea* and *N. crassa* matrixes. Annotation of the 238 kb-sequenced contig was facilitated with Artemis (Rutherford *et al.*, 2000; <http://www.sanger.ac.uk/Software/Artemis/>). This software was also used to compute GC content, using a sliding window of 400 bp.

For novel transposable elements identified here, AT-rich sequences showing no homology with previously characterized retrotransposons (Attard *et al.*, 2005; Gout *et al.*, 2006) were blasted using BLAST2 ([http://bioinfo.hku.hk/services/analyseq/cgi-bin/blast2\\_in.pl](http://bioinfo.hku.hk/services/analyseq/cgi-bin/blast2_in.pl)) against the draft 5 $\times$  whole genome sequence of *L. maculans* available at the time of the study. Highly repeated sequences thus identified were aligned and extended at both extremities to generate a consensus sequence. Presence at both ends of direct or inverted repeats (IR), resembling those of LTR retrotransposons or IR of DNA transposons indicated the complete element was identified. BLAST searches were further used to identify genes and gene organization, thus allowing assignment of the novel repeats to a particular family of transposable element.

#### Vector constructions and transformation of fungal isolates

The binary vector PZPnat1 (a gift from Dr Barbara Howlett, University of Melbourne, Australia), which carries the nourseothricin acetyltransferase *nat1* gene (conferring resistance to the antibiotic nourseothricin) under control of the *A. nidulans* trpC promoter as a selectable marker, was used

for *A. tumefaciens*-mediated transformation of *L. maculans*. v23.1.3 allele of *AvrLm4-7* was isolated by digestion of the BAC clone 1B4 with *PacI*. A ~5 kb fragment was isolated following agarose gel electrophoresis, purified, and digested with *NheI* and *RsaI*. The resulting 2353 bp fragment was ligated into PZPnat1 previously digested with *HpaI* and *SpeI* to generate compatible ends. The construct was termed PZPnat1-A4A7. Cloning of the *avrLm4-AvrLm7* allele from isolate v23.1.2 was performed by PCR amplification using the AVR47-JN2-Up/-Lo primers (Table S3). Following digestion of the PCR product with *SnaBI* and *BstXI*, and gel purification, the 1.8 kb fragment generated was used to replace the excised corresponding sequence of PZPnat1-A4A7. The construct was termed PZPnat1-a4A7. Constructs were then introduced into *A. tumefaciens* strain C58 by electroporation, and transformation of *L. maculans* was performed as previously described (Gout *et al.*, 2006). Fungal transformants were selected on 50 µg ml<sup>-1</sup> nourseothricin (WERNER Bio-Agents, Jena, Germany).

#### PCR-mediated site-directed mutagenesis

Using the PZPnat1-a4A7 construct as template, the primer MD-invers (with a G<sup>358</sup> instead of C<sup>358</sup>) was used in combination with primer AVR47-JN2-Lo (Table S3) in a PCR reaction using the high fidelity DNA polymerase Taq Phusion (Finnzymes) to generate a 425 bp fragment MP1. MP1 was then used as a megaprimer in combination with AVR47-JN2-Up in a Taq Phusion PCR reaction to generate an 1843 bp product termed a4\*A7 for the mutagenized v23.1.2 allele. Following restriction with *SnaBI* and *BstXI* and gel purification, the fragment was used to replace the excised corresponding sequence of PZPnat1-a4A7. The construct, termed PZPnat1-a4\*A7, was cloned in *Escherichia coli*, reextracted and accuracy of the point mutation checked by sequencing. Finally, the construct was introduced into *A. tumefaciens* and used for transformation of isolates Nzt-4 (a4a7) and v23.1.2 (a4A7).

#### Acknowledgements

This work was partly funded by a grant from the CETIOM, and F. Parlange by a fellowship from the ANRT (CIFRE bursary). BAC sequencing was performed via a grant from INRA-PPV (AIP 2005 'Séquençage et CRB'). We are grateful to L. Gout (INRA-Bioger) for critical reading and helpful comments on the manuscript. We thank E. Mendes-Pereira and A. Gautier (INRA-BIOGER, France) for sequencing assistance, E. Remy (INRA-BIOGER, France) for help in qRT-PCR experiments, L. Coudard and M. Chabirand (INRA-BIOGER, France) for plant management.

#### References

Ansan-Melayah, D., Balesdent, M.H., Buee, M., and Rouxel, T. (1995) Genetic characterization of *AvrLm1*, the first avirulence gene of *Leptosphaeria maculans*. *Phytopathology* 85: 1525-1529.  
 Armstrong, M.R., Whisson, S.C., Pritchard, L., Bos, J.I.B., Venter, E., Avrova, A.O., *et al.* (2005) An ancestral

oomycete locus contains late blight avirulence gene *Avr3a*, encoding a protein that is recognized in the host cytoplasm. *Proc Natl Ac Sc USA* 102: 7766-7771.  
 Attard, A., Gout, L., Gourgues, M., Kuhn, M.L., Schmit, J., Laroche, S., *et al.* (2002) Analysis of molecular markers genetically linked to the *Leptosphaeria maculans* avirulence gene *AvrLm1* in field populations indicates a highly conserved event leading to virulence on *Rlm1* genotypes. *Mol Plant Microbe Interact* 15: 672-682.  
 Attard, A., Gout, L., Ross, S., Parlange, F., Cattolico, L., Balesdent, M.H., and Rouxel, T. (2005) Truncated and RIP-degenerated copies of the LTR retrotransposon *Pholy* are clustered in a pericentromeric region of the *Leptosphaeria maculans* genome. *Fungal Genet Biol* 42: 30-41.  
 Balesdent, M.H., Attard, A., Ansan-Melayah, D., Delourme, R., Renard, M., and Rouxel, T. (2001) Genetic control and host range of avirulence toward *Brassica napus* cultivars Quinta and Jet Neuf in *Leptosphaeria maculans*. *Phytopathology* 91: 70-76.  
 Balesdent, M.H., Attard, A., Kuhn, M.L., and Rouxel, T. (2002) New avirulence genes in the phytopathogenic fungus *Leptosphaeria maculans*. *Phytopathology* 92: 1122-1133.  
 Balesdent, M.H., Barbetti, M.J., Li, H., Sivasithamparam, K., Gout, L., and Rouxel, T. (2005) Analysis of *Leptosphaeria maculans* race structure in a worldwide collection of isolates. *Phytopathology* 95: 1061-1071.  
 Balesdent, M.H., Louvard, K., Pinochet, X., and Rouxel, T. (2006) A large-scale survey of races of *Leptosphaeria maculans* occurring on oilseed rape in France. *Eur J Plant Pathol* 114: 53-65.  
 Bent, A.F., and Mackey, D. (2007) Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. *Annu Rev Phytopathol* 45: 399-436.  
 Bernardi, G. (2000) Isochores and the evolutionary genomics of vertebrates. *Gene* 241: 3-17.  
 Blaise, F., Remy, E., Meyer, M., Zhou, L., Narcy, J.P., Roux, J., *et al.* (2007) A critical assessment of *Agrobacterium tumefaciens*-mediated transformation as a tool for pathogenicity gene discovery in the phytopathogenic fungus *Leptosphaeria maculans*. *Fungal Genet Biol* 44: 123-138.  
 Böhnert, H.U., Fudal, I., Dìoh, W., Tharreau, D., Notteghem, J.L., and Lebrun, M.H. (2004) A putative polyketide synthase peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. *Plant Cell* 16: 2499-2513.  
 Bok, J.W., Balajee, S.A., Marr, K.A., Andes, D., Nielsen, K.F., Frisvad, J.C., and Keller, N.P. (2005) *LaeA*, a regulator of morphogenetic fungal virulence factors. *Euk Cell* 4: 1574-1582.  
 Bos, J.I., Kanneganti, T.D., Young, C., Cakir, C., Huitema, E., Win, J., *et al.* (2006) The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*. *Plant J* 48: 165-176.  
 van den Burg, H.A., Westerink, N., Francoijs, K.J., Roth, R., Woestenenk, E., Boeren, S., *et al.* (2003) Natural disulfide bond-disrupted mutants of AVR4 of the tomato pathogen *Cladosporium fulvum* are sensitive to proteolysis, circumvent Cf-4-mediated resistance, but retain their chitin binding ability. *J Biol Chem* 278: 27340-27346.

- Catanzariti, A.M., Dodds, P.N., Lawrence, G.J., Ayliffe, M.A., and Ellis, J.G. (2006) Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* 18: 243-256.
- Daboussi, M.J., and Capy, P. (2003) Transposable elements in filamentous fungi. *Ann Rev Microbiol* 7: 75-99.
- Davière, J.M., Langin, T., and Daboussi, M.J. (2001) Potential role of transposable elements in the rapid reorganization of the *Fusarium oxysporum* genome. *Fungal Genet Biol* 34: 177-192.
- Delourme, R., Pilet-Nayel, M.L., Archipiano, M., Horvais, R., Tanguy, X., Rouxel, T., *et al.* (2004) A cluster of major specific resistance genes to *Leptosphaeria maculans* in *Brassica napus*. *Phytopathology* 94: 578-583.
- Dodds, P.N., Lawrence, G.J., Catanzariti, A.-M., Ayliffe, M.A., and Ellis, J.G. (2004) The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell* 16: 755-768.
- van der Does, H.C., and Rep, M. (2007) Virulence genes and the evolution of host specificity in plant-pathogenic fungi. *Mol Plant Microbe Interact* 20: 1175-1182.
- Eichler, E.E., and Sankoff, D. (2003) Structural dynamics of eukaryotic chromosome evolution. *Science* 301: 793-797.
- Ellis, J.G., Dodds, P.N., and Lawrence, G.J. (2007) Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annu Rev Phytopathol* 45: 289-306.
- Farman, M.L., Eto, Y., Nakao, T., Tosa, Y., Nakayashiki, H., Mayama, S., and Leong, S.A. (2002) Analysis of the structure of the AVR1-CO39 avirulence locus in virulent rice-infecting isolates of *Magnaporthe grisea*. *Mol Plant Microbe Interact* 15: 6-16.
- Flor, H.H. (1971) Current status of the gene-for-gene concept. *Ann Rev Phytopathol* 9: 275-296.
- Fudal, I., Ross, S., Gout, L., Blaise, F., Kuhn, M.L., Eckert, M., *et al.* (2007) Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria maculans* genome: map-based cloning of AvrLm6. *Mol Plant Microbe Interact* 20: 459-470.
- Ganko, E.W., Bhattachajee, V., Schliekelman, P., and McDonald, J.F. (2003) Evidence for the contribution of LTR retrotransposons to *C. elegans* gene evolution. *Mol Biol Evol* 20: 1925-1931.
- Gout, L., Fudal, I., Kuhn, M.L., Blaise, F., Cattolico, L., Balesdent, M.H., and Rouxel, T. (2006) Lost in the middle of nowhere: the AvrLm1 avirulence gene of *Leptosphaeria maculans*. *Mol Microbiol* 60: 67-80.
- Gout, L., Kuhn, M.L., Vincenot, L., Bernard-Samain, S., Cattolico, L., Barbetti, M., *et al.* (2007) Genome structure impacts molecular evolution at the AvrLm1 avirulence locus of the plant pathogen *Leptosphaeria maculans*. *Environ Microbiol* 9: 2978-2992.
- Greenberg, J.T., and Yao, N. (2004) The role and regulation of programmed cell death in plant-pathogen interactions. *Cell Microbiol* 6: 201-211.
- Grewal, S.I., and Moazed, D. (2003) Heterochromatin and epigenetic control of gene expression. *Science* 301: 798-802.
- Houterman, P.M., Cornelissen, B.J.C., and Rep, M. (2008) Suppression of plant resistance gene-based immunity by a fungal effector. *PLoS Pathog* 4: e1000061.
- Huang, Y.J., Li, Z.Q., Evans, N., Rouxel, T., Fitt, B.D.L., and Balesdent, M.H. (2006) Fitness cost associated with loss of the AvrLm4 avirulence function in *Leptosphaeria maculans* (Phoma stem canker of oilseed rape). *Eur J Plant Pathol* 114: 77-89.
- Hua-Van, A., Daviere, J.M., Kaper, F., Langin, T., and Daboussi, M.J. (2000) Genome organization in *Fusarium oxysporum*: clusters of class II transposons. *Curr Genet* 37: 339-347.
- Hua-Van, A., Langin, T., and Daboussi, M.J. (2002) Aberrant transposition of a Tc1-mariner element, impala, in the fungus *Fusarium oxysporum*. *Mol Genet Genomics* 267: 79-87.
- Joosten, M.H.A.J., Vogelsang, R., Cozijnsen, T.J., Verberne, M.C., and De Wit, P.J.G.M. (1997) The biotrophic fungus *Cladosporium fulvum* circumvents Cf-4-mediated resistance by producing unstable AVR4 elicitors. *Plant Cell* 9: 367-379.
- Kang, S., Lebrun, M.H., Farrall, L., and Valent, B. (2001) Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. *Mol Plant Microbe Interact* 14: 671-674.
- Kazazian, H.H. (2004) Mobile elements: drivers of genome evolution. *Science* 303: 1626-1632.
- Kim, J.F., Charkowski, A.O., Alfano, J.R., Collmer, A., and Beer, S.V. (1998) Sequences related to transposable elements and bacteriophages flank avirulence genes of *Pseudomonas syringae*. *Mol Plant Microbe Interact* 11: 1247-1252.
- Kuhn, M.L., Gout, L., Howlett, B.J., Melayah, D., Meyer, M., Balesdent, M.H., and Rouxel, T. (2006) Genetic linkage maps and genomic organization in *Leptosphaeria maculans*. *Eur J Plant Pathol* 114: 17-31.
- Muller, P.Y., Janovjak, H., Miserez, A.R., and Dobbie, Z. (2002) Processing of gene expression data generated by quantitative real-time RT-PCR. *Biotechniques* 32: 1372-1378.
- Nitta, N., Farman, M.L., and Leong, S.A. (1997) Genome organization of *Magnaporthe grisea*: integration of genetic maps, clustering of transposable elements and identification of genome duplications and rearrangements. *Theor App Genet* 95: 20-32.
- Oliver, J.L., Bernaola-Galvan, P., Carpena, P., and Roman-Roldan, R. (2001) Isochore chromosome maps of eukaryotic genomes. *Gene* 276: 47-56.
- Orbach, M.J., Farrall, L., Sweigard, J.A., Chumley, F.G., and Valent, B. (2000) A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*. *Plant Cell* 12: 2019-2032.
- Rehmany, A.P., Grenville, L.J., Gunn, N.D., Allen, R.L., Paninyk, Z., Byrne, J., *et al.* (2003) A genetic interval and physical contig spanning the *Peronospora parasitica* (At) avirulence gene locus *ATR1Nd*. *Fungal Genet Biol* 38: 33-42.
- Rehmany, A.P., Gordon, A., Rose, L.E., Allen, R.L., Armstrong, M.R., Whisson, S.C., *et al.* (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two *Arabidopsis* lines. *Plant Cell* 17: 1839-1850.

- Remy, E., Meyer, M., Blaise, F., Chabirand, M., Wolff, N., Balesdent, M.-H., and Rouxel, T. (2008) The *Lmpma1* gene of *Leptosphaeria maculans* encodes a plasma membrane H<sup>+</sup>-ATPase isoform essential for pathogenicity towards oilseed rape. *Fungal Genet Biol* 45: 1122-1134.
- Rep, M. (2005) Small proteins of plant-pathogenic fungi secreted during host colonization. *FEMS Microbiol Lett* 253: 19-27.
- Rep, M., van der Does, H.C., Meijer, M., van Wijk, R., Houterman, P.M., Dekker, H.L., *et al.* (2004) A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for I-3-mediated resistance in tomato. *Mol Microbiol* 53: 1373-1383.
- Ridout, C.J., Skamnioti, P., Porritt, O., Sacristan, S., Jones, J.D.G., and Brown, J.K.M. (2006) Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *Plant Cell* 18: 2402-2414.
- Rivas, S., and Thomas, C.M. (2005) Molecular interactions between tomato and the leaf mold pathogen *Cladosporium fulvum*. *Annu Rev Phytopathol* 43: 395-436.
- Rohe, M., Gierlich, A., Hermann, H., Hahn, M., Schmidt, B., Rosahl, S., and Knogge, W. (1995) The race-specific elicitor, NIP1, from the barley pathogen, *Rhynchosporium secalis*, determines avirulence on host plants of the Rrs1 resistance genotype. *EMBO J* 14: 4168-4177.
- Rouxel, T., and Balesdent, M.H. (2005) The stem canker (blackleg) fungus, *Leptosphaeria maculans*, enters the genomic era. *Mol Plant Pathol* 6: 225-241.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., and Barrell, B. (2000) Artemis: sequence visualisation and annotation. *Bioinformatics* 16: 944-945.
- Wang, C.-I.A., Gunèar, G., Forwood, J.K., Teh, T., Catanzariti, A.M., Lawrence, G.J., *et al.* (2007) Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. *Plant Cell* 19: 2898-2912.
- Wang, S., Liu, N., Peng, K., and Zhang, Q. (1999) The distribution and copy number of copia-like retrotransposons in rice (*Oryza sativa* L.) and their implications in the organization and evolution of the rice genome. *Proc Natl Acad Sci USA* 96: 6824-6828.

### Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.



## Supplementary data:

**Supplementary Table S1 (1/3).** Genes predicted in *AvrLm7* genomic region and homology in the databases.

Gene #	Position	Best BLAST matches in databases			Homologies with <i>Stagonospora nodorum</i> genome <sup>a</sup>		
		Predicted function ; domains (organism)	E-value	Accessions	Location	E-value	Accessions
1.	AT1	hypothetical protein AN3281.2; Cytochrome P450 (Aspergillus nidula )	2e-82	XP_660885	Supercontig 1.7	0.0	SNOG_04820.1
2.	AT1	unnamed protein product; Esterase_lipase ( <i>Aspergillus oryzae</i> )	3e-12	BAE64659	Supercontig 1.5	2e-11	SNOG_03619.1
3.	GCI	hypothetical protein AN8426.2; GAL4-like Zn2Cys6 binuclear cluster DNA-binding domain found in transcription regulators like GAL4 ( <i>Aspergillus nidulans</i> )	3e-06	XP_681695	-	-	-
4.	GCI	hypothetical protein CIMG_05421; Beta-glucan synthesis-associated protein (SKN1) ( <i>Coccidioides immitis</i> )	8e-141	EAS34397	Supercontig 1.8	0.0	SNOG_05933.1
5.	GCI	chitin synthase class VI ( <i>Coccidioides immitis</i> )	0.0	AAP74955	Supercontig 1.8	0.0	SNOG_05923.1 <sup>b</sup>
6.	GCI	unnamed protein product; Middle domain of eukaryotic initiation factor 4G (eIF4G) ( <i>Aspergillus oryzae</i> )	0.0	BAE55474	Supercontig 1.8	6e-31	SNOG_05925.1 <sup>b</sup>
7.	GCI	hypothetical protein FG07783.1; Histidine acid phosphatase ( <i>Fusarium graminearum</i> )	4e-158	XP_387959	Supercontig 1.8	0.0	SNOG_05929.1
8.	GCI	conserved hypothetical protein; Eps15 homology domain found in proteins implicated in endocytosis, vesicle transport, and signal transduction ( <i>Chaetomium globosum</i> )	3e-130	EAQ87269	Supercontig 1.8	0.0	SNOG_05930.1
9.	GCI	chitin synthase ( <i>Paracoccidioides brasiliensis</i> )	0.0	AAD19613	Supercontig 1.26	0.0	SNOG_13325.1
10.	GCI	chitin synthase ChsE ( <i>Aspergillus fumigatus</i> )	0.0	XP_755677	Supercontig 1.26	0.0	SNOG_13326.1
11.	GCI	putative mitochondrial translation initiation factor IF-2 ( <i>Aspergillus fumigatus</i> )	0.0	XP_750147	Supercontig 1.26	0.0	SNOG_13327.1
12.	GCI	hypothetical protein FG08963.1; ADP-ribose diphosphatase ( <i>Fusarium graminearum</i> )	3e-57	XP_389139	Supercontig 1.8	0.0	SNOG_05921.1
13.	GCI	putative FAD dependent oxidoreductase ( <i>Aspergillus fumigatus</i> )	1e-131	XP_750158	Supercontig 1.8	0.0	SNOG_05922.1
14.	GCI	putative SRP signal recognition particle protein ( <i>Aspergillus fumigatus</i> )	4e-131	XP_750144	Supercontig 1.26	0.0	SNOG_13331.1

**Supplementary Table S1 (2/3)** Genes predicted in AvrLm7 genomic region and homology in the databases.

15.	GCI	unnamed protein product; PPR repeat ( <i>Aspergillus oryzae</i> )	4e-57	BAE58070	Supercontig 1.26	0.0	SNOG_13330.1
16.	GCI	No homology			-	-	-
17.	GCI	<u>domain 1</u> predicted protein ( <i>Fusarium graminearum</i> )	9e-20	XP_381954	Supercontig 1.26	0.0	SNOG_13332.1
		<u>domain 2</u> GATA Zn-finger-containing transcription factor AREB	1e-23	AAG49351			
18.	GCI	gamma ( <i>Aspergillus nidulans</i> ) predicted protein; ATP synthase subunit C ( <i>Magnaporthe grisea</i> )	2e-27	XP_362034	<i>Supercontig 1.26</i>	6e-44	SNOG_13334.1
19.	GCI	thermotolerance protein ( <i>Aspergillus fumigatus</i> )	9e-58	AAT58665	Supercontig 1.26	0.0	SNOG_13342.1
20.	GCI	hypothetical protein CIMG_05000; Probable Rab-GAPs, GTP-activator activity on Rab-like GTPases ( <i>Coccidioides immitis</i> )	1e-85	EAS33976	Supercontig 1.26	0.0	SNOG_13343.1
21.	GCI	No homology	-	-	Supercontig 1.26	9e-16	SNOG_13345.1
22.	GCI	No homology	-	-	-	-	-
23.	GCI	No homology	-	-	Supercontig 1.26	1e-31	SNOG_13348.1
24.	GCI	hypothetical protein; O-methyltransferase ( <i>Chaetomium globosum</i> )	5e-72	EAQ89416	Supercontig 1.16	0.0	SNOG_10068.1
25.	GCI	No homology	-	-	Supercontig 1.12	0.0	SNOG_08285.1
26.	GCI	hypothetical protein FG10290.1; EndoIII-related endonuclease, DNA Glycosylase C ( <i>Fusarium graminearum</i> )	5e-106	XP_390466	Supercontig 1.13	0.0	SNOG_08612.1
27.	GCI	hypothetical protein MGG_07103 ( <i>Magnaporthe grisea</i> )	4e-29	XP_367178	Supercontig 1.8	0.0	SNOG_05911.1
28.	GCI	hypothetical protein ( <i>Neurospora crassa</i> )a	4e-18	XP_961755	Supercontig 1.8	0.0	SNOG_05912.1
29.	GCI	hypothetical protein AN6269.2; Translocation protein Sec62 ( <i>Aspergillus nidulans</i> )	4e-93	XP_663873	Supercontig 1.8	0.0	SNOG_05913.1
30.	GCI	UDP-N-acetylglucosaminyltransferase ( <i>Aspergillus fumigatus</i> )	0.0	XP_750088	Supercontig 1.8	0.0	SNOG_05914.1
31.	GCI	No homology			Supercontig 1.8	1e-07	SNOG_06024.1
32.	GCI	hypothetical protein CIMG_05464; Vesicle transport v-SNARE protein ( <i>Coccidioides immitis</i> )	3e-75	EAS34440	Supercontig 1.8	0.0	SNOG_06026.1
33.	GCI	hypothetical protein MGG_02107; Peptidase_M28 ( <i>Magnaporthe grisea</i> )	6e-103	XP_365405	Supercontig 1.8	0.0	SNOG_06031.1 <sup>b</sup>
					Supercontig 1.8	3.7e-40	SNOG_06032.1 <sup>b</sup>

**Supplementary Table S1 (3/3).** Genes predicted in AvrLm7 genomic region and homology in the databases.

34.	GCI	nitrite reductase ( <i>Williams et al., 1995</i> ) ( <i>Leptosphaeria maculans</i> )	AAA82739		Supercontig 1.8	0.0	SNOG_06023.1
35.	GCI	nitrate reductase ( <i>Williams et al., 1995</i> ) ( <i>Leptosphaeria maculans</i> )	AAA82740		Supercontig 1.8	0.0	SNOG_06022.1
36.	GCI	hypothetical protein CHGG_04791; DEAD-like helicases superfamily ( <i>Chaetomium globosum</i> )	EAQ88172	0.0	Supercontig 1.8	0.0	SNOG_06021.1
37.	GCI	hypothetical protein ( <i>Neurospora crassa</i> )	XP_955825	4e-16	Supercontig 1.8	0.0	SNOG_06020.1
38.	AT2	No homology	-	-	-	-	-
39.	AT2	No homology	-	-	-	-	-
40.	AT2	No homology	-	-	-	-	-

<sup>a</sup> All homologies detected in *S. nodorum* genome concerned hypothetical proteins.

<sup>b</sup> Consecutive *S. nodorum* sequences detected by BLAST P with *L. maculans* genes #5 and #33 potentially correspond to different regions of a same gene.

Supplementary Table S2 .List and characteristics of the reference or field isolates whose AvrLm4-7 allele was sequenced in the present study.

Isolate	Year of isolation	Geographical origin	Phenotypes <sup>a</sup>		PCR amplifications <sup>b</sup>	
			AvrLm4	AvrLm7	AVR47internal primers	AVR47external primers
<b>Reference isolates</b>						
V23.1.3	-	-	+	+	+	+
V23.1.2	-	-	-	+	+	+
Nzt-4	unknown	New Zealand	-	-	-	-
<b>Field isolates</b>						
IBCN17	1988	Australia	+	+	+	+
IBCN74	1979	France	+	+	+	+
IBCN77	1987	Australia	+	+	+	+
IBCN87	1989	Canada	+	+	+	+
IBCN88	1989	Canada	+	+	+	+
WT71	2005	Australia	+	+	nd	+
B1.6	2000	France	+	+	nd	+
M1.2	2000	France	+	+	nd	+
J1.3	2000	France	+	+	nd	+
K11	2000	France	+	+	nd	+
C-04	2006	France	+	+	+	+
IBCN02	1992	Germany	-	+	+	+
IBCN03	1991	Germany	-	+	+	+
IBCN73	unknown	Germany	-	+	+	+
WT14	2005	Australia	-	+	nd	+
a2	1990	France	-	+	+	+
H5	1990	France	-	+	+	+
B1.1	2000	France	-	+	nd	+
M1.25	2000	France	-	+	nd	+
J1.1	2000	France	-	+	nd	+
K1	2000	France	-	+	nd	+
P7	2000	France	-	+	nd	+
R4.8	2005	France	-	+	nd	+
IBCN04	1992	Germany	-	+	+	+
IBCN85	1989	Canada	-	+	+	+
WT02	2005	Australia	-	+	nd	+
WT12	2005	Australia	-	+	nd	+
WT88	2005	Australia	-	+	+	+
WT09	2005	Australia	-	+	nd	+
M1.1	2000	France	-	+	nd	+
P1	2000	France	-	+	nd	+
Hearty 6,1 & 6,2	2005	France	-	+	+	+
B-02	2006	France	-	+	+	+
C-03	2006	France	-	+	+	+
G-01	2006	France	-	+	+	+
L-05	2006	France	-	+	+	+
G06-E115 & E116	2006	France	-	+	+	+
G06-E125	2006	France	-	+	+	+
G06-E204	2006	France	-	+	+	+
G06-202	2006	France	-	+	+	+
G06-203	2006	France	-	+	+	+
S-02	2006	France	-	+	+	+

<sup>a</sup> Interaction phenotypes : (-) virulence; (+) avirulence.

<sup>b</sup> PCR amplification: (-) absence; (+) presence. nd: not done.



# **Chapitre 2**

## **Evolution moléculaire du gène d'avirulence**

### ***AvrLm4-7* sous pression de sélection**



## Introduction

Le premier cultivar de *B. napus* possédant le gène de résistance *Rlm7*, Roxet, a été inscrit en France en 2002 mais n'a été cultivé en 2002-2003 que sur des surfaces très limitées (Pinochet, Comm. Pers.). L'utilisation à grande échelle de cultivars possédant *Rlm7* a lieu peu de temps après avec la réussite commerciale dès 2005-2006 des hybrides Exagone (inscrit en 2004), Extend et Exocet (2005). Ce dernier est dans le « Top 5 » des cultivars les plus cultivés depuis la saison 2006-2007 (1<sup>er</sup> en 2007-2008) (Anonymous, 2010). Etant donné la pression de sélection qu'imposent de tels cultivars utilisés à grande échelle sur la population française de *L. maculans* très largement avirulente vis-à-vis de ce gène de résistance (Balesdent *et al.*, 2006) et l'historique de contournement de gènes de résistance par *L. maculans* (Rouxel *et al.*, 2003; Sprague *et al.*, 2006), le contournement de *Rlm7* était envisageable en seulement quelques années.

L'utilisation récente de cultivars *Rlm7* était l'occasion de mettre en place une expérimentation permettant d'évaluer la durabilité de cette résistance et l'influence des pratiques culturales sur celle-ci. Pour la première fois chez un champignon phytopathogène, nous pouvions nous baser sur la connaissance de la structure d'un gène d'avirulence et de son environnement génomique (Parlange, 2006) pour évaluer comment les populations fongiques évolueraient lorsqu'elles sont soumises à la pression de sélection du gène de résistance spécifique correspondant (ici *Rlm7*) dans les conditions du champ.

Une étude portant sur une période de 4 cycles culturaux menée sur deux sites expérimentaux indépendants et différant seulement par leurs modes de culture a alors été mise en place en 2004. Sur le premier site (Grignon) était cultivé du colza en monoculture avec un travail du sol simplifié tandis que sur le second site (Versailles), le mode de culture incluait rotation culturale ainsi qu'enfouissement par labour des résidus de cultures.

L'isolement de souches de *L. maculans* issues de ces deux sites a permis d'établir une importante collection de souches virulentes et avirulentes vis-à-vis de *Rlm7*. L'analyse de cette collection montre un fort contraste entre les deux sites expérimentaux. En effet, en trois ans, moins de 1% des souches prélevées sur cultivar *rlm7* à Versailles étaient virulentes vis-à-vis de *Rlm7* tandis qu'à Grignon, 30% des souches prélevées sur cultivar *rlm7* après trois années d'expérimentation étaient virulentes vis-à-vis de *Rlm7*. Cette différence démontre l'importance des pratiques culturales dans le maintien de l'efficacité de la résistance *Rlm7* dans le temps.

L'analyse moléculaire des souches de la collection montre par ailleurs que la présence de souches virulentes était principalement due à la génération, par mutation du gène *AvrLm4-7*, de souches virulentes au sein même de la parcelle de culture. En effet, un grand nombre d'évènements de mutation indépendants affectant *AvrLm4-7* et permettant d'expliquer la virulence des souches vis-à-vis de *Rlm7* a pu être répertorié, la majorité de ces mutations étant liée à la reproduction sexuée du



champignon. Le fait que les éléments de pratique culturale qui ont été retenus sur la parcelle de Grignon soient justement ceux qui favorisent la reproduction sexuée et l'accomplissement d'un cycle infectieux complet souligne l'importance de la reproduction sexuée du champignon dans sa capacité à contourner *Rlm7*, et donc plus globalement un gène de résistance spécifique.

L'évaluation de l'influence des pratiques culturales sur le contournement de *Rlm7* associé à l'étude des souches de *L. maculans* avirulentes et virulentes vis-à-vis de *Rlm7* issues des sites expérimentaux de Versailles et de Grignon sur trois ans sont ici présentés sous la forme d'un article intitulé « Experimental evolution of the *AvrLm4-7* gene of *Leptosphaeria maculans* under *Rlm7* resistance gene selection pressure in the field » devant être soumis au journal « PLOS pathogen ».

# **Experimental Evolution of the *AvrLm4-7* Gene of *Leptosphaeria maculans* Under *Rlm7* Resistance Gene Selection Pressure in the Field**

**Guillaume Daverdin<sup>1</sup> , Thierry Rouxel <sup>1</sup>, Lilian Gout<sup>1,3</sup>, Jean-Noël Aubertot<sup>2</sup>, Isabelle Fudal<sup>1</sup>, Julien Carpezat<sup>4</sup> & Marie-Hélène Balesdent <sup>1\*</sup>**

**1** INRA, UR1290 Bioger, Av. Lucien Brétignières, Campus AgroParisTech, BP01, 78850 Thiverval-Grignon, France

**2** INRA, UMR 1248 AGIR, BP 52627, 31326 Castanet Tolosan Cedex, France

**3** AgroParisTech, Av. Lucien Brétignières, Campus AgroParisTech, BP01, 78850 Thiverval-Grignon, France

**4** CETIOM, Av. Lucien Brétignières, Campus AgroParisTech, BP01, 78850 Thiverval-Grignon, France

\* corresponding author, e-mail [mhb@versailles.inra.fr](mailto:mhb@versailles.inra.fr)

## Abstract

*Leptosphaeria maculans*, the agent of the stem canker of oilseed rape, develops “gene-for-gene” interactions with its host plant, and shows a high evolutionary potential to “breakdown” novel resistance sources in the field. Here, we set up an experimental evolution trial to monitor the evolution of populations faced to the newly marketed *Rlm7* resistance over three consecutive years, and we identified molecular events responsible for the loss of the corresponding *AvrLm7* avirulence. The experiment was done in two fields in which either minimum tillage or regular cultural practices were used. Population genetics approaches were also used to evaluate the main population traits linked with birth and dissemination of the mutations in the populations. A total of 1530 isolates was collected over the three years in a single field in which minimum tillage was applied, on one cv. harbouring *Rlm7* and one cv. devoid of it. During this period *avrLm7* virulent isolates collected on the susceptible cv. increased in frequency from 10% to nearly 40%, but remained undetectable in neighbouring fields or in the experimental field in which regular cultural practises were used. All possible types of molecular events leading to virulence were found, with an over-representation of complete deletions and wide-scale Repeat Induced Point mutations of the coding sequence. Our study points out the importance of mutability of genes involved in the interaction, directly linked with their genome environment, and reproduction/mating system rather than geographical dissemination, which appear to be a minor threat for the resistance breakdown in the *L. maculans*-*B. napus* pathosystem. By favouring sexual reproduction, minimum tillage practices are shown here to strongly enhance the mutability of avirulence genes and the overall evolutionary potential of the fungus.

## Author Summary

Plant genetic resistance is the most efficient and environment-friendly strategy to combat plant diseases due to fungi, but fungi show a high evolutionary potential allowing them to overcome novel resistance when these are cropped on large acreages. Here we set up an experimental evolutionary field over three consecutive years to evaluate how and how fast a pathogen of oilseed rape, *Leptosphaeria maculans* will overcome the novel resistance gene *Rlm7* and we compared 'cautious' and 'risky' (more and more used by farmers to-date) cultural practices to evaluate their incidence on disease resistance breakdown. When using the 'risky' practices, but not the 'cautious' ones, the fungus showed a very high evolutionary potential and the populations contained ca. 40 % of virulent isolates at the end of the study. All possible molecular events leading to virulence were identified at the single field scale, and our study strongly suggests that the genome environment and the reproduction regime privileged hypermutability at a single locus to gain virulence. Migration of virulent isolates from surrounding fields is thus less threatening for resistance breakdown than the intrinsic ability of population of the fungus to generate gene hypermutability.

## Introduction

Fungi are the most important pathogens of cultivated plants, with a significant impact in economic, human health and environmental terms, the later being due to the high amount of fungicides used to control plant diseases (Stukenbrock & McDonald, 2008). In contrast to fungicide use, plant genetic resistance to diseases is an environment-friendly strategy. One common scheme allowing effective resistance, designed according to the biological properties of the pathogens and hosts, is the incorporation of major-gene resistance (*R*) genes, into a plant genotype (Hammond-Kosack & Kanyuka, 2007). Unfortunately, fungi show an incredible plasticity and ability to disseminate and adapt rapidly to changes in their environment, including host species or control methods (Stukenbrock & McDonald, 2008). In this respect, emergence of novel, better-adapted strains resulting in “breakdown” of control methods in a few years only, are common traits of fungal phytopathogens (McDonald & Linde, 2002).

Durable disease resistance is a long-sought goal of crop protection and plant breeding programs, and a fertile field for research, since durable resistance is suggested to be cost effective and environmentally sound (Leach et al., 2001). In addition it promotes the conservation of genetic resources that may be rare due to the narrow genetic basis of some “recent” crop species, as is the case for oilseed rape (*Brassica napus*) (Morice, 1963). The durability of a resistance gene (van den Bosch & Gilligan, 2003; Palloix *et al.*, 2009) is dependent on three main factors : (i) the genetic background of the plant genotype, with a recent example showing that the introduction of a *R* gene in a variety expressing a high level of quantitative (polygenic) resistance increases durability of a *B. napus* (oilseed rape) major gene resistance (Brun et al. 2010). (ii) The biology and evolutionary potential of the pathogen population. Pathogens with a high evolutionary potential are more likely to overcome genetic resistance than pathogens with a low evolutionary potential, and the differences in

evolutionary potential are assumed to depend upon population size, mutability of genes involved in the interaction, reproduction/mating system and gene/genotype flow. According to McDonald and Linde (2002), these last two criteria are suggested to be crucial for resistance “breakdown”, with the most critical pathogens being those that possess both a high potential for gene flow and a mixed reproduction system, with at least one sexual cycle per growing season and asexual reproduction during the epidemic phase. (iii) Cropping practices at the field level including crop sequence, and stubble management through tillage that theoretically may directly affect the evolutionary potential of the fungus by reducing the population size, its dissemination ability or interfere with its reproductive regime (Aubertot et al., 2006). At the landscape level, the proportion of the area sown with a resistant variety throughout the years also influences the durability of specific resistances (Lô-Pelzer et al., 2010a; 2010b).

Stem canker of crucifers, also termed blackleg of oilseed rape (OSR), is caused by the ascomycete *Leptosphaeria maculans* (*Phoma lingam*). It is a disease of major economic importance whose magnitude increased dramatically with the development of oilseed rape cropping in all parts of the world (Fitt et al., 2006). Following sexual reproduction taking place on stem debris, leaf infections by ascospores (in autumn in Western Europe) cause leaf lesions, supporting asexual multiplication, whose incidence on yield is very limited. The life/pathogenic cycle of the fungus is completed by a lengthy endophytic systemic colonisation of plant tissues departing from the leaf lesions, eventually resulting in lodging of the plants at the end of the growing season (Spring and early Summer in Western Europe) (Rouxel & Balesdent, 2005).

Resistance to *L. maculans* may be major single-gene resistance expressed since the seedling stage, or polygenic “field” resistance expressed after initial infection, which will eventually reduce disease severity (Delourme et al., 2006). Major gene resistance is race-specific: a *L. maculans* avirulence (*AvrLm*) gene induces resistance when it challenges the

corresponding plant resistance (*Rlm*) gene. Race-specific resistance is effective at protecting plants in crops if the corresponding *Avr* allele is predominant in the local pathogen population (Ansan-Melayah et al., 1997). Major gene resistance has been widely used in the past in OSR (Rouxel et al., 2003a; 2003b). However, surveys of populations, at a time during which the *Rlm1* resistance was used on a large scale in farmer fields, demonstrated that virulent isolates able to attack *Rlm1* plants increased in frequency from less than 10% to more than 90% in only 3 growing seasons, leading to a rapid loss of field efficiency of *Rlm1* (Rouxel et al., 2003b). Additional data suggest a highly dynamic evolution of race structure of *L. maculans* populations that may have sequentially overwhelmed *Rlm2*, *Rlm4* and *Rlm1* in Europe (Rouxel et al., 2003a).

The impact of cropping practices, including crop sequence, return delay, stubble management and tillage, sowing date, or nitrogen availability during the vegetative stages, on stem canker severity has been investigated (Aubertot et al, 2004, 2006). The impact of cropping practices on the durability of R genes is also postulated but has never been experimentally assessed, and there is currently no explanation on how they may affect the speed and pattern of molecular evolution of *Avr* genes submitted to R gene selection pressure.

The cloning and characterization of the molecular determinants of resistance from the plants (*R* genes) and pathogens (*Avr* genes) has not only led to a deeper understanding of their functions and how they interact, but also provided the molecular information and means to detect mutations in these genes. *Avr* genes have been cloned from only seven fungal species, and the manner in which fungal *Avr* genes evolve to evade host recognition is only documented for few *Avr* genes, mostly in *Cladosporium fulvum*, *Magnaporthe oryzae*, *Rhynchosporium secalis* and *L. maculans* (Stergiopoulos & de Wit, 2009; Rouxel and Balesdent, 2010). Moreover the link between laboratory and field studies often is missing,

and when field populations were investigated for mutations in *Avr* gene they often originated from uncharacterized plant genotypes or were collected and analysed well after the corresponding resistance gene was used in the agronomic practice (e.g. Schürch et al., 2004). There is currently no example of identification of the initial events leading to virulence when a fungal pathogen is faced for the first time to a *R* gene deployed under field conditions.

Three *Avr* genes, *AvrLm1*, *AvrLm6* and *AvrLm4-7* have been cloned in *L. maculans* (Gout et al., 2006a; Fudal et al., 2007; Parlange et al., 2009), allowing studies on their molecular evolution under selection pressure to be undertaken. Virulent populations at the *AvrLm1* locus quasi-exclusively acquired virulence by large-sized (>200 kb) deletions, but they were sampled a few years after the *Rlm1* resistance was deployed (Gout et al., 2007).

The *AvrLm4-7* gene is recognised by two distinct resistance genes, *Rlm4* and *Rlm7* (Parlange et al., 2009). The *L. maculans* populations have been submitted to the *Rlm4* selection pressure since the 1970's in Europe, and escape from recognition by *Rlm4* is due to a single-base non-synonymous mutation, maintaining recognition by *Rlm7* (Parlange et al., 2009). In contrast, until recently, *L. maculans* has never been subjected to the *Rlm7* selection pressure in Europe and a population survey done in 2000-2001 in France identified only one virulent isolate of 1787 (0.05 %) (Balesdent et al., 2006). The *Rlm7* resistance has been introduced in commercial cultivars in 2002 in France, and became a commercial success in 2005, thus providing us with the opportunity to survey emergence of virulent *L. maculans* isolates at the time of initial selection pressure and to identify the first molecular events responsible for the overwhelming of the resistance gene. This context was also an unprecedented opportunity to set up an experimental evolution trial aiming at comparing contrasting cropping practices to assess their effect on the rate of adaptive response of *L.*



*maculans* populations to plant resistance changes and identify the molecular events favoured by “risky” practices towards the disease.

The aim of this study was to combine agronomical, molecular genetic and population genetic approaches to identify molecular events responsible for the loss of the *AvrLm7* specificity in *L. maculans* confronted to the *Rlm7* selection pressure. A three-year field experiment was set up to evaluate the impact of two cropping systems on the evolution of *L. maculans* populations and in particular on the selection of new virulent isolates. Molecular analysis of events leading to the virulent phenotype in a single field revealed a tremendous diversity of mutation events and mechanisms. In accordance with the large impact of cultural practices on fungal biology, we demonstrate here the role of population size and gene mutability, driven by sexual reproduction and genome structure, rather than migration on adaptation of populations to a new R gene and R gene durability.

## **Results:**

### **Incidence of the disease in the field**

Two experimental plots were set up in Grignon and Versailles, ca. 12 km away one from another. Both were cropped alongside with a cultivar harbouring the novel *Rlm7* resistance (mostly cv. Exagone) and a cultivar devoid of it (Campala). Crop management plans were identical for both locations in terms of sowing date, sowing rate, nitrogen management, herbicide and insecticide applications, and harvesting date. However, the two sites differed in terms of tillage and crop sequence. At Versailles, the plots were systematically ploughed after harvest and had not contained OSR for at least 4 years. In Grignon, the plot was a OSR monoculture associated to simplified tillage. The two considered cropping systems aimed at mimicking an area with cautious cultural practices

with regards to phoma stem canker (Versailles) and one with risky cultural practices (Grignon; Schneider et al., 2006).

The severity of stem necrosis was evaluated at the end of the growing season according to the G2 disease index ranging between 0 and 9 (Aubertot et al., 2004). In both Summer 2007 and 2008, the disease incidence was very low on the cultivar devoid of *Rlm7*, Campala : 0.96 and 1.05 at Versailles in Summer 2007 and 2008, respectively, 1.53 and 1.03 at Grignon in Summer 2007 and 2008, respectively. This indicates a very low disease pressure in both these years. Even with this low disease pressure, the efficiency of the *Rlm7* resistance expressed in cv. Exagone was obvious at Versailles, with G2 disease index of 0.17 and 0.38 in 2007 and 2008, respectively. Such low G2 index can be interpreted as complete immunity of the plants to the disease. In Grignon, however, the G2 disease index on Exagone was higher than in Versailles (0.86 in 2007) and, in 2008, also higher than the one observed on Campala (2.34).

### **Interaction phenotypes**

A collection of 2621 isolates has been collected over three consecutive years either from the *Rlm7* cv. Exagone or from the non-screening genotype Campala, and from two locations where contrasting agronomical practices were used (Table 1). Of these, 2047 were characterized for their interaction with *Rlm4* and *Rlm7* plant genotypes since *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7* resistance genes.

Depending on the sampling year, the frequency of the *AvrLm4* (i.e., avirulent towards *Rlm4*) isolates ranged between 1.4 and 4.3% in the Grignon sampling, and between 1.6 and 1.9% in the Versailles sampling. In both cases, this frequency is strongly reduced as compared to what was previously found in these locations in 2001 (11.0% for Grignon and 19.6% in Versailles; Balesdent et al., 2006), but consistent with the reduction in the

frequency of *AvrLm4* isolates observed in the surroundings of these experimental plots between 2000 and 2006 (Table 2).

The frequency of *avrLm7* (i.e., virulent towards *Rlm7*) isolates was estimated prior to the set up of the experimental field following samplings on a series of cvs devoid of *Rlm7* sown in the immediate vicinity of the experimental plots, between 2000 and 2004 (Table 2). Of 822 isolates analysed, only one *avrLm7* was found at Versailles in 2004 (Table 2).

Within the experimental evolution trial, only a few *avrLm7* isolates were found on Campala at Versailles (0 in 2006, and 1 [0.6%] in 2007). In contrast, a much higher proportion of *avrLm7* isolates were collected from Campala at Grignon in 2006, and this frequency increased over the three sampling years to reach 36.2 % in the 2008 sampling (Figs. 1; 2). The frequency of virulent isolates was also found to increase as a function of the sampling date in Autumn, as shown when analyzing the five sampling dates of leaves of Campala at the Autumn 2006 (between October 10 and December 7) (Fig. 2). This increased ratio was consistent with that found in stems of the same cv. at the end of the 2006-2007 growing season (Fig. 2; Table 1)

At Grignon, the frequency of *avrLm7* isolates was higher and increased faster when isolated from Exagone than when obtained from Campala (Fig. 1). However, presence of *Rlm7* in the plants did not fully prevent the recovery of avirulent isolates, whether obtained from leaf lesions or from stems, and, in this later case, postulated to have successfully completed the whole endophytic systemic colonization of the plant (Table 1). This may be due to a sampling bias since even non-typical lesions, with sparse pycnidia and surrounded by a dark margin (Fig. 3) were collected the first year, so that we maximize the number of virulent isolates recovered on leaves at a time when they were infrequent in the pathogen population. Often, such atypical leaf lesions corresponded to avirulent isolates (Fig. 3).

Analysis of the combination of interaction phenotypes, i.e, occurrence of the different alleles of the *AvrLm4-7* gene indicated absence of the *AvrLm4-avrLm7* phenotype, and scarcity of the double avirulent phenotype (*AvrLm4-AvrLm7*), as expected due to the scarcity of isolates expressing the *AvrLm4* avirulence (Table 1). The increase of the double virulent phenotype (*avrLm4-avrLm7*) at Grignon reflected the observed increase of the *avrLm7* isolate frequency (Table 1; Fig. 1).

This three-year experiment allowed us to generate a collection comprised of 774 *AvrLm7* isolates and 802 *avrLm7* isolates from a single field (Grignon), for subsequent molecular analyses.

### ***AvrLm4-7* polymorphism in avirulent population**

High Resolution Melting (HRM) analysis of *AvrLm4-7* alleles of isolates expressing the *AvrLm7* avirulence was done on 169 randomly selected isolates of the Grignon sampling. It was completed by sequencing of a subset of each of the classes revealed by HRM analysis.

This revealed a very low level of nucleotide polymorphism of the gene in the coding sequence and the intron, consistent with what was observed previously (Parlange et al., 2009). Only five polymorphic sites were found, all corresponding to non-synonymous mutations in the protein (Table 3). The combination of these polymorphic nucleotides generated five haplotypes (Table 3). One haplotype grouped all the isolates showing the *AvrLm4* specificity and differs from the other haplotypes by the presence of a guanine at base 358 resulting in a Glycine at amino acid 120 (Table 3). In accordance to what was described by Parlange *et al.* (2009) all isolates having lost the *AvrLm4* specificity showed a G358C mutation, resulting in a Gly120Arg change (Table 3; Fig. 4). Of the four alleles having lost the *AvrLm4* specificity but maintaining the *AvrLm7* avirulence, two were largely prevalent, and represented 87.6% of the isolates analysed (Table 3). Each of these four

alleles exhibited only one point mutation (in addition to the G358C mutation) as compared to the allele representative of the double avirulent phenotype (Table 3).

### **Molecular evolution of the *AvrLm4-7* gene: deletion events and major genomic rearrangements in *avrLm7* isolates**

Of 802 *avrLm7* isolates obtained from Grignon, 751 were characterized for the *AvrLm4-7* sequence or expression alteration responsible for the *avrLm7* phenotype. Of these, 274 isolates (36.5%) showed an amplification product of the expected size when amplified with both external and internal primers (Fig. 5). In addition, one isolate, G08-C1022, showed an internal amplification product of the expected size whereas the PCR product obtained with external primers was 270-bp larger than expected.

Of the remaining 476 isolates, 348 isolates (46.3%) did not produce an amplification product when using external or internal primers and 128 isolates (17.0%) produced an amplification product of the expected size with internal primers but no amplification with external primers. The most common event, corresponding to a complete absence of amplification was interpreted as complete deletion of the gene as commonly observed for other avirulence genes of *L. maculans* (Gout *et al.*, 2007; Fudal *et al.*, 2009) or for *AvrLm4-7* in isolate NZ-T4 (Parlange *et al.*, 2009). In one other case, that of isolate G07-E1059, a putative small-size, 28-kb deletion could be postulated. TAIL-PCR amplification departing from the amplified internal *AvrLm4-7* fragment showed that the gene was interrupted within the intron and that the following sequence corresponded to the closest GC-equilibrated region, 28 kb away. This could indicate deletion of the 3' part of the gene (including STOP codon and 3' UTR) along with the following ca. 28 kb mosaics of degenerated Transposable Elements (TEs) and/or major rearrangement in this genomic region as compared to the v23.1.3 reference whole-genome sequence (Rouxel *et al.*, 2011).

The case of the other isolates for which PCR amplification of the *AvrLm4-7* gene could be done with internal, but not external primers was uneasy to resolve due to the low efficiency of the TAIL-PCR recovery of flanking sequences. TAIL-PCR was performed onto 30 isolates in which only the internal amplification was feasible, but only four allowed us to recover an exploitable sequence data, a ratio lower than the 33%, previously described in *L. maculans* for recovery of the T-DNA insertion sites (Blaise et al, 2007). Apart from the previously described case of isolate G07-E1059, all three sequences corresponded to *AvrLm4-7* sequences truncated in their 3' part and in which the interrupting sequence was a portion of a TE inactivated by Repeat Induced Point (RIP) mutations (Attard *et al.*, 2005; Rouxel *et al.*, 2011). These corresponded to internal parts of the Long Terminal Repeat (LTR)-retrotransposon RLG\_*Rolly* or of the DNA transposon DTM\_*Sahana* (for classification and analysis of TE families in *L. maculans*, see Rouxel *et al.*, 2011). In the third case, the interrupting sequence corresponded to one LTR of RLC\_*Pholy*, but, due to sequence length, it was not possible to conclude whether the whole TE, or only one solo-LTR was interrupting the *AvrLm4-7* sequence. This, however could be done in the case of isolate G08-C1022, whose external PCR product was larger than expected, and for which the whole *AvrLm4-7* sequence could be obtained. In this case, only the 281-bp LTR of RLC-*Pholy* was found to be inserted within the gene. Examination of the whole genome sequence of isolate v23.1.3 revealed that RLC-*Pholy* is present in the proximity of the *AvrLm4-7* sequence, 7 to 12 kb away, both in 5' and 3', suggesting that minor deletion may have generated interruption of the gene by the TE, but not directly integration of the solo-LTR within the coding sequence. The interruption by the two other TEs is even more difficult to resolve as none is present in the vicinity of *AvrLm4-7*. The closest DTM\_*Sahana* found is located 220 kb away on the 3' side, and this TE has the particularity to be mostly targeted to GC-equilibrated isochores and not to AT-rich regions where *AvrLm4-7* resides (Rouxel et al., 2011); the closest RLG\_*Rolly*

found is located more than 1.1 Mb away on the 3' side of the gene, but 22 kb away on its 5' side. In the absence of TE having maintained their transposition activity in the genome (Rouxel *et al.*, 2011), all these data suggest that major genome rearrangements were at the origin of *AvrLm4-7* inactivation in alleles for which only the internal fragment of the gene could be amplified.

### **Molecular evolution of the *AvrLm4-7* gene : point mutations and other minor events**

The *AvrLm4-7* allele was sequenced for the 274 isolates for which the whole sequence could be amplified using the external primers and showed a great diversity of mutational events responsible for the *avrLm7* virulence, including point or dinucleotide deletions, point mutations non attributable to RIP and RIP mutations (Fig. 4).

Single nucleotide deletion was a rare event and regarded only two sites in four isolates. The deletions of the A<sup>79</sup> nucleotide (one isolate) and the G<sup>200</sup> nucleotide (three isolates) both introduced a frameshift likely to result in production of a severely truncated protein of 23 and 46 amino acids, respectively (Fig. 4).

A much more common event, found in 55 isolates (7.3% of the *avrLm7* isolates) was a dinucleotide deletion that occurred at two locations within the coding sequence : AA<sup>78-79</sup> (53 isolates), taking place close to a GAGA motif, and AA<sup>149-150</sup> (2 isolates). These two dinucleotide deletion events have also been found in two French isolates of the world-wide *avrLm7* collection of isolates (data not shown). To our knowledge, such an event has never been described for inactivation of avirulence genes. These deletions also introduced a frameshift likely to generate severely truncated proteins of 20 and 34 amino acids, respectively (Fig. 4).

In addition to previously identified Single Polymorphic Nucleotides (SNP) found in *AvrLm4-7* that do not affect the *AvrLm7* interaction specificity, seven polymorphic sites were

found in 30 isolates to be responsible for the *avrLm7* phenotype (Fig. 4; Table 4). Only SNPs in the coding sequence were found (e.g. none was present in the intron) and none regarded the signal peptide sequence. As compared to the reference sequences (v23.1.3 or v23.1.2) never more than one polymorphic site was found per isolate. Finally all mutations were either non-sense or missense mutations (Fig. 4; Table 4). The two most commonly found mutations regarded a C103T mutation changing the Gln<sup>35</sup> residue into a stop codon, and the G299C change resulting into Arg100Pro change (Table 4). The C103T mutation has features reminiscent of RIP mutations and C<sup>103</sup> is often one of the mutated nucleotide among others in more heavily RIPped alleles (see below) and probably represents an extremely mild RIP event. Two SNP resulted in changes of Cys<sup>22</sup> and Cys<sup>79</sup> residues into another amino acid (Table 4). Another mutation affecting another cysteine (Cys90Phe) was found in isolate S06 from the world-wide collection of isolates (data not shown). In all three cases, modification of a cysteine residue involved in disulphide bridge formation (I. Fudal and H. van Tilbeurgh, unpublished data) is likely to have drastic effects on the 3-D structure of the avirulence protein. Similarly the SNP resulting in the Leu<sup>55</sup> to STOP change (Table 4) is likely to result in inactivation of the protein via production of a severely truncated protein. The other non-synonymous mutations (G299C, T305C and C336G) drastically change the physico-chemical properties of the residues (Arg100Pro, Phe102Ser and Ser112Arg, respectively) (Table 4).

### **Molecular evolution of the *AvrLm4-7* gene: RIP mutations**

In addition to the seven isolates with a single C<sup>103</sup> to T<sup>103</sup> mutation probably attributable to RIP, 179 isolates (23.8%) showed 3 to 31 point mutations corresponding to 84 polymorphic sites throughout the *AvrLm4-7* sequence. Alignment of 127 fully sequenced alleles showed these mutations were exclusively C  $\leftrightarrow$  T or G  $\leftrightarrow$  A mutations, with one exception where the polymorphic nucleotide was a G, an A, or a C (Fig. 5). This resulted in a



reduced GC content of the alleles (mean: 40.45% instead of 43.93% found in the reference v23.1.3 allele). The mutations were not random within sequences, and were preferentially affecting CpA or TpG di-nucleotides (CpA → TpA and TpG → TpA) (60.7% of the cases). Other dinucleotide affected comprised CpG, CpC and GpG, but more rarely than CpA or TpG (Fig. 5). These mutations thus typically took place in the favoured dinucleotide context for RIP occurrence (Hane and Oliver 2008). Confirmation of the occurrence of RIP in these alleles was obtained with a RIPcal analysis (Hane and Oliver 2008) and by calculating the typical RIP indices TpA/ApT and (CpA+TpG)/(ApC+GpT) (Galagan et al., 2003; data not shown). RIPcal analysis also pointed out that the whole gene sequence was severely affected by RIP, except for the intron and part of the sequence coding for the peptide signal (Fig. 5). In contrast the region located immediately in 5' of the intron showed a very massive RIP degeneracy (Fig. 5). This however does not indicate any RIP target preference in the gene since examination of the RIP targets (i.e., CpA or TpG dinucleotides) revealed that 51 of 56 (91.1%) of favored RIP targets were mutated in one allele or the other. The scarcity of RIP degeneracy in the part of the gene encoding the signal peptide or the intron was uniquely due to the scarcity of RIP targets in these parts of the sequence (e.g., only one CpA dinucleotide in the intron). Sequencing of ca. 550 bp of the promoter region and 60 bp of the 3' UTR also showed the whole 5' UTR and 3' UTR were severely affected by RIP (Fig. 5). In contrast the sequenced part of the promoter downstream of the 5' UTR was only weakly affected by RIP mutation (Fig. 5) even though the density of RIP target is similar to that found in other parts of the gene (data not shown). Within the coding sequence, and if only considering the typical CpA → TpA and TpG → TpA mutations, the number of mutations caused by RIP ranged between 3 and 29, depending on the allele (Fig. 4).

As generally described for RIP mutations, G → A or C → T transitions introduced STOP codons in the sequence, likely to result into a severely truncated protein, that may only contain 13 amino acids in a series of alleles (Fig. 4).

The RIP-generated polymorphic sites resulted in high allele diversity with 98 different alleles out of 127 sequences. Only seven alleles were found to be shared by more than one isolate (2 to 4, if excluding isolates obtained from ascospores on the same stubble and likely to result from the same meiosis), whereas all other ones were found in only one isolate each.

### **Molecular evolution of the *AvrLm4-7* gene : unexplained phenomena**

In the case of six isolates from Grignon, the sequencing of the *AvrLm4-7* gene encompassing 550-bp of promoter region and 60 bp of 3'UTR sequencing did not reveal any nucleotide polymorphism as compared to reference *AvrLm7* alleles (data not shown). It was thus investigated whether *in planta* altered expression could be responsible for the *avrLm7* phenotype. If including isolates of the international collection of *avrLm7* isolates, only three with an unaltered *AvrLm4-7* sequence showed a markedly reduced level of expression *in planta* (Fig. 6). This was found in three isolates from Western Australia, but not in two of the isolates obtained from the Grignon experimental evolution plot (Fig. 6 and data not shown). Other alleles showing sequence alteration and conferring the *avrLm7* phenotype were also not found to be altered in their level of expression *in planta* 12 dpi (Fig. 6).

### **Dynamics of mutation events during the three-year survey**

Investigation of the genetic variability for the *AvrLm4-7* gene in the 751 virulent isolates from Grignon revealed numerous independent molecular events leading to alteration of this gene. The three year survey showed that major genome rearrangements leading to complete or partial gene deletion and RIP mutations were the two main events to explain

evolution of virulence toward *Rlm7*, being respectively present in 63% and 24% of the virulent isolates over the three years. However, our three-year survey indicated a contrasted sequence of events (Fig. 7). Minor events (SNP, single or di-nucleotide frequency, unaltered gene sequence), along with major rearrangements leaving internal part of the gene unaltered tend to remain steady over the three-year survey, even though a low but constant increase in frequency of AA dinucleotide can be postulated (Fig. 7). In contrast, a dramatic change in frequency was observed over the three year for two of the most common events, RIP mutations and large-scale deletions. Whereas RIP mutation was the most frequent event leading to virulence on the first year of the survey, at a time when only very few virulent isolates could be found (Table 1), complete gene deletion became widely prevalent in years 2 and 3 of the survey, while RIP mutations decreased the second year and then remained stable at ca. 20% of the events (Fig. 7). When analysing intensity of RIP mutations in the *AvrLm4-7* allele, there was no clear pattern regarding the intensity of mutation as a function of the year of isolation with for example, alleles with the least number of RIP mutations found in 2006 as in 2008 (Figs. 4, 8), or in contrast, the alleles with the higher number of RIP mutations found in 2006 (isolate G06-436 in Fig. 4). Moreover, even though an increase in the amount of RIP-mutated sites was observed from one year to the other (Fig. 8), this increase was not statistically significant (non parametric Kruskal-Wallis test;  $P = 0.371$ ).

### **Population analysis**

To investigate the origin of *avrLm7* isolates and how their rapid increase in frequency has impacted the genetic structure of *L. maculans* populations, a subset of 891 *avrLm7* and *AvrLm7* isolates collected in the Grignon experimental plots were genotyped using four minisatellites markers (Suppl. Table S1). Gene and genotypic diversity for the eight populations are summarized in Table 5. All markers were polymorphic and a total of 65

alleles were detected. The average number of alleles over these loci ranged from 4.80 for the CaStem07 sample to 12.50 for the ExStem08 sample (Table 5). The small number of alleles observed in the CaStem07 and ExLeaf06 samples may be due to sample size. Indeed all samples shared the same high frequency alleles at the 4 minisatellites but the number of rare alleles (frequencies  $< 0.05$ ) was associated with the number of collected isolates. If only the more common alleles are considered (frequency  $> 0.05$ ), the average number of alleles is not significantly different across all samples (Table 5). The number of private alleles was estimated to 4 and 8 in the samples collected from cvs. Campala (CaLeaf06, CaLeaf07, CaStem07 and CaStem08) and Exagone (ExLeaf06, ExLeaf07, ExStem07 and ExStem08), respectively. Despite a higher number of private alleles in Exagone samples, private allelic richness, when corrected for sample size, was not significantly different between samples from the two cultivars (data not shown).

Over the 3 years of collection, samples collected from Exagone and Campala had similar level of average gene diversity (mean across samples  $H = 0.74$ ). The level of genotypic diversity was also very high in all samples, with G values close to 1.00, and was not significantly different across samples. Overall, a total of 492 multilocus genotypes were differentiated using the 4 minisatellites, of which 158 were shared by 2 to 26 isolates. These repeated genotypes were distributed within and among the samples and were composed of a combination of the most frequent alleles for the four minisatellites. Furthermore, polymorphisms are often observed for the mating type and for the *AvrLm4-7* loci for isolates sharing a genotype. Repeated genotypes detected in the samples are then more indicative of lack of discrimination power of the four minisatellites rather than clonality in these *L. maculans* populations. Indeed, analysis per sample showed that the number of genotypes often was very close to the number of isolates. The mean numbers of isolates per genotype

ranged from 1.00 to 1.39, but the differences were not significant, and 591 isolates (66%) had unique genotypes (Table 5).

Both mating types were present within all samples and occurred in equal frequencies for most of them, except for the Exleaf07 sample collected from leaf lesions occurring on the *Rlm7* cv. Exagone, in which a significant deviation ( $P < 0.05$ ) from a 1:1 ratio was detected (Table 5). The reason why this biased frequency is observed is unclear and one may hypothesize a resampling bias. The multilocus linkage disequilibrium values ( $r_d$ ) obtained for the eight samples were close to zero and did not deviate significantly from the expectations under the null hypothesis of random mating in all samples ( $P > 0.05$ ) (Table 5). These results suggested that there was no genotypic disequilibrium in the studied samples and that recombination regularly generates new genotypes in the experimental plots from Grignon.

To estimate differentiation between the samples, we firstly used Principal Coordinate Analysis (PCoA) to summarize information contained in a matrix of genetic distance among samples (data not shown). The first three Principal Coordinates (PCo) captured 59.2%, 28.4% and 8.46% of total variation, respectively. PCoA results showed no clear pattern of samples differentiation related to the date of sampling or the cultivar sampled. PCo1 separated CaStem07 and ExLeaf06 from the 6 remaining samples and PCo2 mainly separated ExLeaf06 and CaLeaf07 from the other samples.

Analysis of  $F$ -statistics (Table 6) showed low but significant level of population differentiation for all loci and overall, with a mean  $F_{ST}$  of 0.004 ( $P = 0.002$ ) over loci and samples. However, no significant genetic differentiation was detected between the two groups of isolates collected from Campala or Exagone ( $F_{ST} = 0.001$ ,  $P = 0.92$ ). Pairwise levels of genetic differentiation also were estimated between all pairs of samples collected from these two cultivars. Most of these  $F_{ST}$ -estimates were not significantly different from zero or, when significant, ranged only from 0.002 to 0.013, with the highest estimates ( $F_{ST} > 0.010$ )

corresponding to pairwise comparisons between samples collected from the Campala and Exagone cultivars (Table 6). Similar results were obtained by testing for heterogeneity in allele frequencies for each pair of samples using the Fisher exact test (data not shown). These estimates are consistent with the hypothesis that there is no genetic differentiation between these field samples.

Genetic differentiation also was calculated between different combinations of isolates, using the avirulence or virulence towards *Rlm7* cultivars or the type of molecular event leading to the *avrLm7* virulence as a grouping factor (Table 7). There was no evidence of a significant differentiation between avirulent and virulent isolates at the *AvrLm7* locus ( $P > 0.05$ ). Pairwise comparisons indicated non-significant divergences between the group of avirulent isolates and most of the groups of virulent isolates, except for the group of isolates in which the *AvrLm4-7* gene was altered by SNPs ( $F_{ST} = 0.01$ ,  $P = 0.02$ ). Most of the pairwise comparisons between virulent isolates indicated no significant differentiation but minisatellite data indicated low but significant genetic structure between groups of isolates for which virulence *avrLm7* was either related to small scale deletion or to large scale deletions, RIP mutations and single nucleotide polymorphisms (Table 7). The mean  $F_{ST}$  over all pairwise comparisons was very low ( $F_{ST} = 0.004$ ,  $P = 0.01$ ).

## Discussion

Whereas “breakdown” of resistance has been described many times when a new resistance gene is used in the field (Stukenbrock & McDonald, 2008), the incidence of agronomical practices on the “evolutionary potential” of a pathogen and the impact they may have on the mutability and selection of genes involved in the plant-pathogen interaction has never been investigated at a molecular level. For this objective, we set up an experimental

evolution trial (two fields) over three continuous years, taking advantage of the commercial release in France of the novel *Rlm7* resistance in oilseed rape, while the French populations of the pathogen were mostly avirulent at the corresponding avirulence locus *AvrLm7* (Balesdent et al., 2006). The two trials were submitted to contrasted agronomical practises (“cautious” in Versailles, and “risky” in Grignon), and surrounding fields cropped with oilseed rape commercial cultivars devoid of the *Rlm7* gene allowed us to monitor the local populations. We analysed population evolution in terms of virulence and in terms of molecular events responsible for the loss of the *AvrLm7* phenotype. Population genetic analyses were used to investigate changes in population structure linked with the surge and dissemination of virulent isolates.

Population evolution towards virulence was extremely different between the two sites. While virulent isolates did not appear and disseminate in the Versailles population, a dramatic increase in the frequency of virulent isolates was observed over the three years in Grignon, as determined following sampling from the cultivar devoid of *Rlm7* cropped alongside with Exagone. In contrast, virulent isolates were not detected in surrounding fields cropped with cvs devoid of *Rlm7*.

One of the most surprising finding of this study is that we found all possible molecular events reported in the literature for loss of fungal avirulence, in a single field. These included complete or partial deletion (Farman et al., 2002; Schürch et al., 2004; Stergiopoulos et al., 2007; Gout et al., 2007; Fudal et al., 2009; Rep et al., 2004; Zhou et al., 2007), amino acid substitutions (Stergiopoulos et al., 2007; Ellis et al., 2007; Schürch et al., 2004), point deletions and production of truncated proteins (Stergiopoulos et al., 2007), “insertion” of a transposon (Kang et al., 2001; Fudal et al., 2005; Zhou et al., 2007), RIP mutations (Fudal et al., 2009) and probable alteration of the expression level of the gene, this latter having been described previously in the case of a TE insertion in the promoter of AVR-Pita of *M. oryzae*

(Kang et al., 2001). In addition, two other phenomena previously unreported were observed, loss of avirulence without sequence alteration of the gene nor of its expression and the common deletion of an AA dinucleotide close to a GAGA motif. One possible explanation for the first case would lie with the presence of another gene masking the effect of *AvrLm7* as reported for *Avr1* masking the avirulence effects of *Avr2* and *Avr3* in *Fusarium oxysporum* (Houterman et al., 2008). Whereas most of the events of mutation or point deletion are infrequent within the Grignon populations, and their generation probably attributable to the normal mutation rate within the genome, the AA deletion genesis is unclear. This phenomenon has resemblance with the still understood mechanism by which DNA replication/RNA transcription and/or translation results in erroneous products with different numbers of A as compared to the original DNA sequences (Antonarakis and Cooper, 2010). Interestingly in cases associated with neurodegenerative diseases in human, dinucleotide deletion took place close to GAGAG motifs (van Leeuwen et al., 2000).

With *L. maculans* being a fungal species in which the sexual cycle is indispensable to pathogenicity (Rouxel & Balesdent, 2005), the driver of all most frequent mechanisms for *AvrLm7* inactivation, i.e., large scale deletions and RIP mutations, amounting to 87.1% of the virulent isolates, is meiotic recombination. *L. maculans* is a fungal species in which Chromosome Length Polymorphisms generated by meiosis were previously identified (Plummer and Howlett, 1995; Leclair et al., 1996). At this time, unequal crossovers arising from unequal pairing of chromosomes due to the presence of numerous repetitive sequences were postulated to lead to the production of chromosomes of novel sizes in the progeny (Plummer and Howlett, 1995). The recent sequencing of the genome of *L. maculans* substantiated this postulate by showing that one third of the genome is made of TEs of which the major bulk consists of only four widely expanded families of LTR retroelements sized 7 to 12 kb (Rouxel et al., 2011). In addition these are grouped in large AT-isochores (Rouxel et



al., 2011) thus providing multiple targets for mispairing between sister chromatids during meiosis. These AT-isochores contain all currently known avirulence genes of *L. maculans*, and large-size deletions are the main event leading to virulence at the *AvrLm1* and *AvrLm6* loci (Gout et al., 2007; Fudal et al., 2009) as found here for the *AvrLm4-7* locus. Similarly, cases where the central part of the gene was conserved whereas external amplifications were not possible are also postulated here to be due to TE-driven genome rearrangements, likely to be due to unequal crossovers. Mainly, the case where TEs are inserted cannot be explained by a recent transposition activity of TEs unlike what is known to occur in *M. oryzae* (Fudal et al., 2005), since all TEs of the genome are severely mutated by RIP and no active copy was found in the sequenced isolate (Rouxel et al., 2011).

Inactivation of an avirulence gene by RIP mutations has been reported previously for *AvrLm6* (Fudal et al., 2009), and was also indicated to be acting on *AvrLm1* (Gout et al., 2007). To date it has never been described to be involved in inactivation of avirulence genes in other fungal species. RIP has initially been described in the model fungus *Neurospora crassa* as a process that efficiently detects and mutates duplicated sequences. Acting only during the sexual cycle, RIP in *N. crassa* identifies duplications that are greater than ca. 400bp (or ca.1kb in the case of unlinked duplications) (Galagan & Selker, 2004). However, in *L. maculans*, the embedding of effector genes (including avirulence genes) in AT-rich isochores made up of mosaics of RIP-altered TEs, and the presence of RIP signatures in the sequence of effector genes indicated that RIP could act on unduplicated sequences to promote gene diversification and that it could “leak” from the neighbouring RIP-affected sequences to generate mutations in single-copy genes (Rouxel et al., 2011). This was further substantiated by the work of van de Wouw et al. (2010) showing a RIP gradient along affected effector gene sequences departing from the closest RIP-affected TEs, and by our own findings showing that RIP intensity along the *AvrLm4-7* alleles is higher in the 3' part of

the gene whose 3' UTR is directly bordered by TEs whereas its promoter is globally unaffected by RIP mutations (Fig. 5).

Ever since the first epidemics of stem canker, it was recommended to apply adequate cultural control methods to reduce the viability and dissemination of the pathogen. These included crop rotation to maximise separation between fields grown with OSR and minimize the probability that the inoculum maintained on stubbles from the previous years is disseminated to surrounding fields (Lacoste et al., 1969; Aubertot et al., 2006), burial and destruction of stubbles to reduce survival of the pathogen during its saprophytic stage and reduce the production of the inoculum following sexual reproduction (Alabouvette & Brunin, 1970). These later practices both favour the degradation of stem debris and repress sexual reproduction in the absence of light, as they prevent ascospore ejection in the air. Tillage regime thus has a direct incidence on stubble persistence and on inoculum availability (Schneider et al., 2006). Here we used two very different tillage regimes and noticed that minimum tillage not only favour the intensity of disease when the fields are sown with susceptible cvs as reported in the literature (Aubertot et al., 2006), but also dramatically improves the ability to breakdown a novel resistance source. The “risky” cropping practises used in the Grignon field are likely to have two major consequences: a large local increase in size of the local population due to the lack of rotation and the close contact of the crop with unburied infected residues from the previous years, and an increased rate of sexual reproduction due to the maintenance of infected debris on the soil. As underlined above the prevalent modes of loss of the avirulence at the AvrLm7 locus directly depend upon the ability the pathogen has to undergo sexual mating. Favouring this part of its life cycle along with population size directly impact the ability to generate a large number of virulent progeny and the chance they will be selected for by the resistant cvs, eventually improving their ability to mate between two virulent parent isolates.

Analysis of surrounding populations isolated from *rlm7* cvs. indicated an increase in *avrLm7* isolates at the borders of the Exagone experimental field, but no detectable virulent isolate in the surrounding fields located less than 600 meters away. This result is consistent with analyses of ascospore dispersal indicating most spores are deposited 500 meters away from the source, especially the first 100 meters (West et al., 2001; Marcroft et al., 2004) and, along with our results suggests that migration is of a low relevance to introduce a new virulence as compared to the intrinsic ability of the population to generate mutation at the selected loci. In addition, new virulent populations and sub-populations discriminated as a function of the molecular events leading to virulence were not accompanied by reduction in neutral diversity of populations, nor in new population structure. Moreover, when this could be investigated (i.e. in the case of RIP patterns of mutations) an extreme diversity of RIPped alleles was found, and alleles with a low intensity of RIP were found each year suggesting new RIP event occurred at each sexual stage. The lack of population differentiation and reduction of diversity strongly suggest that molecular events leading to virulence are generated *de novo* at each sexual cycle within the local population, and that the intensity of mutation favoured by the genomic environment of the avirulence gene and by the intense sexual multiplication of the fungus is much higher than the probability that a new virulent isolate migrates in surrounding population to establish a founding event, eventually disseminating within a population following selection. In this respect, the sequence of events observed in our experimental field can be summarized as follows: (i) multiple events of mutation favoured by the genomic environment of *AvrLm4-7* and driven by sexual reproduction appear at each sexual cycle and are selected for by the resistant cultivar; (ii) RIPped (or other inactivated) alleles eventually are replaced by large-size deletions since the maintenance of the otherwise dispensable AT-isochores hosting it is no longer selected. (iii) Depending on the fitness penalty linked with the loss of function of *AvrLm4-7*, in the near

future, the population submitted to the selection pressure will either mainly consist of isolates for which the whole genomic region is deleted (as observed for current European populations at the AvrLm1 locus; Gout *et al.*, 2007) or of populations in which one of the amino acid substitution event at location 100, 102 or 112 results in a still functional allele and will disseminate in the populations as observed for the loss of the AvrLm4 specificity (Parlange *et al.*, 2009; this study).

Lastly, our study provides new indicators and strategies for a durable management of major resistance genes. According to McDonald and Linde (2002), the evolutionary potential of a pathogen is a crucial point to predict the durability of a novel resistance. However, the main criteria the authors use to rank pathogens as highly prone to breakdown disease resistance are the mating system, the population size and the dispersion ability (McDonald and Linde, 2002). Our study clearly indicates that a fourth criteria has to be taken into account, the mutability of the corresponding avirulence gene, that is prevalent upon the ability to disseminate when the rate of mutation of a given gene is much higher than the probability that a new virulent isolate is propagated from other fields. The gene mutability may, as in the case of *L. maculans*, depend on the reproduction behaviour of the fungus, but it may also be extremely high for asexual fungi e.g. bearing avirulence genes on B chromosomes (Ma *et al.*, 2010). In addition, it will be a criteria to consider for given species in which avirulence genes are located in TE-rich environments such as *Phytophthora infestans* or *Blumeria graminis* (Haas *et al.*, 2009; Spanu *et al.*, 2010), but will also be a criteria to consider on a gene-by-gene basis for those fungal species for which only part of the avirulence genes are located in a dispensable genome environment, as exemplified for the few telomeric avirulence genes of *M. grisea* (Farman, 2007). This strongly indicates that, in addition to biological and epidemiological parameters pointed out by McDonald and Linde (2002), knowledge of the avirulence genes, their genome environment, and their involvement

in the fungal fitness are of utmost importance to predict how easily they can be lost and how long the corresponding resistance will last. This study also stress that long-time recommended cultural practises (burying through ploughing and rotations to favour the degradation of debris) influence mutability of genes and drastically limit the ability to generate the hypermutability at the avirulence locus it needs to breakdown disease resistance.

## **Materials and methods:**

### **Field experiments**

Isolates of *L. maculans* used in this study were collected from two independent experimental fields located in France (Grignon, 48° 50' 28.40" N latitude; 1° 56' 13.83" E longitude) and Versailles (48° 48' 27.59" N latitude; 2° 5' 12.30" E longitude), located ca. 12 km away one from another). These two locations have been used in previous studies to describe *L. maculans* population race structure in the wild (e.g., Ansan-Melayah et al., 1997; Balesdent et al., 2006). The plot in Grignon was a right triangle with a 50 m base and 100 m height. The plots in Versailles were 50 m x 100 m rectangles sown alongside with a cv. devoid of *Rlm7* (Campala) and a cv. harbouring *Rlm7* (Roxet in 2004 and 2005, and Exagone in 2006, 2007, and 2008). Crop management plans (defined by a set of decision rules following the CETIOM's regional recommendations) were identical for both locations in terms of sowing date, sowing rate, nitrogen management, herbicide and insecticide applications, and harvesting date. No fungicide was used in the experiment. However, the two sites differed in terms of tillage and crop sequence (see results section).

Data collected included severity of stem necrosis evaluated at crop maturity (GS 6.6, Sylvester-Bradley, 1985), at the summer of 2005, 2007 and 2008. The severity of stem necrosis was evaluated according to the G2 disease index (Aubertot et al., 2004) on 160 randomly chosen plants per cv. This index summarises the proportions of plants observed

within six canker severity classes. It is calculated using arbitrarily chosen coefficients associated to each class of canker severity and ranges between 0 (healthy plants) to 9 (all plants totally cankered).

### **Biological sampling in the experimental evolution trial**

During three continuous years corresponding to two cultural cycles (2006-2007; 2007-2008), but to three generations of the fungus, isolates from the Versailles and Grignon fields were sampled either from leaf lesions (single conidium isolates) or from stubbles (single-ascospore isolates) using methods described by Balesdent *et al.* (2006) and West *et al.* (2002), respectively. Due to the peculiar life cycle of *L. maculans* in which ascospores are the origin of leaf lesions in Autumn, which in turn initiate the systemic colonisation of plants eventually causing the stem canker in Summer, populations collected on stem canker at year *n*, following meiosis, were considered as similar to those collected from leaf lesions on the non-screening genotype (here Campala) at Autumn of year *n*. Isolates sampled from leaf or stem lesions from the Campala cv. were identified as CaLeafxx or CaStemxx, respectively, with xx standing for the year of collection and samples isolated from the *Rlm7* cv. Exagone were similarly identified as ExLeafxx or ExStemxx.

Depending on the presence of typical leaf lesions caused by *L. maculans* on the *Rlm7* genotype, number of leaves collected in Grignon and Versailles ranged between 20 and 200 (Table 1). For example, only 24 symptomatic leaves were found in Grignon on Exagone in Autumn 2006 and all of them were sampled. When more leaf lesions were present (subsequent years for Exagone, all years for Campala), infected leaves were randomly selected. In all cases, except the ExLeaf2006 sampling, only one isolate was obtained from each individual plant. As the frequency of virulent (*avrLm7*) isolates was expected to be low on both sites the first sampling year, 500 distinct leaf lesions were collected on Campala in

Autumn 2006 (Table 1), so that a frequency  $\geq 0.5\%$  of virulent isolates could be detected with a 95% confidence interval. According to the frequency of *avrLm7* isolates found in Grignon in 2006, this number was reduced to 200 leaves for the second year of the experiment (Autumn 2007). Similarly, 100 to 200 infected stubbles were collected each year, on each cv. and each site, except on Versailles in 2008 due to the lack of virulent isolates observed in the previous seasons and years (Table 1). Collected stubbles were washed with tap water and then maintained on the soil covered with a black plastic sheet under natural weather conditions for 3-4 months before pieces of stems bearing pseudothecia were used for ascospore isolation (West et al., 2002). Due to the scarcity of pseudothecia differentiated, more than one (average = 3) ejected ascospores were usually isolated from a single stem. Of these, only 3 isolates per stubble or less were analysed using molecular markers (see below).

### **Additional biological samplings**

The frequency of virulent *avrLm7* isolates before the set up of the experimental field trials (both sites) and in the close vicinity of the field trial (Grignon site) was estimated using collection of isolates collected prior or during the experiment, on a range of cvs. that do not contain the *Rlm7* gene (Table 2). Isolates were recovered as described above either from leaf lesions or from ascospores released from stem residues. In total, 822 additional isolates were thus phenotyped for avirulence on *Rlm7* and *Rlm4*, including around 100 isolates collected in Grignon in Autumn 2006, in two OSR fields located less than 600 m from the experimental trial (Table 3).

### **Additional isolates and maintenance of isolates**

The reference isolates v23.1.3 (*AvrLm4-AvrLm7*) whose genome sequence is available (Rouxel et al., 2011), v23.1.2 (*avrLm4-AvrLm7*), and NZ-T4 (*avrLm4-avrLm7*) (Balesdent et al. 2001; 2002) were used as controls for inoculation tests, and as sequence reference for the *AvrLm4-AvrLm7* and *avrLm4-AvrLm7* alleles of the *AvrLm4-7* gene (Parlange et al., 2009). In order to diversify the sources of molecular events leading to the *avrLm7* virulence, 47 *avrLm7* isolates from world-wide collections were also included. Of these, three French isolates, termed S06, L21 and B18 were obtained from leaf lesions of *Rlm7* cvs. in the Autumn of 2006 in Senne (cv. Exagone), Francillon (cv. Extend) and Le Blanc (cv. Exagone), and three isolates from the Western Australia collection of isolates, WT06, WT20 and WT37 (Gout et al., 2007) were included.

All fungal cultures were maintained on V8 juice agar and highly sporulating cultures were collected from 12-15 day-old cultures according to the procedure described by Ansan-Melayah et al. (1995).

### **Pathogenicity tests**

*AvrLm4* and *AvrLm7* avirulence/virulence phenotypes were determined by the inoculation of 15 days old *B. napus* cotyledons with 10  $\mu$ L of  $10^{-7}$  mL<sup>-1</sup> conidia suspension as described by Balesdent et al. (2005). Four different cultivars were used: the susceptible control Westar without any known resistance gene, Jet Neuf or Pixel (harbouring *Rlm4*) and line 01-23-2-1 (harbouring *Rlm7*). Symptoms were scored on 10 to 12 cotyledons per interaction two or three times at 14 to 21 days after inoculation by using the IMAScore rating scale, comprising 6 infections classes (IC1 to IC6), IC1 to IC3 representing the phenotypes of avirulent isolates (*AvrLm*) and IC4 to IC6 corresponding to virulent isolates (*avrLm*) (Balesdent et al. 2002). This analysis was also used to discriminate *L. maculans* from the closely-related species of the species complex, *L. biglobosa* (Mendes-Pereira et al., 2004).



These two species are commonly associated and can be isolated from the same infected organs (West et al., 2002). However their growth pattern on synthetic media and symptom expression on cotyledons is different (West et al., 2002) and allows to discriminate them.

### **DNA manipulations**

Genomic DNA was extracted from conidia suspensions using the DNeasy 96 Plant Kit and the QIAGEN BioRobot 3000 in accordance with the manufacturer's recommendations.

Polymerase Chain Reaction (PCR) primers used for mating type amplification, *AvrLm4-7* analyses and minisatellite analyses were designed with PRIMER 3 (Rozen and Skaletsky, 2000) and are described in Supplementary Table S1. In the case of the *AvrLm4-7* gene, PCR amplifications were done with (i) "external" primers spanning part of the promoter region and part of the 3' UTR and generating a 1434 bp fragment and (ii) "internal" primers located within the coding sequence, between the ATG and intron and spanning 478 bp (Fig. 5). PCR were performed in an Eppendorf Mastercycler EP Gradient thermocycler (Eppendorf, Le Pecq, France), with 30 cycles of 94°C for 30s, 60°C for 30s, 72°C for 80s, with a final extension at 72°C for 5min. Hybridization temperature and elongation time differed in function of the PCR products (Supplementary Table S1). Sequencing was performed using a Beckman Coulter CEQ 8000 automated sequencer (Beckman Coulter, Fullerton, CA, USA) according to the manufacturer's instructions.

### **TAIL PCR experiments**

To recover DNA sequences flanking *AvrLm4-7*, Thermal Asymmetric Interlaced PCR (TAIL-PCR-Liu et Whittier., 1995) was used. Nested *AvrLm4-7* sequence-specific primers were designed from the *AvrLm4-7* allele sequence of isolate v23.1.2 to match Liu and Whittier (1995) primers' characteristics (Supplementary Table S1). Arbitrary Degenerated (AD)

primers used in association with *AvrLm4-7* TAIL-PCR primers were identical to primers AD1, AD2 and AD3 described by Liu and Whittier (1995), but AD2 was more efficient to amplify the genomic region surrounding *AvrLm4-7* than the two other degenerated primers and was used for the two subsequent rounds of PCR amplification. First and second rounds of TAIL PCR were performed as described by Liu and Whittier (1995). Secondary TAIL-PCR products were purified using the Nucleospin® Extract II purification Kit (Macherey-Nagel, Hoerd, Fr) and were used either for the third round of TAIL PCR, whenever the amount of amplified product was insufficient, or as template for DNA sequencing using the specific tertiary border primer GD3 as a sequencing primer (Supplementary Table S1).

### **Quantitative RT-PCR**

Cotyledons of cv. Westar were inoculated as described above, and sampled 12 days after inoculation. Total RNA extraction and single-strand cDNA synthesis were performed as described by Fudal *et al.* (2007). In a first round, one RNA extraction and two reverse transcriptions were performed for each sample. Water and RNA from plants inoculated with isolate NZ-T4, lacking the *AvrLm4-7* gene, were used as a negative controls. Primers for *AvrLm4-7* amplification were as described by Parlange *et al.* (2009). qRT-PCR was performed using 7700 real-time PCR equipment (Applied Biosystems, Foster City, CA, USA) and ABsolute SYBR Green ROX dUTP Mix (ABgene, Courtaboeuf, France), as described by Fudal *et al.* (2007). *β-tubulin* was used as a constitutive reference gene. Ct values were analyzed as described by Muller and associates (2002). Whenever altered expression level was postulated, one or two additional biological samples were analysed.

### **High Resolution Melting PCR of *AvrLm4-7* allelic variants**

For high-throughput identification of *AvrLm4-7* allelic variants prior to allele sequencing (when relevant), High Resolution Melting PCR (HRM PCR) was used as an alternative to sequencing. A set of control isolates with known *AvrLm4-7* sequences was included in each HRM-PCR run and each sample differing from avirulent controls were sequenced. The HRM PCR mix used consisted of 5 $\mu$ L of genomic DNA (ca. 3 ng  $\mu$ L<sup>-1</sup>) added to 3  $\mu$ L of ultra-pure water, 1 $\mu$ L of each primer (at 6  $\mu$ M) and 10 $\mu$ L of MeltDoctor HRM Master Mix 2X (Applied Biosystems, Courtaboeuf, Fr). PCR was performed with the following conditions: an initial denaturation step at 95°C for 10 min, followed by 30 cycles of 95°C for 15s, 54°C for 20s, 72°C for 30s, and finally, an extension step at 72° C for 5 min. After the PCR was completed, high-resolution melting curve analysis was performed by collecting data from 70 to 90 C at a ramp rate of 1%. PCR and melting curves were done in a row on a qPCR 7500 Fast Real-Time PCR equipment (Applied Biosystems). Melting curves were analyzed with the High Resolution Melting software v2.0 (Applied Biosystems).

### **Markers for population genetics analyses**

A multiplex PCR was used to characterize the distribution of the Mat1-1 and Mat1-2 alleles within the collection isolates (Cozijnsen & Howlett, 2003). Four genetically independent minisatellite markers (Supplementary Table S1) and *MinLm9*, a minisatellite marker located 1.9 cM away from the *AvrLm4-7* locus (Parlange et al., 2009; Dilmaghani, 2010) were used for population genetic analyses. For each isolate, the allele sizes were determined using quantity one 1-D Analysis software (BioRad, Marnes-la-Coquette, France) by comparison with band sizes of the 1-kb<sup>+</sup> ladder (Invitrogen, Cergy Pontoise, France) and internal control with known allele size and known number of repeats of the core motif (i.e, the sequenced reference v23.1.3 isolate). Data were scored as the number of repeat units for each isolate and each minisatellite locus.

### **DATA analyses : sequence manipulation**

Sequences of the *AvrLm4-7* gene encompassing part of the promoter and part of the 3'UTR were compared following sequence alignment using MULTALIN and CLUSTALX (Corpet, 1988; Thompson et al., 1997). Automated analysis of RIP in *AvrLm4-7* alleles was performed using RIPCAL (<http://www.sourceforge.net/projects/ripcal>), a software tool that computes RIP indexes and performs alignment-based analyses (Hane et al., 2008).

### **DATA analyses : population genetics**

The frequency of each molecular event leading to the *avrLm7* virulence was calculated in the different samples. In cases where more than one isolate was recovered from a single plant, as it was the case for isolates obtained from stem residues, a sub-sample was obtained by random selection of a single isolate from each individual plant. The molecular event for each sampled isolate was then recorded and the frequency of each molecular event was calculated on the basis of this sub-sampling. This randomized sub-sampling was repeated 10000 times, and the mean frequency for each molecular event calculated.

To analyze minisatellites variability, the software GENEALex (Peakall and Smouse, 2006) was used to compute allele frequencies, number of alleles per minisatellites (A), number of private alleles, and Nei's gene diversity (Hd) (Nei, 1973) at each locus and over all loci within and over the samples (i.e., for each sampled OSR cv., year and period of sampling). The observed number of alleles in a sample is highly dependent on sample size, therefore allelic richness and private allelic richness was calculated using HP-Rare 1.1 (Kalinowski, 2005) correcting for sample size (n) in each sample using the rarefaction method, which fixes n as the smallest number of individuals typed for a locus in a sample. Genotypic diversity (G) was

calculated using  $G=[n/(n-1)(1- \sum p_i^2)]$  as implemented in MULTILOCUS (Agapow & Burt, 2001) where  $p_i$  is the frequency of  $i^{\text{th}}$  genotype and  $n$  is the number of individuals.

Linkage disequilibrium was evaluated using two different approaches. First, minisatellites were tested pairwise within and across samples using the genotypic disequilibrium test implemented in Genepop (Raymond & Rousset, 1995). The statistical significance of each pairwise test of linkage disequilibrium was tested by Fisher's exact test. For each of these tests, P-values were sequentially Bonferroni corrected for multiple tests (Rice, 1989). MULTILOCUS was also used to estimate the associations of alleles among different loci with the standardized version of the index of association  $r_d$ , using MULTILOCUS (Agapow and Burt 2001). Multilocus linkage disequilibrium (Brown *et al.*, 1980; Haubold *et al.*, 1998; Maynard-Smith *et al.*, 1993) was also estimated using the index of association, which gives information on whether two different individuals sharing the same allele at one locus are more likely to share the same allele at another locus. The significance of  $r_d$  was established by comparing the observed value to the distribution obtained from 1000 randomizations with alleles at each locus being resampled without replacement to simulate the effect of random mating. The hypothesis of random mating was tested as follows: the distribution of mating types was compared to the 50:50 ratio expected under random mating for a haploid fungus using  $\chi^2$  tests. Principal component analyses as implemented in Genalex (Peakall and Smouse 2006) were used to facilitate visualization in a multi-dimensional space of the pattern of genetic differentiation based on Nei's unbiased genetic distance (Nei 1973). Heterogeneity of allele frequencies among samples was investigated for each locus with contingency tables analysis and the Fisher exact test as implemented in the GENEPOP program. The null hypothesis tested was H0: "allele frequencies are homogeneous across populations." Genetic structure was analyzed with standard  $F_{ST}$  coefficient of population differentiation, which indicates the proportion of the total genetic variation attributable to population differentiation. F-statistics

were calculated using Genalex version 6.4 and  $F_{ST}$  values were tested for significance using 1000 permutations. To further analyze the effect of allelic variation of the *AvrLm4-7* gene on the genetic structure of *L. maculans* populations, analyses of genetic differentiation between samples were performed, using the type of molecular event leading to the *avrLm7* virulence as a grouping factor.

### **Acknowledgements**

The authors wish to thank Laurent Coudard (INRA-Bioger) for its major contribution to isolate collection management, Martin Willigsecker (INRA-Bioger) for plant management, Angélique Gauthier (INRA-Bioger) for help in sequencing, Jonathan Grandaubert (INRA-Bioger) for bioinformatic analyses, Martine Leflon (CETIOM, Thiverval-Grignon, France) for help in setting up the HRM protocol, Christophe Montagnier (Unité Expérimentale Grandes Cultures, INRA-Versailles) for setting up and supervision of the experimental evolution fields. Martin Barbetti (University of Western Australia, Crawley, Australia), Gunter Leis (Limagrain, France) and partners from “Chambre d’Agriculture de l’Indre”, contributed to the International collection of *avrLm7* isolates. This work was funded by the ANR project CEDRE (ANR-PADD-005), the CASDAR Project “Bioviresdur”, and a grant from the Santé des Plantes et Environnement (SPE) INRA department. G. Daverdin was funded by the French ministry of agriculture.

### **References**

Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. *Mol Ecol Notes* 1: 101-102.

- Ansan-Melayah D, Balesdent MH, Buee M, Rouxel T (1995). Genetic characterization of *AvrLm1*, the first avirulence gene of *Leptosphaeria maculans*. *Phytopathology* 85: 1525-1529.
- Ansan-Melayah D, Rouxel T, Bertrand J, Letarnec B, Mendes-Pereira E, Balesdent MH (1997) Field efficiency of *Brassica napus* specific resistance correlates with *Leptosphaeria maculans* population structure. *Eur J Plant Pathol* 103: 835-841.
- Antonarakis SE, Cooper DN (2010) Human gene mutation: mechanisms and consequences. In Vogel and Motulsky's human genetics: problems and approaches, IV<sup>th</sup> edition, Speicher MR, Antonarakis SE & Motulsky (Eds.), Springer-Verlag, Berlin, pp. 319-363.
- Attard A, Gout L, Ross S, Parlange F, Cattolico L, *et al.* (2005) Truncated and RIP-degenerated copies of the LTR retrotransposon *Pholy* are clustered in a pericentromeric region of the *Leptosphaeria maculans* genome. *Fungal Genet Biol* 42: 30-41.
- Aubertot JN, Schott JJ, Penaud A, Brun H, Doré T (2004) Methods for sampling and assessment in relation to the spatial pattern of phoma stem canker (*Leptosphaeria maculans*) in oilseed rape. *Eur J Plant Pathol* 110: 183-192.
- Aubertot JN, West JS, Bousset-Vaslin L, Salam MU, Barbetti MJ, Diggle AJ (2006) Improved management for durable disease control: a case study of phoma stem canker of oilseed rape (*Brassica napus*). *Eur J Plant Pathol* 114: 91-106.
- Balesdent MH, Attard A, Ansan-Melayah D, Delourme R, Renard M, Rouxel T (2001) Genetic control and host range of avirulence toward *Brassica napus* cultivars Quinta and Jet Neuf in *Leptosphaeria maculans*. *Phytopathology* 91: 70-76.
- Balesdent MH, Attard A, Kuhn ML, Rouxel T (2002) New avirulence genes in the phytopathogenic fungus *Leptosphaeria maculans*. *Phytopathology* 92: 1122-1133.

- Balesdent MH, Barbetti MJ, Li H, Sivasithamparam K, Gout L, Rouxel T (2005) Analysis of *Leptosphaeria maculans* race structure in a worldwide collection of isolates. *Phytopathology* 95: 1061-1071.
- Balesdent MH, Louvard K, Pinochet X, Rouxel T (2006) A large scale survey of races of *Leptosphaeria maculans* occurring on oilseed rape in France. *Eur J Plant Pathol* 114: 53-65.
- Balesdent MH, Daverdin G, Carpezat J, Rouxel T, Pinochet X, Leflon M (2011) *AvrLm7* allelic diversity in *Leptosphaeria maculans* nevertheless allows survey of virulent isolates using the HRM (High Melting Resolution) method. Proc. 13<sup>th</sup> International rapeseed congress, Prague 2011
- Balesdent MH, Louvard K, Pinochet X, Rouxel T (2006) A large-scale survey of races of *Leptosphaeria maculans* occurring on oilseed rape in France. *Eur J Plant Pathol* 114: 53-65.
- Belkheir K, Borsa P, Chiki L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations Interactions. CNRS UMR 5171, Université de Montpellier II, Montpellier (France).
- Blaise F, Remy E, Meyer M, Zhou L, Narcy JP, Roux J, Balesdent MH, Rouxel T (2007) A critical assessment of *Agrobacterium tumefaciens*-mediated transformation as a tool for pathogenicity gene discovery in the phytopathogenic fungus *Leptosphaeria maculans*. *Fungal Genet Biol* 44: 123-138.
- Brown AHD, Feldman MW, Nevo E (1980) Multilocus structure of natural-populations of *hordeum-spontaneum*. *Genetics* 96: 523-536.
- Brun H, Chèvre AM, Fitt BDL, Powers S, Besnard AL, Ermel M *et al.* (2010). Quantitative resistance increases the durability of qualitative resistance to *Leptosphaeria maculans* in *Brassica napus*. *New Phytol* 185: 285-299.



- Corpet F (1988). Multiple sequence alignment with hierarchical-clustering. *Nucleic Acids Res* 16: 10881-10890.
- Cozijnsen AJ, Howlett BJ (2003) Characterisation of the mating type locus of the plant pathogenic ascomycete *Leptosphaeria maculans*. *Current Genet* 43: 351-357.
- Delourme R, Chèvre A-M, Brun H, Rouxel T, Balesdent MH *et al.* (2006) Major gene and polygenic resistance to *Leptosphaeria maculans* in oilseed rape (*Brassica napus*). *Eur J Plant Pathol* 114: 41-52.
- Dilmaghani A (2010) World-wide population structure of the *Leptosphaeria maculans*–*Leptosphaeria biglobosa* species complex. PhD Thesis, Université Paris-Sud-XI, 256 pp.
- Dilmaghani A, Balesdent MH, Didier JP, Wu C, Davey J *et al.* (2009) The *Leptosphaeria maculans*–*Leptosphaeria biglobosa* species complex in the American continent. *Plant Pathol* 58: 1044-1058.
- Ellis JG, Lawrence GJ, Dodds PN (2007) Further analysis of gene-for-gene disease resistance specificity in flax. *Mol Plant Pathol* 8: 103-109.
- Farman MK (2007) Telomeres in the rice blast fungus *Magnaporthe oryzae*: the world of the end as we know it. *FEMS Microbiol Lett* 273: 125-132.
- Fitt BDL, Brun H, Barbetti MJ, Rimmer SR (2006) World-wide importance of phoma stem canker (*Leptosphaeria maculans* and *L. biglobosa*) on oilseed rape (*Brassica napus*). *Eur J Plant Pathol* 114: 3-15.
- Fudal I, Bohnert HU, Tharreau D, Lebrun MH (2005). Transposition of MINE, a composite retrotransposon, in the avirulence gene *ACE1* of the rice blast fungus *Magnaporthe grisea*. *Fungal Genet Biol* 42: 761-772.
- Fudal I, Ross S, Gout L, Blaise F, Kuhn ML *et al.* (2007) Heterochromatin-like regions as ecological niches for avirulence genes in *Leptosphaeria maculans* genome: map-based cloning of *AvrLm6*. *Mol Plant-Microbe Interact* 20: 459-470.

- Fudal I, Ross S, Brun H, Besnard AL, Ermel M *et al.* (2009) Repeat-Induced Point mutation (RIP) as an alternative mechanism of evolution toward virulence in *Leptosphaeria maculans*. *Mol Plant-Microbe Interact* 22: 932-941.
- Galagan JE, Selker EU (2004) RIP: the evolutionary cost of genome defense. *Trends Genet* 9: 417-423.
- Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND *et al.* (2003). The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422: 859-868.
- Gladders P, Musa TM (1980) Observations on the epidemiology of *Leptosphaeria maculans* stem canker in winter oilseed rape. *Plant Pathol* 29: 28-37.
- Gout L, Fudal I, Kuhn ML, Blaise F, Cattolico L *et al.* (2006) Lost in the middle of nowhere: the *AvrLm1* avirulence gene of *Leptosphaeria maculans*. *Molecular Microbiol* 60: 67-80.
- Gout L, Eckert MR, Rouxel T, Balesdent MH (2006) Genetic variability and distribution of mating type alleles in field populations of *Leptosphaeria maculans* from France. *Appl Environ Microbiol* 72: 185-191.
- Gout L, Kuhn ML, Vincenot L, Bernard-Samain S, Cattolico L *et al.* (2007) Genome structure impacts molecular evolution at the *AvrLm1* avirulence locus of the plant pathogen *Leptosphaeria maculans*. *Environ Microbiol* 9: 2978-2992.
- Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE *et al.* (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461: 393-398.
- Hammond-Kosack KE, Kanyuka K (2007) Resistance genes (*R* genes) in plants. In: *Encyclopedia of Life Sciences (ELS)*. John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0020119
- Hane JK, Oliver RP (2008) RIPCAL: a tool for alignment-based analysis of repeat-induced point mutations in fungal genomic sequences. *BMC Bioinformatics* 9: 478.

- Haubold B, Travisano M, Painey PB, Hudson RR (1998) Detecting linkage disequilibrium in bacterial population. *Genetics* 150: 1341-1348.
- Houterman PM, Cornelissen BJC, Rep M (2008) Suppression of plant resistance gene-based immunity by a fungal effector. *PLoS Path* 4: e1000061.
- Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Mol Ecol Notes* 5: 187-189.
- Kang S, Lebrun MH, Farrall L, Valent B (2001) Gain of virulence caused by insertion of a *Pot3* transposon in a *Magnaporthe grisea* avirulence gene. *Mol Plant-Microbe Interact* 14: 671-674.
- Lacoste L, Louvet J, Anselme C, Alabouvette C, Brunin R, Pierre JG (1969) Rôle de *Phoma lingam* (Tode) Desm. et de sa forme parfaite *Leptosphaeria maculans* (Desm.) Ces. & de Not. dans les épidémies de nécrose du collet du colza (*Brassica napus* L. var. *oleifera* Metzger). *CR Acad Agric France* 55: 981-989.
- Langella O (1997) Populations: a population genetic software. 1.2.28 ed: CNRS. UPR9034.
- Leclair S, Ansan-Melayah D, Rouxel T, Balesdent MH (1996) Meiotic behaviour of the minichromosome in the phytopathogenic ascomycete *Leptosphaeria maculans*. *Current Genet* 30: 541-548.
- Leach JE, Vera Cruz CM, Bai J, Leung H (2001) Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Ann Rev Phytopathol* 39: 187-224.
- Linde CC, Zhan J, McDonald BA (2002) Population structure of *Mycosphaerella graminicola*: from lesions to continents. *Phytopathology* 9: 946-955.
- Liu YG, Whittier RF (1995). Thermal asymmetric interlaced PCR-automatable amplification and sequencing of insert end fragments from p1 and YAC clones for chromosome walking. *Genomics* 25: 674-681.

- Lô-Pelzer E, Bousset L, Jeuffroy MH, Salam MU, Pinochet X *et al.* (2010) SIPPOM-WOSR: A Simulator for Integrated Pathogen POpulation Management of phoma stem canker on Winter Oil Seed Rape. I. Description of the model. *Field Crops Res* 118: 73–81.
- Lô-Pelzer E, Aubertot JN, Bousset L, Salam MU, Jeuffroy MH (2010) SIPPOM-WOSR: A Simulator for Integrated Pathogen POpulation Management of phoma stem canker on Winter Oil Seed Rape. II. Sensitivity analysis. *Field Crops Res* 118: 82-93.
- Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ *et al.* (2010). Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464: 367-373.
- Marcroft SJ, Sprague SJ, Pymer SJ, Salisburu, PA, Howlett, BJ (2004) Crop isolation, not extended rotation length, reduces blackleg (*Leptosphaeria maculans*) severity of canola (*Brassica napus*) in south-eastern Australia. *Aust J Exp Agric* 44: 601-606.
- McDonald BA, Linde C (2002) The population genetics of plant pathogens and breeding for durable resistance. *Euphytica* 12: 163-180.
- Morice J (1963) Les relations entre les espèces cultivées du genre *Brassica* et les possibilités d'amélioration du colza (*Brassica napus* L. var. *oleifera* Metzger) au moyen de croisements interspécifiques. *Ann Amel Plantes* 13: 221-252.
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sciences USA* 12: 3321-3323.
- Palloix, A, Ayme V, Moury B (2009) Durability of plant major resistance genes to pathogens depends on the genetic background, experimental evidence and consequences for breeding strategies ? *New Phytol* 183: 190–199.
- Parlange F, Daverdin G, Fudal I, Kuhn ML, Balesdent MH *et al.* (2009) *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4*

- and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change. *Molecular Microbiol* 71: 851- 863.
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6: 288-295.
- Plummer, KM, Howlett BJ (1995) Inheritance of chromosomal length polymorphisms in the ascomycete *Leptosphaeria maculans*. *Mol Gen Genet* 247: 416-422.
- Raymond M, Rousset F (1995) GENEPOP (version-1.2) - Population-genetics software for exact tests and ecumenicism. *J Heredity* 86: 248-249.
- Rice WR (1989) Analyzing tables of statistical tests. *Evolution* 43: 223-225.
- Rep M, van der Does HC, Meijer M, van Wijk R, Houterman PM *et al.* (2004). A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for *I-3*-mediated resistance in tomato. *Mol Microbiol* 53: 1373-83.
- Rouxel T, Willner E, Coudard L, Balesdent MH (2003). Screening and identification of resistance to *Leptosphaeria maculans* in *Brassica napus* accessions. *Euphytica* 133: 219-231.
- Rouxel T, Penaud A, Pinochet X, Brun H, Gout L, Delourme R, Schmit J, Balesdent MH (2003) A ten-year survey of populations of *Leptosphaeria maculans* in France indicates a rapid adaptation towards the *Rlm1* resistance gene in oilseed rape. *Eur J Plant Pathol* 109: 871-881.
- Rouxel T, Balesdent MH (2005) The stem canker (blackleg) fungus, *Leptosphaeria maculans*, enters the genomic era. *Mol Plant Pathol* 6: 225-241.
- Rouxel T, Balesdent MH (2010) Avirulence Genes. In: *Encyclopedia of Life Sciences (ELS)*. John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0021267.
- Rouxel T, Grandaubert J, Hane JK, Hoede C, van de Wouw AP *et al.* (2011) Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat Induced Point mutations. *Nat Comm* in press.

- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386.
- Schneider O, Roger-Estrade J, Aubertot JN, Doré T (2006) Effects of seeders and tillage equipment on vertical distribution of oilseed rape stubble. *Soil Tillage Res* 85: 115-122.
- Schurch S, Linde CC, Knogge W, Jackson LF, McDonald BA (2004). Molecular population genetic analysis differentiates two virulence mechanisms of the fungal avirulence gene *NIP1*. *Mol Plant-Microbe Interact* 17: 1114-1125.
- Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes D M *et al.* (2010). Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* 330: 1543-1546.
- Stachowiak A, Olechnowicz J, Jedryczka M, Rouxel T, Balesdent MH *et al.* (2006) Frequency of avirulence alleles in field populations of *Leptosphaeria maculans* in Europe. *Eur J Plant Pathol* 114: 67- 75.
- Stergiopoulos I, De Kock MJ, Lindhout P, de Wit PJ (2007). Allelic variation in the effector genes of the tomato pathogen *Cladosporium fulvum* reveals different modes of adaptive evolution. *Mol Plant-Microbe Interact* 20: 1271-1283.
- Stergiopoulos I, de Wit PJGM (2009) Fungal effector proteins. *Annu Rev Phytopathol* 47: 233-263.
- Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annu Rev Phytopathol* 46: 75-100.
- Sylvester-Bradley R (1985) Revision of a code for stages of development in oilseed rape (*Brassica napus* L.). *Aspects Appl Biol* 10: 395-400.

- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876-4882.
- van den Bosch F, Giligan CA (2003) Measures of durability of resistance. *Phytopathology* 93: 616-625.
- van de Wouw A, Cozijnsen AJ, Hane JK, Brunner PC, McDonald BA *et al.* (2010) Evolution of linked avirulence effectors in *Leptosphaeria maculans* is affected by genomic environment and exposure to resistance genes in host plants. *PLoS Pathogens* 6:e1001180.
- van Leeuwen FW, Fischer DF, Kamel D, Sluijs JA, Sonnemans MA *et al.* (2000) Molecular misreading: a new type of transcript mutation expressed during aging. *Neurobiol Aging* 21:879-891.
- West JS, Kharbanda PD, Barbetti MJ, Fitt BDL (2001) Epidemiology and management of *Leptosphaeria maculans* (phoma stem canker) on oilseed rape in Australia, Canada and Europe. *Plant Pathol* 50: 10-27.
- West JS, Balesdent MH, Rouxel T, Narcy JP, Huang YJ, Roux J, Steed JM, Fitt BDL, Schmit J (2002) Colonization of winter oilseed rape tissues by A/Tox(+) and B/Tox(0) *Leptosphaeria maculans* (phoma stem canker) in France and England. *Plant Pathol* 51: 311-321.
- Zhou E, Jia Y, Singh P, Correll JC, Lee FN (2007). Instability of the *Magnaporthe oryzae* avirulence gene *AVR-Pita* alters virulence. *Fungal Genet Biol* 44: 1024-1034.

Table 1. Number and origin of isolates collected in the experimental evolution trial and phenotyped for virulence towards *Rlm4* and *Rlm7*.

Season/ Year	Location	Cultivar and organ sampled <sup>a</sup>	No of leaves/ stems collected <sup>b</sup>	No of isolates obtained	No of <i>L.</i> <i>maculans</i> isolates phenotyped <sup>c</sup>	No (%) of isolates in each phenotypic class <sup>d</sup>			
						<i>avrLm4</i> - <i>avrLm7</i>	<i>avrLm4</i> - <i>AvrLm7</i>	<i>AvrLm4</i> - <i>AvrLm7</i>	nd
Autumn 2006	Grignon	Campala/leaves	500	500	445	41 (9.3%)	384 (86.3%)	12 (2.7%)	8 (1.8%)
		Exagone/leaves	24	35	35	14 (40%)	21 (60%)	0 (0%)	0 (0%)
	Versailles	Campala/leaves	500	500	246	0 (0%)	242 (98.4%)	4 (1.7%)	0 (0%)
		Exagone/leaves	37	37	35	0 (0%)	35 (100%)	0 (0%)	0 (0%)
Summer 2007	Grignon	Campala/stems	100	43	40	9 (22.5%)	31 (77.5%)	0 (0%)	0 (0%)
		Exagone/stems	100	108	101	85 (84.2%)	16 (15.9%)	0 (0%)	0 (0%)
	Versailles	Campala/stems	100	7	5	1 (20%)	4 (80%)	0 (0%)	0 (0%)
		Exagone/stems	100	12	5	3 (60%)	2 (40%)	0 (0%)	0 (0%)
Autumn 2007	Grignon	Campala/leaves	200	184	178	56 (31.5%)	118 (66.3%)	3 (1.7%)	1 (0.6%)
		Exagone/leaves	200	191	187	176 (94.2%)	10 (5.4%)	1 (0.6%)	0 (0%)
	Versailles	Campala/leaves	200	200	160	1 (0.7%)	153 (95.7%)	3 (1.9%)	3 (1.9%)
		Exagone/leaves	21	21	20	0 (0%)	20 (100%)	0 (0%)	0 (0%)
Summer 2008	Grignon	Campala/stems	200	260	235	85 (36.2%)	139 (59.2%)	10 (4.3%)	1 (0.5%)
		Exagone/stems	200	523	355	336 (94.7%)	12 (3.4%)	7 (2.0%)	0 (0%)
Total				2621	2047	807	1187	40	13

<sup>a</sup> Exagone is an oilseed rape cv. harbouring the resistance gene *Rlm7*, whereas Campala is devoid of it

<sup>b</sup> One single-conidium isolate was obtained from each leaf lesion, except for Exagone/leaves 2006 for which very few leaf lesions were found, whereas more than one single-ascospore isolate could be obtained from each stem residues

<sup>c</sup> Excluding isolate belonging to the other species of the species complex, *L. biglobosa*, and those for which the interaction phenotype could not be clearly established

<sup>d</sup> *avrLm4-avrLm7*, isolates virulent on both *Rlm4* and *Rlm7*; *avrLm4-AvrLm7*, isolates avirulent on *Rlm7* only; *AvrLm4-AvrLm7*, isolates avirulent on both *Rlm4* and *Rlm7*; nd, phenotypic interaction on *Rlm4* is missing



Table 2. Number and origin of isolates obtained and phenotyped for virulence towards *Rlm4* and *Rlm7* : collections obtained before or close to the experimental evolution trial

Season / year	Location	Cultivar and organ sampled <sup>a</sup>	No of leaves /stems collected <sup>b</sup>	No of isolates obtained	No of <i>L. maculans</i> isolates phenotyped <sup>c</sup>	Interaction phenotype (%) <sup>d</sup>		
						<i>avrLm4-avrLm7</i>	<i>avrLm4-AvrLm7</i>	<i>AvrLm4-AvrLm7</i>
Autumn 2000	Grignon	Drakkar / leaves	100	100	100	0.00	89.00	11.00
	Versailles	Drakkar / leaves	100	87	87	0.00	80.46	19.54
Summer 2002	Grignon	Bristol / stem	47	98	98	0.00	82.65	17.35
		Pollen / stem	36	93	93	0.00	100.00	0.00
Autumn 2002	Versailles	Darmor / leaves	46	46	46	0.00	93.48	6.52
		Mohican / leaves	49	49	49	0.00	97.96	2.04
Summer 2003	Versailles	Mohican / stem	29	29	29	3.45	86.21	10.34
Autumn 2003	Versailles	Mohican / leaves	17	17	17	0.00	94.12	5.88
summer 2004	Versailles	Mohican / stem	210	210	210	0.00	87.62	12.38
Autumn 2006	Grignon, field 2 <sup>e</sup>	Grizzly / leaves	100	100	47	0.00	0.00	100.00
	Grignon, field 3	Campala / leaves	100	84	46	0.00	93.48	6.52

<sup>a</sup>Drakkar is an oilseed rape cv. devoid of any *Rlm* gene; Mohican, Campala and Darmor contain *Rlm9*, Bristol *Rlm2* and *Rlm9*, Pollen *Rlm4*, and Grizzly *Rlm2* and *Rlm3*

<sup>b</sup> One single-conidium isolate was obtained from each leaf lesion, whereas more than one single-ascospore isolate could be obtained from each stem residues

<sup>c</sup> Excluding isolates belonging to the other species of the species complex, *L. biglobosa*, and those for which the interaction phenotype could not be clearly established

<sup>d</sup> *avrLm4-avrLm7*, isolates virulent on both *Rlm4* and *Rlm7*; *avrLm4-AvrLm7*, isolates avirulent on *Rlm7* only; *AvrLm4-AvrLm7*, isolates avirulent on both *Rlm4* and *Rlm7*.

<sup>e</sup> Collected from fields located less than 600 m away from the experimental evolution trial

Table 3. Predicted amino acid polymorphism found in AvrLm4-7 proteins produced by different alleles of the avirulent, *AvrLm7*, form of the gene, and their occurrence in populations of the experimental evolution trial in Grignon.

Interaction phenotype alleles <sup>a</sup>	Amino acid position <sup>b</sup>					No (frequency) of isolates	Representative isolate for this allele
	45	74	80	86	120		
<i>AvrLm4-AvrLm7</i>	Leu	Val	Ile	Asp	Gly	7 (4.1%)	v23.1.3
<i>avrLm4-AvrLm7</i>	Leu	Val	<u>Thr</u>	Asp	<u>Arg</u>	62 (36.7%)	v23.1.2
<i>avrLm4-AvrLm7</i>	Leu	Val	Ile	<u>Asn</u>	<u>Arg</u>	86 (50.9%)	G08-C1049
<i>avrLm4-AvrLm7</i>	<u>Ser</u>	Val	Ile	Asp	<u>Arg</u>	13 (7.7%)	G06-485
<i>avrLm4-AvrLm7</i>	Leu	<u>Ile</u>	Ile	Asp	<u>Arg</u>	1 (0.6%)	G08-C1226

<sup>a</sup> *avrLm4-AvrLm7*, isolates avirulent on *Rlm7* only; *AvrLm4-AvrLm7*, isolates avirulent on both *Rlm4* and *Rlm7*.

<sup>b</sup> Comparisons are based on the sequence of the reference *AvrLm4-7* protein of isolate v23.1.3 inducing avirulence towards *Rlm4* and *Rlm7* genotypes. Leu45Ser is due to a T to C mutation at base 123; Val74Ile is due to G to A mutation at base 220; Ile80Thr is due to a T to C mutation at base 239; Asp86Asn is due to a G to A mutation at base 256; Gly120Arg is due to a G to C mutation at base 358. The underlined amino acids are those showing polymorphism as compared to the reference sequence.

Table 4. Point mutations responsible for the *avrLm7* phenotype and their frequency in Grignon experimental evolution population.

Amino acid change	No (frequency) of isolates <sup>a</sup>	Representative isolate <sup>b</sup>
Cys <sup>22</sup> → Arg <sup>22</sup>	1 (4.35%)	G07-E210
Gln <sup>35</sup> → STOP	7 (30.4%)	G07-E480 (B18)
Leu <sup>55</sup> → STOP	1 (4.35%)	G08-E1288
Cys <sup>79</sup> → Ser <sup>79</sup>	1 (4.35%)	G07-E295
Arg <sup>100</sup> → Pro <sup>100</sup>	10 (43.5%)	G07-E1006
Phe <sup>102</sup> → Ser <sup>102</sup>	1 (4.35%)	G07-E497
Ser <sup>112</sup> → Arg <sup>112</sup>	2 (8.7%)	G06-E202

<sup>a</sup> Excluding isolates obtained from the same stubble which may represent sister isolates from the same cross

<sup>b</sup> The deduced protein sequence for representative isolates are displayed in Fig. 4

Table 5. Estimates of genetic diversity and linkage disequilibrium revealed by 4 minisatellites in samples of *Leptosphaeria maculans* collected at the Grignon experimental evolution site between 2006 and 2008

Population <sup>a</sup>	CaLeaf06	ExLeaf06	CaLeaf07	CaStem07	ExLeaf07	ExStem07	CaStem08	ExStem08
N <sup>b</sup>	81	35	75	12	164	97	118	309
Mat1.1/Mat1.2	48/33	18/17	38/37	5/7	95/69*	50/47	58/60	154/155
Na	9.30 ± 1.75	7.30 ± 1.31	9.30 ± 1.65	4.80 ± 0.63	10.80 ± 1.75	10.50 ± 1.66	9.30 ± 1.03	12.50 ± 1.76
Na Freq. ≥ 5%	4.50 ± 0.50	4.50 ± 0.65	5.00 ± 0.58	4.80 ± 0.63	4.00 ± 0.41	4.80 ± 0.85	5.50 ± 0.65	4.50 ± 0.87
MLG	72	35	65	11	124	86	92	221
N <sub>MLG</sub>	1.13 ± 0.13	1.00 ± 0.00	1.15 ± 0.12	1.09 ± 0.18	1.32 ± 0.15	1.13 ± 0.10	1.28 ± 0.14	1.39 ± 0.13
G	1.00	1.00	0.99	0.98	0.99	1.00	0.99	1.00
H	0.76 ± 0.04	0.75 ± 0.05	0.75 ± 0.06	0.69 ± 0.08	0.73 ± 0.06	0.75 ± 0.07	0.74 ± 0.03	0.75 ± 0.05
r <sub>d</sub>	0.002 <sup>ns</sup>	-0.032 <sup>ns</sup>	0.011 <sup>ns</sup>	-0.005 <sup>ns</sup>	0.014 <sup>ns</sup>	0.003 <sup>ns</sup>	0.003 <sup>ns</sup>	0.000 <sup>ns</sup>

<sup>a</sup> Populations sampled from leaf or stem lesions from the Campala cv. were identified as CaLeafxx or CaStemxx, respectively, with xx standing for the year of collection and samples isolated from the Rlm7 cv. Exagone were similarly identified as ExLeafxx or ExStemxx

<sup>b</sup> N, number of isolates; Mat1.1/Mat1.2, 2 is the ratio of the number of isolates carrying mating type 1 (mat1-1) to isolates carrying mating type 2 (mat1-2). (\*) Significant deviation from a 1:1 ratio was tested using a  $\chi^2$  with 1 degree of freedom ( $P < 0.05$ ); Na, mean number of different alleles; Na Freq. ≥ 5%, mean number of different alleles with frequency >0.05; MLG, number of multilocus genotypes; N<sub>MLG</sub>, mean number of multilocus genotypes; G, index of genotypic diversity; H, unbiased gene diversity; r<sub>d</sub> standardized version of the index of association as calculated in Multilocus v3.1 (ns : not significant).

Table 6.  $F_{ST}$  estimates for all pairs of samples of *L. maculans* collected from the Campala and Exagone cultivars at the Grignon experimental evolution site between 2006 and 2008

	CaLeaf06 <sup>a</sup>	ExLeaf06	CaLeaf07	CaStem07	ExLeaf07	ExStem07	CaStem08
ExLeaf06	0.001	-					
CaLeaf07	0.000	<b>0.013*</b> <sup>b</sup>	-				
CaStem07	0.000	0.009	0.003	-			
ExLeaf07	<b>0.006*</b>	0.003	0.001	0.002	-		
ExStem07	0.001	0.007	0.000	0.000	0.002	-	
CaStem08	0.000	<b>0.010*</b>	0.002	0.000	<b>0.009*</b>	<b>0.011*</b>	-
ExStem08	0.000	0.003	<b>0.004*</b>	0.006	<b>0.006*</b>	<b>0.002*</b>	<b>0.008*</b>

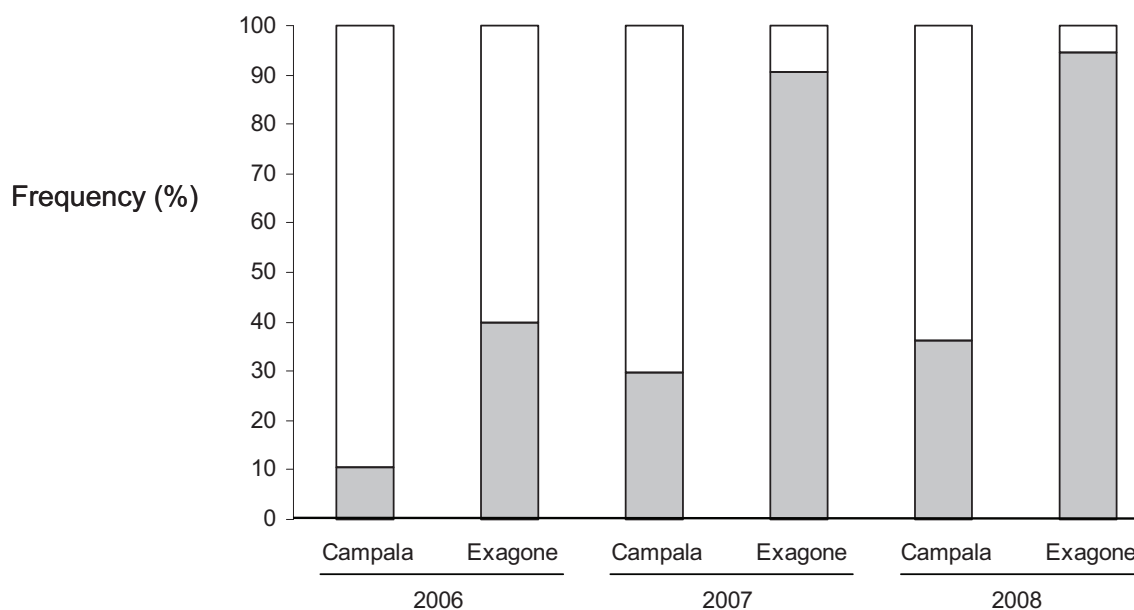
<sup>a</sup> Populations sampled from leaf or stem lesions from the Campala cv. were identified as CaLeafxx or CaStemxx, respectively, with xx standing for the year of collection and samples isolated from the Rlm7 cv. Exagone were similarly identified as ExLeafxx or ExStemxx

<sup>b</sup> Significant values are shown in bold type (\* $P < 0.05$ )

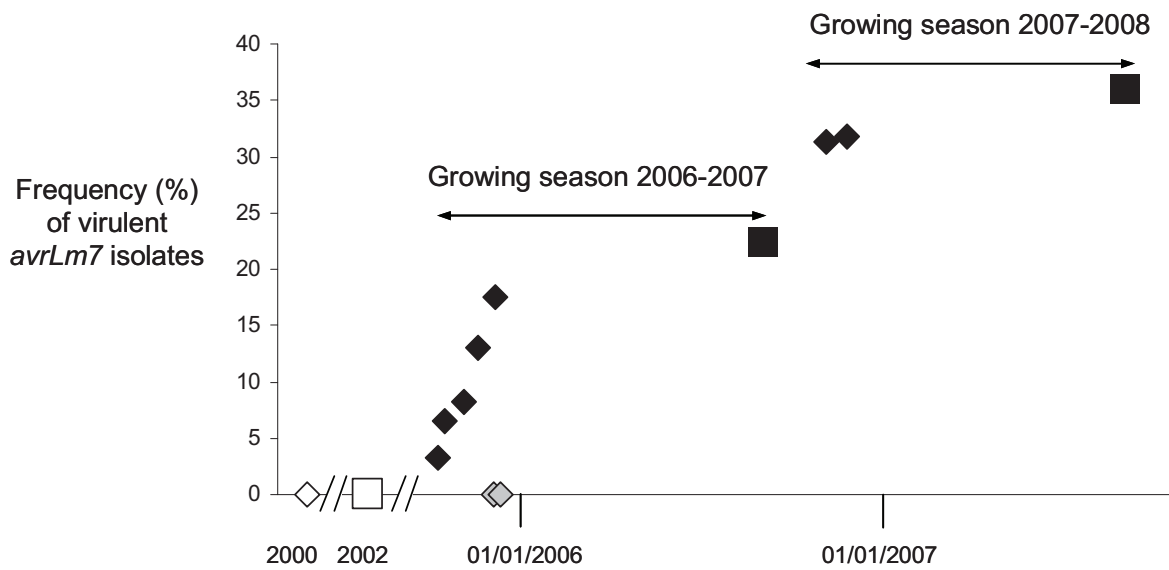
Table 7.  $F_{ST}$  estimates for all pairs of groups of *L. maculans* isolates collected from the Campala and Exagone cultivars at the Grignon experimental evolution site between 2006 and 2008. Combination of isolates across samples was done using the type of molecular event leading to virulence on *R/m7* cultivars by comparison with local populations showing avirulence towards *R/m7*.

	AA dinucleotide deletion and point deletions	Avirulent	Large scale deletions	RIP mutations	Single Nucleotide Polymorphisms	Small-scale deletions
Avirulent	0.000	-				
Large scale deletions	<b>0.003*</b> <sup>a</sup>	0.002	-			
RIP mutations	0.001	0.000	<b>0.003*</b>	-		
Single Nucleotide Polymorphisms	0.010	<b>0.010*</b>	0.005	0.003	-	
Small-scale deletions	0.003	0.000	<b>0.006*</b>	<b>0.005*</b>	<b>0.016*</b>	-
Unaltered gene sequence	0.000	0.000	0.000	0.000	0.000	0.000

<sup>a</sup> Significant values are shown in bold type (\*P < 0.05)

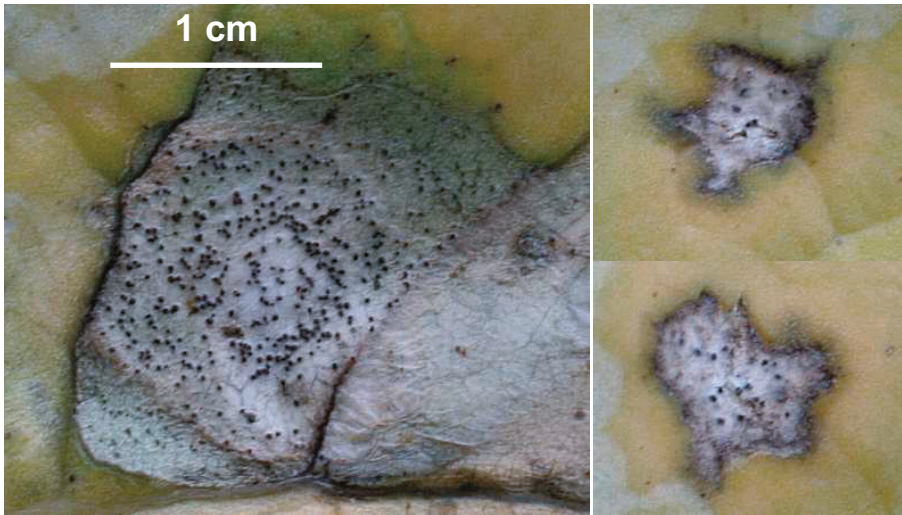


**Figure. 1. Evolution of the frequency of virulent, *avrLm7* isolates, at the Grignon experimental evolution site.** Isolates were obtained from the *Rlm7* cv. Exagone and the cv. devoid of *Rlm7* (Campala) during the three years of the sampling. The plot was an oilseed rape monoculture associated to simplified tillage. Grey bars, virulent, *avrLm7* isolates; white bars, avirulent, *AvrLm7* isolates.



**Figure 2. Evolution of the frequency virulent avrLm7 isolates of *Leptosphaeria maculans* obtained from susceptible cultivars.** A total of 1278 isolates virulent toward *Rlm7* have been recovered from non-screening cvs (without the resistance gene *Rlm7*) in the experimental evolution trial and its immediate surroundings between 2000 and 2008. Isolates were obtained either at the leaf stage (diamonds; beginning of the infection cycle) or at the stem base (squares, end of the infection cycle, after one sexual reproduction cycle). Isolates were collected before the establishment of the assay (white dots), within the field trial (black dots) or in fields sown close (<600m) to the experimental trial (grey diamonds).

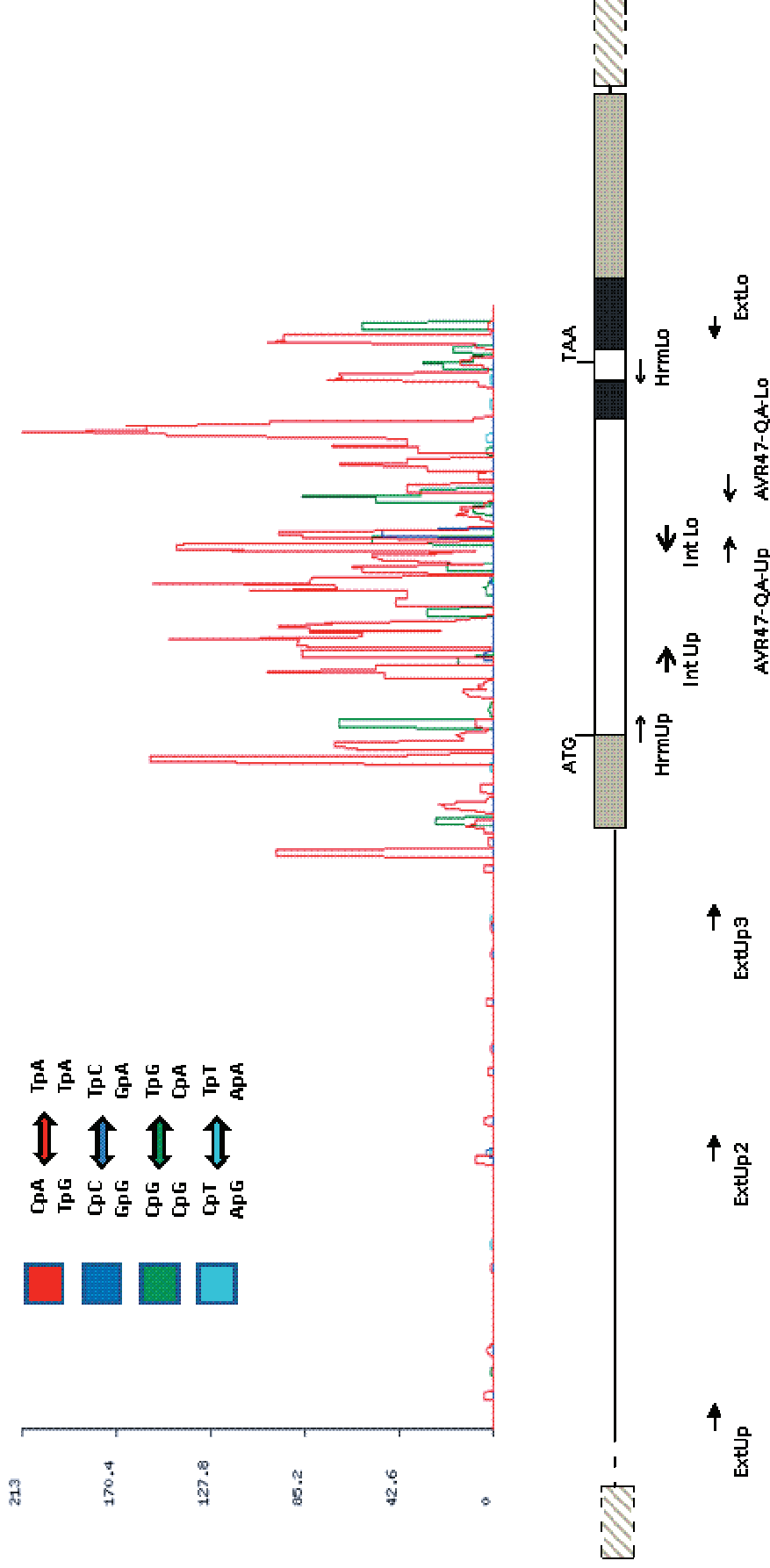




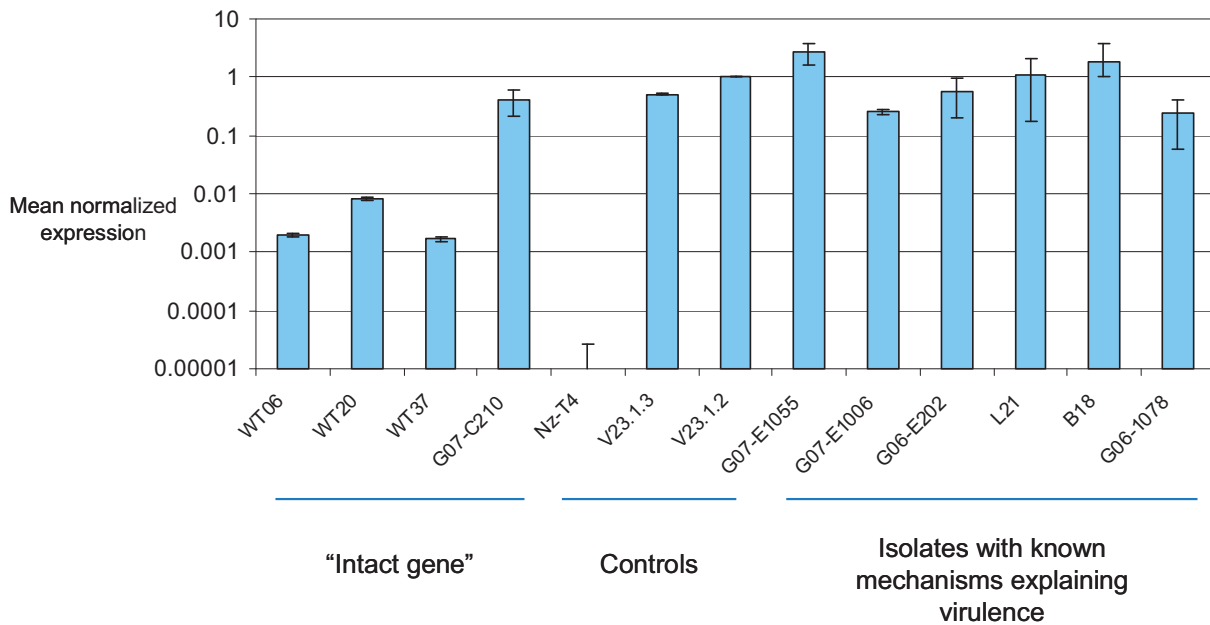
**Figure 3.** An example of *Leptosphaeria maculans* leaf lesions developed on the *Rlm7* cv. **Exagone**. Left, the isolate recovered from this lesion, G06-E101, displayed a virulent, *avrLm7* phenotype following further inoculation on a *Rlm7* line; right, the isolates recovered from these lesions, G06-E102 (up) and G06-E103 (down) displayed an avirulent, *AvrLm7* phenotype following inoculation on a *Rlm7* line. These show limited lesion development seemingly blocked by a late necrotic response of the plant. The three isolates were recovered from a single leaf, collected in October 2006 in the Grignon experimental field.



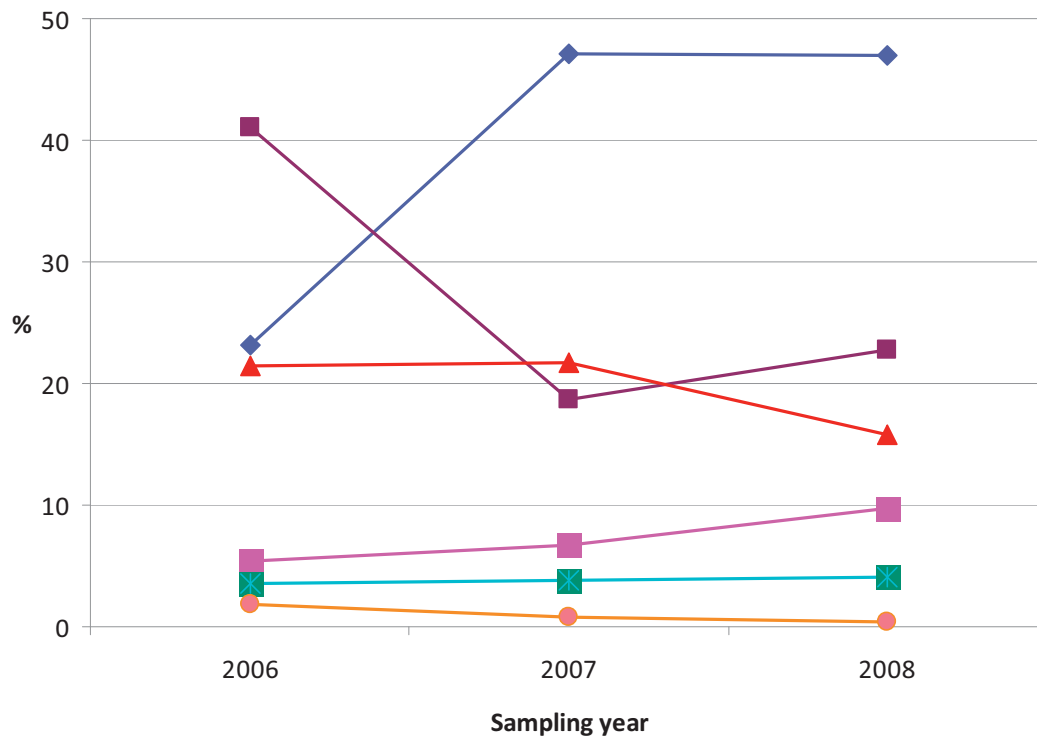
**Figure 4. GeneDoc-edited alignment of representative variants of the AvrLm4-7 protein.**



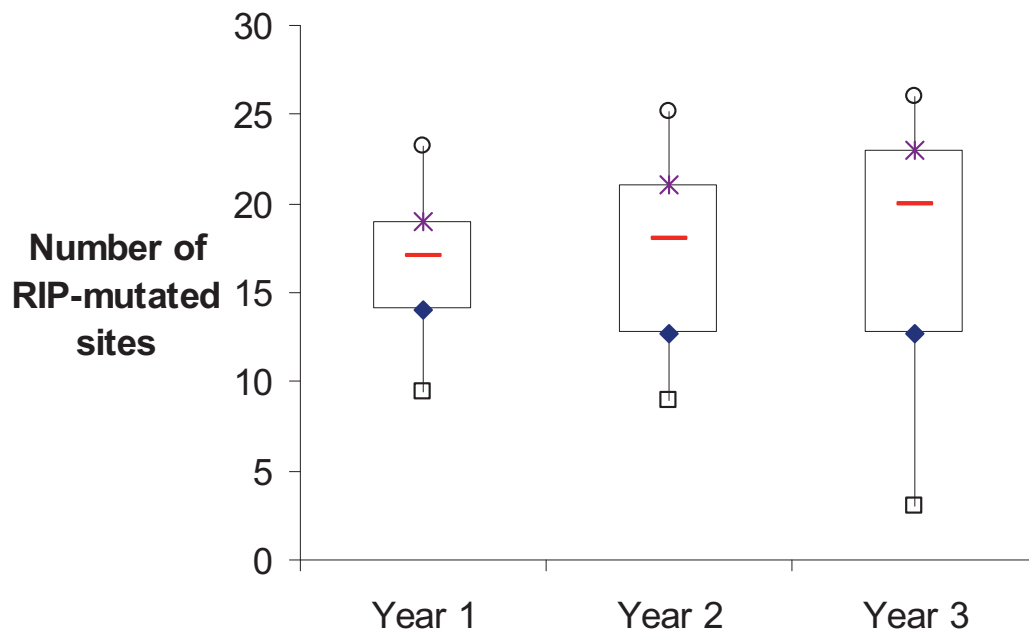
**Figure 5. RIPCAL analysis of 127 RIPped *AvrLm4-7* alleles and schematic representation of the *AvrLm4-7* gene.** The alleles were identified in virulent, *avrLm7* isolates obtained 2006 and 2008 in the Grignon experimental evolution plot. The sequence with the highest total G+C content (v23.1.2) was chosen as the least RIP-mutated model for comparison of all aligned sequences. Physical distribution of RIP along the *AvrLm4-7* gene is illustrated by the overall RIP mutation frequency over a 50 bp scanning window. A strong over-representation of the CpATpG↔TpA mutation (red) over CpN↔TpN or NpG↔NpA mutations (blue and green curves) is observed. The *AvrLm4-7* sequence analysed here (encompassing ca. 550 bp of promoter region and ca. 60 bp of the 3'UTR) is represented along with the location of primers used in this study for PCR, Q-RT-PCR, sequencing and HRM analyses. White, coding sequence; black, introns; grey, 5' and 3' UTR; hashed boxes, Transposable elements



**Figure. 6. Expression of the *AvrLm4-7* gene of virulent, *avrLm7* isolates, *in planta* 12 days post inoculation.** Virulent isolates showed either an unaltered coding sequence (G07-C210, and collection isolates WTxx, L21 and B18) or were representative of a few SNP events leading to inactivation of the *AvrLm7* activity (see Fig. 4). L21 allele is similar to that of G07-E497 in Fig. 4 and B18 allele is similar to that of G07-E480 in Fig. 4; G06-1078, mildly RIPped allele resulting in two non-synonymous changes in the protein. *AvrLm4-7* expression was analysed by qRT-PCR and measured relatively to  $\beta$ -tubulin. Each data point is the average of two technical repeats and two biological repeats were done for WTxx isolates for which an altered expression level was observed. Standard error of mean normalized expression level is indicated by error bars.



**Figure 7. Dynamics of the mutational events leading to the *avrLm7* phenotype over three years.** Dark blue diamonds, large scale deletions; purple squares, RIP mutations; red triangles, small-scale deletions with part of the gene being unaffected; pink squares, AA dinucleotide deletion and point deletions; green squares, Single Nucleotide Polymorphisms; orange circles, unaltered gene sequence



**Figure 8. Box plot analysis of the number of RIP-mutations in the *AvrLm4-7* gene.** RIP mutations were responsible for the *avrLm7* phenotype, and the analysis was done over the three years of the experimental evolution plot.

**Supplementary Table S1.** Primers and PCR conditions used in this study

Markers	Primers (5'-3')		Hybridation temperature (°C)	Elongation time (s)
	Forward	reverse		
<b><u>AvrLm4-7 amplification and sequencing</u></b>				
AvrLm4-7 ext	TATCGCATACCAACATTAGGC	GATGGATCAACCGCTAACAA	60	90
AvrLm4-7 Int	ATATCTGGAGAAATTCGCTATC	CCAAGGGTCGGTAGTTATGC	60	30
AvrLm4-7 ext-Up2	ATTTGCTACACTAGATTATAC	-	60	-
AvrLm4-7 ext-Up3	AACCCTGCTAGATAGTAAAGCT	-	60	-
<b><u>High Resolution Melting PCR</u></b>				
AvrLm7-HRM	AACATGCCACTATCCCTC	ACCTCCGTATCTTTAGTC	54	45
<b><u>Quantitative PCR</u></b>				
AVR47-QA (AvrLm4-7)	GCCCTGCATAACTACCGAC	TCCTGGCCAAATATAACTCC	60	15
qPCRTub ( $\beta$ -tubulin)	AAGAACTCATCCTACTTCGT	TGAATAGCTCCTGAATGG	60	15
<b><u>Minisatellite markers</u></b>				
MinLm3	GGCTCGGTCGGTTAGTTA	AATGATGTACAGGACGGGATTT	60	60
MinLm6	GGAAGGAACACACGGTGAC	AATTGAATGATTTGCGACACA	60	60
MinLm8	ATTTGCTGGCGGTG TAGGTA	TGTTTGTACATGTGGTAAAGTAAAGCA	60	60
MinLm9	GCATTTGTTGGCTGGCATTGGA	TGCGTTGGCGTGTGACATGACT	60	60
MinLm1377	CGTCCAGTCGTGCCCTGCTTTG	GGTGCAATGCTTGGCGGACCATT	58	60
MinLm2451	GGGGCGAATGGTATGTTTATAGT	CGGACACAATACTCACCCACCTC	58	60
MinLm4-22	GCAGAACGAGACCCGAGTACA	GTTGGCCCTGGTCATAGAGA	60	60
MinLm6-59	GAGAAACTTGCTGGCCCTGG	GAGGCCAATGTCCAATGTGT	60	60
<b><u>Tail PCR</u></b>				
Tail-GD1	CTCAAGGCACCTTGTCACCAC	-	59	-
Tail-GD2	CCACAAAGACTGAAGCTTTGAATG	-	61	-
Tail-GD3	ATATGACGGCCAAAGTCCA	-	59	-
AD1	-	NTCGA(G/C)T(A/T)T(G/C)G(A/T)GTT	-	-
AD2	-	NGTCGA(G/C)(A/T)GANA(A/T)GAA	-	-
AD3	-	(A/T)GTGNAG(A/T)ANCANAGA	-	-







# **Chapitre 3**

## **Interaction *AvrLm4-7* – *AvrLm3***



## Introduction

Deux clusters génétiques de gènes d'avirulence ont été répertoriés chez *L. maculans*, *AvrLm1-2-6* et *AvrLm3-4-7-9*. Parmi les gènes du second cluster, le clonage d'*AvrLm7* (Chapitre 1) a permis de montrer que *AvrLm4* et *AvrLm7* étaient deux allèles du même gène, appelé dès lors *AvrLm4-7*, tandis que *AvrLm9* a été localisé à 22,6 cM d'*AvrLm4-7* (Balesdent *et al.*, 2005). *AvrLm3* n'a quant à lui jamais pu être localisé physiquement du fait de la constante liaison en répulsion des phénotypes *AvrLm3* et *AvrLm7* dans la descendance des croisements testés à cet effet (co-localisation génétique). Il était alors envisagé que *AvrLm3* puisse être un nouvel allèle d'*AvrLm4-7*.

La proportion de souches avirulentes vis-à-vis de *Rlm3* (A3) en France et en Europe s'avère extrêmement faible (Balesdent *et al.*, 2006; Stachowiak *et al.*, 2006). Le fait que la première souche virulente vis-à-vis de *Rlm7* (a7) à avoir été trouvée en France (CB-M3.2) (Balesdent *et al.*, 2006) soit une souche A3 a soulevé de nombreuses questions. On pouvait émettre l'hypothèse que l'avirulence vis-à-vis de *Rlm7* (A7) ait été perdue dans une souche Française A3 ou encore que cet événement ait eu lieu dans une population comportant une plus grande fréquence de souches A3. Une troisième hypothèse étant que la mutation impliquée dans la perte de l'avirulence A7 soit à l'origine de la résurgence du phénotype A3.

Afin de déterminer laquelle de ces hypothèses était valide nous avons entamé les démarches suivantes :

- 1) Une exploitation des données de la bibliographie concernant la présence des différentes races phénotypiques de *L. maculans* ainsi qu'une analyse phénotypique plus poussée de la collection enrichie en souches a7 obtenue dans le chapitre 2 de cette thèse devaient permettre de déterminer si la perte de l'avirulence A7 constatée dans de très nombreuses souches provoquait chez celles-ci un changement de phénotype envers *Rlm3*. L'hypothèse selon laquelle *AvrLm3* seraient un nouvel allèle d'*AvrLm4-7* devait également être testée par la même occasion.
- 2) L'analyse rétrospective d'un croisement entre deux souches a3 dans la descendance duquel des souches A3 sont réapparues
- 3) La mise en place d'un nouveau croisement visant à l'analyse génétique et à la cartographie génétique d'*AvrLm3*.
- 4) Des analyses de complémentation fonctionnelle et d'extinction d'expression par RNAi pour élucider l'interaction fonctionnelle entre *AvrLm4-7* et *AvrLm3*.
- 5) L'initiation d'une marche chromosomique visant au clonage d'*AvrLm3* et prenant avantage de la disponibilité de la séquence du génome de *L. maculans*.

Ces démarches ainsi que les résultats obtenus ayant permis de démontrer l'existence d'un unique gène *AvrLm3* à l'origine de l'avirulence A3 et d'une relation antagoniste entre *AvrLm3* et *AvrLm4-7* sont ici présentés sous la forme d'un article intitulé « A game of hide and seek: the complex interplay between the avirulence genes *AvrLm4-7* and *AvrLm3* in *Leptosphaeria maculans* » devant être soumis au journal « New Phytologist ».

A game of hide and seek: the complex interplay between the avirulence genes  
*AvrLm4-7* and *AvrLm3* in *Leptosphaeria maculans*

**Authors**

Guillaume Daverdin, Bénédicte Ollivier, Isabelle Fudal, Françoise Blaise, Alexandre  
Degrave, Thierry Rouxel and Marie -Hélène Balesdent\*

INRA-Biogère, UR1290, Avenue Lucien Brétignières, BP 01, F-78850 Thiverval-Grignon,  
France

\* Corresponding author; Tel 33 (1) 30 81 45 73, [mhb@versailles.inra.fr](mailto:mhb@versailles.inra.fr)

**Word count :**

**Introduction: 1004**

**Material and Methods: 1785**

**Results: 1866**

**Discussion: 1970**

**Acknowledgements: 75**

**Total: 6700**

## Summary

- The relationship between two avirulence genes, *AvrLm4-7* and *AvrLm3*, of *Leptosphaeria maculans*, the filamentous ascomycete causing stem canker of oilseed rape, was investigated using both molecular and genetic evidences. The avirulence gene *AvrLm4-7* confers a double specificity of interaction toward the resistance genes *Rlm4* and *Rlm7*. Co-segregation of *AvrLm3* and *AvrLm4-7* suggested *AvrLm3* could be a potential novel allele of *AvrLm4-7*.

- Survey of natural populations of *L. maculans* submitted or not to *Rlm7* selection pressure and crosses between isolates possessing or not functional alleles of *AvrLm4-7* and *AvrLm3* led to the determination of genetic relationships between *AvrLm4-7* and *AvrLm3* and to genetic mapping of *AvrLm3*. Complementation and silencing experiments were used to determine whether presence of *AvrLm4-7* or change in its expression modified *AvrLm3* recognition.

- Population survey, crosses and genetic mapping showed that the avirulent phenotype toward *Rlm3* is actually governed by a single gene, *AvrLm3*, distinct from *AvrLm4-7*, located in a telomeric or sub-telomeric region. Complementation and silencing experiments demonstrated that the presence of a functional allele of *AvrLm4-7* prevents *Rlm3*-mediated recognition of *AvrLm3*.

- The increase of *AvrLm3* isolates following *Rlm7* selection pressure could be exploited with either a pyramiding strategy combining *Rlm3* and *Rlm7* in the same genotype, or by alternation strategies where *Rlm7* and *Rlm3* genotypes successively used.

## Key words (5-8) (words in the title must be in)

*Leptosphaeria maculans*, *Brassica napus*, oilseed rape, avirulence, interaction, resistance

## **Introduction**

This last decade, an increasing number of avirulence genes of phytopathogenic fungi have been genetically identified, cloned and characterized (Stergiopoulos & de Wit 2009, Rouxel & Balesdent, 2010). In its first acceptation, an avirulence (*Avr*) gene encodes a protein which is specifically “recognized” by genotypes of the host plant harbouring the corresponding resistance (*R*) gene. Identification and molecular characterization of fungal *Avr* genes firstly aimed at understanding the molecular dialogue that plants and their pathogens establish during the primary steps of infection, leading either to resistance or to susceptibility. It also provided tools for a better management of resistance sources, such as molecular markers to monitor virulent isolates in natural populations, or indications on the potential durability of the matching *R* genes whenever the fungal fitness penalty due to the *Avr* gene loss could be estimated (Huang *et al.*, 2010; Leach *et al.*, 2001).

*Leptosphaeria maculans* is a filamentous ascomycete causing stem canker disease of oilseed rape (*Brassica napus*), a damaging disease world-wide (Fitt *et al.*, 2006). The main way to control the disease is the use of resistant *B. napus* varieties, that may combine quantitative resistance and specific, major resistance genes, termed *Rlm* (Delourme *et al.*, 2006). Population genetic studies indicated that *L. maculans* displays a very high level of gene and genotypic diversity, a low level of population differentiation and a high gene flow between populations (Gout *et al.*, 2006a). Accordingly, this fungus also has the ability to very rapidly overcome major gene resistances when used in fields where the local population is mainly composed of avirulent isolates (Sprague *et al.*, 2006).

Recently, the genome of *L. maculans* was sequenced (Rouxel *et al.*, 2011). This genome is organised into distinct GC-equilibrated and AT-rich blocks, referred to as isochores. AT-



isochores are mainly composed of mosaics of repeated elements that have undergone Repeat-Induced Point (RIP) mutation degeneracy, a premeiotic repeat inactivation mechanism specific to fungi and targeting repeated elements. Effectors of the fungus, including its avirulence genes, were found to be located within AT-isochores (Rouxel *et al.*, 2011). Numerous *L. maculans* avirulence genes (*AvrLm*) have been identified by formal genetics. Of the ten genes (*AvrLm1-AvrLm10*) described, most of them are located in two independent genetic clusters (Balesdent *et al.*, 2002). The first one contains *AvrLm1*, *AvrLm2*, *AvrLm6*, along with two additional genes, *LmCys1* and *LmCys2*, encoding proteins with characteristics similar to those encoded by the *Avr* genes (Fudal *et al.*, 2007, Van de Wouw *et al.*, 2010). The second cluster contains *AvrLm3*, *AvrLm4*, *AvrLm7* and *AvrLm9* (Balesdent *et al.*, 2002, 2005). Following the cloning of *AvrLm7*, it has been demonstrated that *AvrLm4* and *AvrLm7* are two distinct alleles of a single gene, renamed *AvrLm4-7*, with a single base mutation, leading to the change of a glycine to an arginine residue, being responsible for the loss of *AvrLm4* specificity whereas *AvrLm7* recognition is unaltered (Parlange *et al.*, 2009). *AvrLm9* has been found to be located at 22.6 cM of *AvrLm4-7*, but attempts to estimate the genetic distance between *AvrLm3* and *AvrLm4-7* failed since the interaction phenotypes toward the corresponding resistance genes *Rlm3*, *Rlm4* and *Rlm7* always co-segregated in *in vitro* crosses (Balesdent *et al.*, 2002). It was therefore questioned whether *AvrLm3* is distinct from, or an allelic form of *AvrLm4-7*.

In 2000 and 2001, a large-scale sampling of isolates was performed to establish the race structure of French *L. maculans* population (Balesdent *et al.*, 2006). This study revealed that only one isolate out of the 1797 isolates collected, isolate CB-M3.2, was virulent toward *Rlm7* (*avrLm7*, hereafter referred to as a7). Although the frequency of isolates avirulent toward *Rlm3* (*AvrLm3*, hereafter referred to as A3) had been found to be very low in France (<0.2%) and Europe (0%) (Balesdent *et al.*, 2006, Stachowiack *et al.*, 2006), isolate CB-M3.2 was

phenotyped as A3. This observation could be explained by three distinct hypotheses: the first one would be in favour of the appearance in an A3 isolate of a mutational event affecting *AvrLm7* function. Due to the very low proportion of A3 isolates in France, this hypothesis is unlikely. The second would be that of a foundation event, i.e the introduction of a  $\Delta$ A3 isolate from a region where the frequency of A3 isolates is higher than in France. For instance, the frequency of A3 isolates was found to be around 70% in Canada (Balesdent *et al.*, 2005) with values ranging from 0 to 100%, depending of the sampling location (Dilmaghani *et al.*, 2009). The third hypothesis would be that of a complex interplay between the *AvrLm4-7* and *AvrLm3* genes so that the mutated allele of *AvrLm4-7* in isolate CB-M3.2 restores a hidden functionality for *AvrLm3* allele. Such an interaction has been described between *Avr* genes *Avr1*, *Avr2* and *Avr3* of *Fusarium oxysporum* (Houterman *et al.*, 2008). In France, the *Rlm7* resistance gene has not been used in the field before 2002, when it was first released with the cv. Roxet. *Rlm7* has then more and more widely been used with the new hybrid varieties Exagone (2004), Exocet and Extend (2005). The introduction of a new resistance gene in a context where the frequency of the corresponding avirulent allele in pathogenic populations was higher than 99% gave us the unique opportunity to analyse the first molecular events that led to the gain of virulence in a real field situation (Balesdent *et al.*, 2011, Daverdin *et al.*, submitted). It was also the opportunity to investigate the relationship between *AvrLm3* and *AvrLm4-7* in order to determine which of the three above-mentioned hypotheses is correct.

In this paper, we report molecular and genetic evidences that *AvrLm3* and *AvrLm4-7* are indeed two distinct genes, with *AvrLm3* being genetically localised close to *AvrLm4-7*. We also demonstrate the conflicting relationship existing between these two genes, *AvrLm7* preventing the expression of *Rlm3* resistance in spite of the presence of *AvrLm3* in the genome. Finally, we report evidence showing that resurgence of *AvrLm3* in French *L.*

*maculans* populations is not due to genetic hitchhiking but is the result of independent events of loss of *AvrLm4-7* functionality.

## **Material and methods**

### Terminology

A specific terminology is employed hereafter to facilitate the distinction between isolate genotypes and phenotypes. Considering the interaction between the avirulence gene *AvrLmi* and the resistant gene *Rlmi*, the phenotype of an isolate is termed “Ai” or “ai”, where “Ai” corresponds to an avirulent phenotype and “ai” to a virulent phenotype. The genotype of an isolate with the functional allele of an avirulence gene ‘i’ is written *AvrLmi* whereas *avrLmi* is assigned to isolates which are harbouring a non-functional or a deleted allele. When genotypic or phenotypic virulence of several avirulence genes are considered, the corresponding abbreviations were added in a row: for instance the phenotype of an isolate virulent toward *Rlm4* and avirulent toward *Rlm7* is termed as a4A7 whereas its genotype is termed by *avrLm4-AvrLm7*.

When considering the complete set of *AvrLm* genes, the race nomenclature defined by Balesdent *et al.* (2005) was used; each *L. maculans* race is identified by the successive numbers of Avr loci for which the isolate has been characterized as avirulent, preceded by the letters ‘Av’. For instance, race Av1-2-4-7 is composed of isolates possessing the Avr alleles *AvrLm1*, *AvrLm2*, *AvrLm4* and *AvrLm7*. Numbers into brackets indicate Avr loci for which the genotyping has not been done.

### *L. maculans* isolates

Isolate collection from Daverdin *et al.* (Submitted) used here originated from an oilseed rape experimental field located in Grignon (France) where one cultivar harboring the resistance

gene *Rlm7* (Exagone) and one without *Rlm7* (Campala) were grown. *L. maculans* isolates were sampled during three continuous years on both cvs, starting in autumn 2006. This collection has been phenotyped for virulence toward *Rlm4* and *Rlm7* and the mutational events generating virulence at the *AvrLm7* locus has been determined (Daverdin, 2011, Daverdin *et al.*, submitted). From this collection, 1514 isolates (722 A7 and 792 a7) were phenotyped here for their virulence toward *Rlm3*.

G06-E101 and G06-E107 are two A3a7 isolates from this collection that were collected from leaf lesions in 2006 on cv. Exagone. *AvrLm4-7* sequencing showed these isolates display two different mutational events in the *AvrLm4-7* gene sequence: deletion of a dinucleotide in position 78-79 for G06-E101 and 16 point mutations caused by RIP for G06-E107. These mutational events introduce premature stop codons in the gene sequence, thereby explaining the virulence of each isolate toward *Rlm7*. These isolates were used in complementation experiments and for *in vitro* crosses. Despite their virulent phenotype, *AvrLm4-7* PCR amplification was still possible in these isolates and thus allowed us to follow *AvrLm4-7* segregation in the progeny of *in vitro* crosses.

Isolate v23.1.3, from which *AvrLm4-7* was cloned (Parlange *et al.*, 2009), was used in silencing experiments. The a3a4a7 isolate Nz-T4, in which *AvrLm4-7* is deleted, (Parlange *et al.*, 2009), was used for *in vitro* crosses. Isolate 19.4.24 (race Av3-5-6-8) and isolate v23.1.2 (race Av5-6-7-8) were used as A3a4a7 and a3a4A7 controls, respectively, in inoculation tests (Balesdent *et al.*, 2002).

### Inoculation tests

All isolates were phenotyped for their virulence on *Rlm3*, *Rlm4* and *Rlm7* using a cotyledon-inoculation test as follows. Spore suspensions of each isolate were inoculated on 10-12 plants of each of the *B. napus* cvs or lines Westar (no R gene), 01.23.2.1 (*Rlm7*), Pixel (*Rlm4*) and

03.22.3.1 (*Rlm3*) as described previously (Balesdent *et al.*, 2001, 2002). Symptoms were scored 12-21 days after inoculation using a 1 (avirulent) to 6 (virulent) rating scale. The isolates were classified as avirulent or virulent for a given locus whenever the mean disease rating on the corresponding differential line was  $\leq 3$  or  $> 3$ , respectively.

#### *In vitro* crosses

*In vitro* crosses were performed as described earlier (Gall *et al.*, 1994; Balesdent *et al.*, 2001). When pseudothecia had developed and ascospores started to be ejected, mature pseudothecia were stuck upside down on the top of a Petri dish containing 2% water-agar. One to two hours later, ejection of ascospores was checked using a stereo microscope (Olympus SZH10, 70X). Ungerminated, well separated ascospores were randomly selected and moved away from the others, a few millimetres from the ejection spot. After 24 h at room temperature, single germinated ascospores were transferred onto individual V8-agar Petri dishes. Segregation analysis of *AvrLm3* was performed using 2 distinct *in vitro* crosses (Table 1).

#### DNA manipulation

Genomic DNA was extracted from conidia suspensions using the DNeasy 96 Plant Kit and the QIAGEN BioRobot 3000 according to manufacturer's recommendations. Polymerase Chain Reactions (PCR) were performed in an Eppendorf Mastercycler EP Gradient thermocycler (Eppendorf, Le Pecq, France), as previously described (Fudal *et al.*, 2007). Gel electrophoresis of PCR products has been reported (Attard *et al.*, 2001). Sequencing was performed using a Beckman Coulter CEQ 8000 automated sequencer (Beckman Coulter, Fullerton, CA, USA) according to the manufacturer's instructions.

#### Genetic mapping

*AvrLm4-7* PCR amplification was used as a marker in crosses where it was polymorphic. Additional markers were selected from previous studies or developed using the v23.1.3 whole genome sequence (Rouxel *et al.*, 2011) (Figure 1; Table 2, Supplementary Table S1). From the molecular markers used by Parlange *et al.* (2009) for *AvrLm4-7* map-based cloning, one minisatellite (*MinLm9*), physically close to *AvrLm4-7*, and three Cleaved Amplified Polymorphic Sequence (CAPS) markers were found to be polymorphic in cross #69 (Figure 1; Table 2). CAPS markers were first amplified (Table 2) and the PCR products were digested with *AvaII*, *DdeI* and *BglII* for markers 46F10\_R, 133E3\_R and 124F12\_F, respectively. Five additional minisatellite markers were designed using the FONZIE pipeline (Bally *et al.*, 2010), one of them, Min12-96, being polymorphic in cross #69. Nineteen single-locus PCR markers spanning two truncated copies of degenerated transposons were generated as described by Gout *et al.* (2006b) and using the annotation of repeated elements described in Rouxel *et al.* (2011) (Figure 1; Table 2, Supplementary Table S1). Twenty predicted genes present in the GC-rich isochore GC-13 (Figure 1) were amplified to look for presence/absence or size polymorphism and five of them were partly sequenced in the two parental isolates of cross #69 to look for potential SNP markers (Supplementary Table S2) but none of these approaches generated additional markers. All primers were generated using Primer3 (Rozen & Skaletsky, 2000).

Linkage analyses among loci were performed using the MapMaker/EXP version 3.0b software (available online; Whitehead/MIT Center for Genome Research, Cambridge, MA, USA) with a LOD score of 3.0 and a maximum recombination frequency of 30 centimorgans (Kuhn *et al.*, 2006). Linkage groups were drawn using Mapchart 2.2 (Voorrips, 2002). Position of the markers and of the repeated elements along the Super-Contig, and the GC content along the sequence were visualized using Artemis v12 (Rutherford *et al.*, 2000)

### **RNA extraction, reverse transcription and qRT-PCR analysis.**

Total RNA was extracted from *B. napus* cv. Westar cotyledons 14 days after infection using the TRIzol reagent (Invitrogen, Cergy Pontoise, France) according to manufacturer's protocol. Total RNA was treated with DNase I RNase-Free (New England Biolabs, Ipswich, MA, U.S.A.). All samples were adjusted to 3 µg of RNA and single-strand cDNA was generated using oligo-dT primed reverse transcription with PrimerScript Reverse Transcriptase (Clontech, Palo Alto, CA, U.S.A.) according to the manufacturer's protocol.

Quantitative RT-PCR was performed using model 7700 real-time PCR equipment (Applied Biosystems, Foster City, CA, U.S.A.) and ABsolute SYBR Green ROX dUTP Mix (ABgene, Courtaboeuf, France) as described by Fudal *et al.* (2007). For each isolate tested, two technical replicates from two biological repeats were performed. Water was used as negative control. Primers used to amplify *AvrLm4-7* mRNA were described in Parlange *et al.* (2009) and those used to amplify *actin* mRNA by Fudal *et al.* (2007). Samples were analyzed from a 10-fold dilution of the original RT products as described by Fudal *et al.* (2007). The fluorescence threshold (Ct) value was determined at 0.1 of fluorescence intensity. Ct values were analyzed according to the  $2^{-\Delta\Delta Ct}$  method (Livak & Schmittgen, 2001). *Actin* was used as a constitutive reference gene and results were reported to the Ct value of the wild-type isolate v23.1.3.

### Transformation experiments

*L. maculans* transformation using *Agrobacterium tumefaciens* was performed as described by Gout *et al.* (2006). Fungal transformants were selected on 50 µg ml<sup>-1</sup> nourseothricin (WERNER BioAgents, Jena, Germany). Vectors and isolates used for transformations are described in Table 3. Two constructs from Parlange *et al.* (2009) have been used in complementation experiments. Both constructs are based on the vector pPZPnat1 and differ

for the allele of *AvrLm4-7* they carry. The first one termed pPZPnat1-*A4A7* carries v23.1.3 wild-type *AvrLm4-7* allele (*AvrLm4-AvrLm7*) whereas pPZPnat1-*a4A7* carries the wild-type v23.1.2 allele (*avrLm4-AvrLm7*) (Table 3).

Vector pPZPnat1-sil*AvrLm4-7* was constructed for RNA-mediated gene silencing of *AvrLm4-7*, using the method described by Fudal *et al.* (2007). Vector pJK11 (a gift from K. Plummer, Horticulture and Food Research Institute of New Zealand) was used for the construction of inverted repeats of the target gene *AvrLm4-7*. pJK11 consists of a backbone of pBluescript II containing the *Glomerella cingulata gpdA* promoter fragment and the *Aspergillus nidulans trpC* terminator fragment, separated by unique *XmaI* and *BamHI* restriction sites. The hairpin inverted repeat fragments were created from the coding region of *AvrLm4-7* by PCR, using cDNA of v23.1.3 as a template. cDNA were obtained by reverse transcription of RNA extracted from infected cotyledons of oilseed rape cv. Westar at seven days post inoculation. The *AvrLm4-7* sense fragment of the inverted repeat was amplified using primers sil*AvrLm4-7-XmaI-F* (GAGAGACCCGGGCTTCCACCAAGTATAAACCC) and sil*AvrLm4-7-AscI-R* (GAGAGAGGCGCGCCTAGTCGCAACCACGAGTCCT). The antisense region was amplified with primers sil*AvrLm4-7-BamHI-F* (GAGAGAGGATCCCTTCCACCAAGTATAAACCC) and sil*AvrLm4-7-AscI-R* (GAGAGAGGCGCGCCCAGGCACGTTGATGCGATTG). Sense and antisense fragments were ligated independently into a pGEMt-easy vector (Promega) and inserts were sequenced for verification of correct PCR amplification. Sense and antisense inserts were respectively excised by digestion with *XmaI* and *AscI* or *AscI* and *BamHI* and ligated into an *XmaI-BamHI*-digested pJK11. The resulting plasmid was named pJK11-sil*AvrLm4-7*. In contrast to Fudal *et al.* (2007) the expression cassette could not be excised by digestion with *SpeI* and *XhoI* because of the presence of an *XhoI* restriction site in the amplified sense and antisense regions of *AvrLm4-7*. The *gdpA* promoter of pJK11 was therefore excised by digestion with



*XhoI* and *HindIII* and this fragment was ligated into the polylinker of an *XhoI-HindIII*-digested pPZPnat1 vector. This plasmid was named pPZPnat1-*gdpA*. The hairpin inverted repeat fragments of *AvrLm4-7* and the *trpC* terminator fragments were excised from pJK11-*silAvrLm4-7* by digestion with *SmaI* (cutting blunt at the same site as *XmaI*) and *SpeI*. This fragment was ligated into pPZPnat1-*gdpA* which had been previously digested with *HpaI* (cutting blunt downstream of *gdpA*) and *SpeI*, creating plasmid pPZPnat1-*silAvrLm4-7*. This plasmid was introduced into *Agrobacterium tumefaciens* C58-pGV2260 by electroporation at 2.5 kV, 200 Ohms and 25  $\mu$ F.

#### Data analyses

Statistical analyses were performed using XLSTAT v2010.5.01. Mainly, two populations were compared for their frequencies of each type of molecular events in the *AvrLm4-7* gene using non-parametric tests, i.e. either Approximate Pearson's Chi squared test or Fisher's exact test.

## **Results**

### Reappraisal of phenotypic data from the literature

Previous attempts to determine whether *AvrLm3* and *AvrLm4-7* were allelic failed, but these studies relied on a limited number of progeny analysed (i.e, less than 40 in Balesdent *et al.*, 2002). Wide scale phenotyping of field populations for their avirulence toward *Rlm3*, *Rlm4* and *Rlm7* was undertaken as an alternative to genetic studies. To be informative, these analyses have to include significant amounts of virulent and avirulent isolates toward *Rlm7* and information about virulence toward *Rlm3*. Thus, three studies have been selected in the literature and re-analysed specifically for *AvrLm3* and *AvrLm7* virulence (Balesdent *et al.*, 2005, Dilmaghani *et al.*, 2009, Kutcher *et al.*, 2010). From the information on race structure described in these studies, we deduced the number of isolates belonging to each of the following classes: avirulent toward both *Rlm3* and *Rlm7* (A3A7), avirulent toward *Rlm3* but virulent toward *Rlm7* (A3a7), virulent toward *Rlm3* but avirulent toward *Rlm7* (a3A7) and virulent toward both *Rlm3* and *Rlm7* (a3a7) (Table 4).

In two of the three studies, the four phenotypic classes were found, suggesting for the first time that *AvrLm3* and *AvrLm4-7* were probably distinct genes. The frequency of isolates in each phenotypic class was not equilibrated. Phenotypic classes a3A7 and A3a7 represented nearly 90% of the isolates whereas on the contrary, A3A7 isolates were especially uncommon with only 1% of the isolates in this class (Table 4).

### Characterization of a French field population submitted to *Rlm7* selection pressure

Virulence toward *Rlm3* and *Rlm7* of the 1514 isolates of the collection established by Daverdin (Daverdin 2011; Daverdin *et al.*, submitted) was determined. Isolates were distributed in the four phenotypic classes with an unequal distribution, similar to that found in the above-mentioned studies (Table 4). A3a7 and a3A7 phenotypic classes encompassed

almost 98 % of the isolates whereas A3A7 and a3a7 isolates represented a very limited proportion of the collection with about 1 % of the isolates for each class (Table 4). Therefore, for a majority of the isolates collected in a single field submitted to *Rlm7* selection pressure, the avirulent phenotypes toward *Rlm3* and *Rlm7* were linked in repulsion, but a few recombinant phenotypes were nevertheless observed. These data suggest a tight genetic or physical linkage, and/or an antagonist interaction between *AvrLm3* and *AvrLm7*. To explain why, for a majority of isolates, avirulence toward *Rlm3* correlates with virulence toward *Rlm7*, we hypothesised that *AvrLm3* and *AvrLm4-7* are two distinct linked genes, and that the *AvrLm3*-mediated *Rlm3* response is prevented by the presence of a functional allele of *AvrLm4-7*.

#### Comparison of molecular events explaining virulence at the *AvrLm7* locus between A3 and a3 isolates

In Daverdin *et al.* (submitted), the mutational events at the origin of the virulence toward *Rlm7* were determined following *AvrLm4-7* PCR amplification and sequencing in 735 of the 789 a7 isolates for which the phenotype at the *AvrLm3* locus could be determined. Virulence toward *Rlm7* can be due to complete or partial deletion of the gene, to inactivation due to stop codons generated by RIP mutation, to single or di-nucleotide deletion in the gene leading to frameshifts generating stop codons, or to single nucleotide mutations (SNP) that may lead either to stop codons in the sequence or to changes in the amino acid sequence of the protein (Daverdin, 2011; Table 5). All categories of molecular events affecting *AvrLm4-7* were identified for A3a7 isolates and only three for the a3a7 isolates (Table 5). An Approximate Pearson's Chi-Squared Test with 20,000 random samplings was used to test the null hypothesis, i.e. that the type of events were similarly distributed in the two populations (A3a7 or a3a7). This test showed that the two populations differed significantly, at a 5% level, (chi-

squared = 20.045,  $P$ -value = 0.037, 99% confidence interval for the  $P$ -value [0.033, 0.040]). This difference was mainly due to the difference in frequency of the SNP category; 30% of a3a7 isolates have SNP in *AvrLm4-7* compared to only 3.7 % for A3a7 isolates (Table 5). Taking into consideration the SNP category of events only, the A3a7 and a3a7 populations differed significantly ( $p$ -value [Fisher exact test] = 0.006). Seven distinct SNP in *AvrLm4-7* have been described in the *avrLm7* isolates (Daverdin, 2011). Four of them are considered as having a major impact in the peptide produced; two of them introduce a premature stop codon after 35 or 55 amino acids, respectively, and the two others change a cysteine into an arginine or a serine, respectively, with probable consequences on the 3D structure of the protein (Daverdin, 2011; Fudal, Blaise and van Tilbeurgh, pers. comm.). These types of SNP were only found in A3a7 isolates (Table 5). The three other SNP identified have less obvious consequences on the protein. The first one (R100 → P100) was shared by A3a7 and a3a7 isolates (15 and 1 isolate, respectively); the second, F102 → S102, was found only once in an A3a7 isolate, and the last one, S112 → R112, was only found in two a3a7 isolates. The distribution of SNP type differed significantly between a3a7 and A3a7 ( $p$ -value [Fisher exact test] = 0.018).

#### Complementation of *AvrLm3-avrLm7* isolates with functional alleles of *AvrLm4-7*

Complementation of two *avrLm7* isolates, G06-E101 and G06-E107, both displaying an A3 phenotype, was performed with functional *AvrLm4-7* alleles. Each isolate was complemented with either the *avrLm4-AvrLm7* or the *AvrLm4-AvrLm7* allele of *AvrLm4-7*. All transformants were virulent on the susceptible cv. Westar (data not shown) and the integration of a functional *AvrLm7* allele restoring the avirulent phenotype toward *Rlm7* was confirmed for all isolates but two following inoculation on the *Rlm7* line (Figure 2). Inoculation of the

transformants on the *Rlm3* lines showed that the presence of a functional copy of *AvrLm4-7*, regardless of its allele, resulted in the gain of virulence toward *Rlm3* (Figure 2).

#### RNAi silencing of *AvrLm4-7* restores the *AvrLm3* specificity

RNAi silencing experiments aimed to confirm the existence of a functional interplay between *AvrLm4-7* and *AvrLm3*, but also to identify whether the functional allele of *AvrLm3* was present in the sequenced isolate v23.1.3, in spite of its a3 phenotype.

The pPZPnat1-sil*AvrLm4-7* vector, containing a fragment of *AvrLm4-7* cDNA in sense and antisense orientation, was transformed into v23.1.3 (a3A4A7). Twenty nine transformants were obtained and analyzed for their interaction phenotypes on *Rlm4*, *Rlm7* and *Rlm3*. The integration of the construction modified the interaction phenotypes only for one transformant (transformant T4, Figure 3). This transformant displayed an A3a4a7 phenotype thereby not only showing that the functional allele of *AvrLm3* is present in v23.1.3, but also confirming the functional antagonistic interplay between *AvrLm4-7* and *AvrLm3*. One transformant, T24, displayed an avirulent phenotype toward some *Rlm3* plants, leading to an intermediate mean disease rating on *Rlm3*, but without being virulent on *Rlm4* and *Rlm7* plants (Figure 3).

Quantitative RT-PCR analysis of the transformants whose virulence toward *Rlm3* was affected (T4, T24) showed that their expression level of *AvrLm4-7* is reduced. Indeed, T4 exhibited a low expression level of *AvrLm4-7* (less than 10% compared to v23.1.3) whereas T24 exhibited an intermediate expression level of *AvrLm4-7* (around 50%). These results thus suggest that the recovery of the A3 phenotype does not require a silencing level as high as that necessary for displaying the a4a7 phenotype.

#### Occurrence of avirulent progeny isolates (A3) following an *in vitro* cross between two virulent isolates (a3) toward *Rlm3*

Cross #33 (Nz-T4 x V23.1.11) was previously established for the map-based cloning of *AvrLm4-7* (Parlange *et al.*, 2009). Parental isolates of cross #33 are both virulent toward *Rlm3*. Characterisation of their progeny for virulence toward *Rlm3* revealed here the presence of five A3 isolates out of 238 (Table 1). This observation, combined to complementation and silencing experiment results, suggests that, despite their virulent phenotype toward *Rlm3*, at least one of the parental isolates of cross #33 must have an avirulent allele of *AvrLm3*, but that the presence of *AvrLm4-7* prevents it to generate the *AvrLm3*-mediated *Rlm3* resistance response. Two types of virulent isolates toward *Rlm3* (a3) are thus suggested to exist: “true” virulent isolates that do not possess a functional allele of *AvrLm3* (*avrLm3-AvrLm7* isolates) and virulent isolates possessing a functional allele of *AvrLm3*, whose recognition by *Rlm3* is impeded by the presence of a functional *AvrLm4-7* allele (genotype, *AvrLm3-AvrLm7*; phenotype, a3A7). Considering this hypothesis, and knowing that Nz-T4 lacks *AvrLm4-7* (Parlange *et al.*, 2009), it is thus suggested that Nz-T4 (a3a7) correspond to a “true” a3 isolate, and that the second parental isolate of cross #33, v23.1.11, possesses both *AvrLm3* and *AvrLm4-7*, leading to the a3A7 phenotype (Table 1). The five isolates in the progeny with the A3a7 phenotype therefore should correspond to recombinant isolates harbouring *AvrLm3* (from v23.1.11) but devoid of *AvrLm4-7* (from Nz-T4). The limited number of recombinant isolates indicated that *AvrLm4-7* and *AvrLm3* are genetically linked.

#### Genetic control of avirulence toward *Rlm3*

The *in vitro* cross #69 was performed between G06-E107 (phenotype A3a7, putative genotype *AvrLm3-avrLm7*) and Nz-T4 (phenotype a3a7, putative genotype *avrLm3-avrLm7*). In this cross, the segregation of *AvrLm3* could be analysed independently from that of *AvrLm7*. One hundred and thirty progeny were phenotyped for avirulence toward *Rlm3* and the 65:60 avirulent:virulent segregation ratio confirmed that A3 phenotype was under a single

gene control (Tables 1, 6) ( $\text{Chi}^2 = 0.199$ , p-value = 0.654). Nz-T4 lacks *AvrLm4-7* whereas G06-E107 possesses a mutated *AvrLm4-7* allele whose PCR amplification is possible. PCR analysis of *AvrLm4-7* segregation revealed that recombination occurred between *AvrLm4-7* and *AvrLm3*, proving that *AvrLm3* and *AvrLm4-7* are not allelic (Table 6). The low number of recombinant isolates compared to parental types confirmed the tight genetic linkage between *AvrLm3* and *AvrLm4-7*.

### Genetic mapping

Progeny of cross #69 was used to initiate a genetic mapping and map-based cloning of *AvrLm3*. The polymorphism of 35 potential molecular markers along Super-Contig (SC) 12 was evaluated between parental isolates of cross #69. Ten polymorphic markers were identified and used to construct a genetic map of the *AvrLm3-AvrLm4-7* region (Figure 1). Four of these markers were designed at the junction of two repetitive elements (SC12\_2, SC12\_14, SC12\_16 and SC12-19 in Figure 1) in the last AT-isochore of SC\_12, named AT\_13. A linkage group (LG) containing eight molecular markers along with *AvrLm3* was obtained, and marker order in this linkage group was consistent with the physical map deduced from the genome sequence of the v23.1.3 isolate (Rouxel *et al.*, 2011) (Figure 1). Genetic mapping indicated that *AvrLm3* is located in the 3' region of AT\_13 (Figure 1). AT\_13, sized 162 kb, encompassed the typical mosaics of RIP degenerated and truncated transposable elements making up AT-isochores in *L. maculans*. The 3' part of AT\_13 comprised 50 kb of degenerated and truncated copies of the telomere-associated repeat *LmTelo2*, containing a RecQ helicase and typical of the most proximal regions of telomeric extremities of *L. maculans* chromosomes (Figure 1). Markers SC\_2, SC\_14 and SC\_16 covered a genetic distance of 4.1cM for a physical distance of 125 kb. The next marker, SC12-19, located 13 kb downstream from SC12-16, was neither linked to *AvrLm3* nor to

other markers of the LG (Figure 1). Finally a minisatellite marker, min12-96 located 16.8 kb from SC12-19 and at 6.2 kb of the 3' end of SC12, was also found to be unlinked to SC12-19 and other markers of the region.

## Discussion

The avirulence gene *AvrLm4-7* confers a double specificity of interaction, toward *Rlm4* and *Rlm7*, the mutation of a single amino acid being sufficient to escape *Rlm4*-mediated recognition of *AvrLm4-7* (Parlange *et al.*, 2009). Co-segregation of *AvrLm3* and *AvrLm4-7* (Balesdent *et al.*, 2002) suggested *AvrLm3* could be a potential novel allele of *AvrLm4-7*. In this study we show for the first time that the avirulent phenotype toward *Rlm3* is actually governed by a single gene, *AvrLm3*, distinct from *AvrLm4-7*, located in a telomeric or subtelomeric region, and we demonstrate by different approaches that the presence of a functional allele of *AvrLm4-7* prevents *Rlm3*-mediated recognition of *AvrLm3*.

The analysis of previously published data regarding interaction phenotypes in field isolates, along with the phenotyping of a large collection of isolates recently obtained in one field submitted to the *Rlm7* selection pressure in France, revealed both a strong negative correlation between the avirulent phenotypes toward *Rlm3* and *Rlm7*, and the occurrence of *ca* 2 % of isolates being either virulent or avirulent toward both genes. These data suggested that *AvrLm3* and *AvrLm4-7* were probably two distinct genes. By formal genetic approaches using a cross in which *AvrLm4-7* did not segregate (cross #69), and a large progeny of a cross in which A3 isolates were recovered although both parents were a3 (cross #33) we confirmed that *AvrLm3* was distinct from, but linked to, *AvrLm4-7*.

Three avirulence genes of *L. maculans*, *AvrLm1*, *AvrLm6* and *AvrLm4-7*, were previously cloned using a map-based strategy. These three genes are located within AT-rich isochores of



*L. maculans*, mainly composed of RIP-degenerated copies of repeated elements. This context complicated the chromosome walk toward the genes, but the polymorphism of the mosaics, or of the sequence, of repeats, along with the availability of large genome sequence data spanning the genetic interval containing the genes, nevertheless allowed successful cloning of the three genes (Gout *et al.*, 2006b; Fudal *et al.*, 2007; Parlange *et al.*, 2009). Therefore, map-based cloning of *AvrLm3* was undertaken, taking advantage (i) of the recent release of the *L. maculans* genome sequence, including an exhaustive repertory of its repeated elements (Rouxel *et al.*, 2011), (ii) of the availability of molecular markers generated during the chromosome walk toward *AvrLm4-7* (Parlange *et al.*, 2009) and (iii) of the availability of the progeny of cross #69, designed to allow the phenotyping of the progeny for the *AvrLm3* / *Rlm3* interaction without interference with the *AvrLm3* / *AvrLm4-7* interaction (see below). Unfortunately the map-based cloning of *AvrLm3* failed, since it was not possible to generate markers that could border *AvrLm3* on its 3' side from the available sequence assembly of the region, despite the great number of potential markers generated. The last marker linked to *AvrLm3* in its 5' side, marker SC12\_16, was generated on the basis of a complex truncation of three telomeric repeats of the *LmTelo2* family. *AvrLm3* being still located 7.8 cM away from it (Figure 1), it is more than likely that *AvrLm3* is located in the distal part of the telomere, wrongly assembled to an unrelated telomeric region. Incompletely assembled telomeric regions are common in genome assembly and only a few telomeres of *L. maculans* were found to contain the most distal *LmTelo1* and RPP\_ *Circe* mosaics (Rouxel *et al.*, 2011). A systematic generation of markers in *L. maculans* genomic regions containing telomeric repeats, with the same strategy as that developed here, and their mapping in cross #69 and in the reference genetic map (Kuhn *et al.*, 2006), will hopefully allow the identification of the correct 3' part of SC12 where *AvrLm3* should be looked for.

In spite of the failure to clone *AvrLm3*, we demonstrated here a negative interaction between *AvrLm3* and *AvrLm4-7*. This was achieved with three complementary approaches: (i) RNA silencing of *AvrLm4-7* in the a3A7 sequenced isolate v23.1.3 restored an A3 phenotype in the only transformant that had lost ability to induce the *Rlm7*-mediated resistance. This result indicated that *AvrLm3* was indeed present in the sequenced isolate v23.1.3, and suggested that its recognition was hidden by *AvrLm4-7*. This negative interaction was also supported by (ii) the diversity and the type of mutational events found in *AvrLm4-7* in virulent a7 isolates recovered from a single experimental field. In 98.6% of the cases, these a7 isolates were also A3, and for 97 % of A3a7 isolates, the mutations in *AvrLm4-7* had drastic effects on the length or structure of the *AvrLm4-7* protein. This suggests that resurgence of A3 isolates in this field is correlated to the loss of function of *AvrLm4-7* in most isolates. Finally (iii) complementation of two A3a7 isolates, both displaying drastic mutations in *AvrLm4-7*, with a functional *AvrLm4-7* allele, induced the loss of the *Rlm3*-mediated recognition of *AvrLm3*. All these data clearly support the hypothesis of a negative interaction between *AvrLm3* and *AvrLm4-7*, with *AvrLm4-7*, when present and functional, inhibiting the *Rlm3*-mediated recognition of *AvrLm3*. This is a new illustration of the complex arms race co-evolution of plant and pathogens. Lawrence *et al.* (1981) demonstrated by formal genetics the occurrence of a pair of dominant genes, an avirulence gene (A) and an inhibitory gene (I), in *Melampsora lini*; in this diploid organism, only isolates with *iiAa* or *iiAA* genotypes are avirulent on the corresponding resistance source. Houterman *et al.* (2008) also demonstrated that the *Avr1* effector of *F. oxysporum*, inducing *I-1* triggered resistance, also suppresses the protective effect of two other tomato resistance genes, *I-2* and *I-3*. Similarly *AvrLm4-7* and *AvrLm3* are parts of the arms race between *L. maculans* and *B. napus*, with two distinct strategies to escape recognition by the resistance genes. While the *Rlm7*-mediated recognition is prevented by a range of mutational events leading in most cases to the lack of, or altered production of

the protein (Daverdin, 2011), the strategy to escape the *Rlm3*-mediated resistance does not affect *AvrLm3*, but implies the presence of *AvrLm4-7*, acting as an ‘inhibitor’ *sensus* Lauwrence *et al.* (1981) or a suppressor of plant resistance gene-based immunity as proposed by Houterman *et al.* (2008).

To explain how *AvrLm4-7* prevents *AvrLm3* recognition, several hypotheses can be proposed. Firstly, *AvrLm4-7* could directly act on *AvrLm3* integrity, and hence suppress a recognition site between *AvrLm3* and *Rlm3* or a related guard protein. This model could be extended with *AvrLm3* being the target of the protein *Rlm7* itself, after being modified by *AvrLm7*. Secondly, the *AvrLm3* and *AvrLm4-7* proteins could have the same host target but may act differently on it. This case has been described in the interaction between *Arabidopsis thaliana* and the bacteria *Pseudomonas syringae* (Axtell *et al.*, 2003). *AvrRpt2* acts by cleaving the plant protein *RIN4*, which triggers *RPS2*-mediated resistance, but the action of *AvrRpt2* on *RIN4* also prevents the phosphorylation of *RIN4* by *AvrRpm1*, which is a prerequisite to trigger *RPM1*-mediated resistance. Alternatively, *AvrLm3* and *AvrLm4-7* could still have the same target, but they could differ in their expression level, their kinetics of expression or their affinity to the target. This last hypothesis would be in agreement with our silencing experiment results showing that, with an intermediate level of *AvrLm4-7* silencing, an A3A7 isolate has been generated, which still induces the *Rlm7* mediated resistance whereas its virulence toward *Rlm3* is not fully restored. The few A3A7 isolates that have been identified at a very low frequency in our collection of isolates may also present a specific regulation of *AvrLm3* and *AvrLm7*, however such a hypothesis can not be checked until *AvrLm3* has been cloned.

With the identification in 2000 of the first French virulent isolate toward *Rlm7* being A3, it was questioned whether this isolate could have been introduced from a region with higher frequency of *AvrLm3* than that found in Europe. The high number of mutational events explaining the acquisition of virulence toward *Rlm7* along with the demonstration that the loss of *AvrLm4-7* function restores the A3 phenotype in most field isolates analysed clearly allow us to reject the foundation event hypothesis. In European populations of *L. maculans*, the *AvrLm3* allele was thought to be very rare (Balesdent *et al.*, 2006, Stachowiak *et al.*, 2006). However, the analysis of a field population submitted to the *Rlm7* selection pressure revealed that the majority of isolates that became a7 because of a loss of functionality of *AvrLm4-7* were indeed avirulent toward *Rlm3*. It is therefore likely that the frequency of *AvrLm3* allele was very high in the *L. maculans* population present in the region where our experimental field was located. In the absence of the cloned *AvrLm3* gene, it is difficult to confirm this local observation on a wider scale, but the lack of population structure between fields, the very high gene flow between French populations (Gout *et al.*, 2006a) and the lack of race structure at the European scale (Stachowiak *et al.*, 2006) suggest the frequency of *AvrLm3* in Europe could be very high. The phenotyping on the *Rlm3* line of virulent a7 populations that will probably appear with the extensive use of *Rlm7* will allow us to determine more precisely the frequency of *AvrLm3* in France.

The maintenance of *AvrLm3* at high frequencies in populations seems contradictory with its location in a telomeric region. The location of *AvrLm3* in a telomeric or a sub-telomeric region is consistent with the assumption that *L. maculans* effectors are located within RIP-affected AT-rich isochores (Rouxel *et al.*, 2011) but such a telomeric location was not reported yet for *L. maculans* avirulence genes. Telomeres are often enriched in genes involved in niche adaptation in eukaryotes (Farman, 2007). For instance, many avirulence genes, including *AVR-Pii* and *AVR-Pita*, have been found in telomeric regions in *Magnaporthe*

*oryzae*, and loss of chromosome tips is suggested to be one frequent mechanism for gain of virulence (Orbach *et al.*, 2000; Farman, 2007, Yoshida *et al.*, 2009). Maintenance of *AvrLm3* in populations in spite of its telomeric location suggests *AvrLm3* is likely to have an important role in the pathogenic or life cycle of the fungus, so that loss of *AvrLm3* would lead to a serious fitness penalty. Such a fitness penalty is known for *Avr3* of *F. oxysporum*, escaping to its recognition by *I-3* through acquisition of *AVR1*, which suppresses the *AVR3-I-3* interaction, and not through loss of *AVR3* (Houterman *et al.*, 2008). As discussed by Houterman *et al.*, (2008), such a complex interplay may lead to initially unpredicted strategies for disease control. Until now, the use of specific *R* genes to control *L. maculans* has resulted in very rapid breakdown of these resistances, as shown for *Rlm1*, *Rlm6* or the *B. rapa* resistance source *RlmS* (Sprague *et al.*, 2006). Experimental trials also revealed that the *Rlm7* resistance can be rapidly overcome whenever agronomic practices favour *L. maculans* sexual reproduction and increase population size (Daverdin *et al.*, submitted). The increase of A3 isolates following *Rlm7* selection pressure can be exploited with either a pyramiding strategy combining *Rlm3* and *Rlm7* in the same genotype, or by alternation strategies where *Rlm7* and *Rlm3* genotypes could successively be used so that the few a7 isolates selected by the *Rlm7* selection pressure would immediately be controlled the next growing season by *Rlm3*. Because of the scarcity of isolates displaying an A3 phenotype in Europe (Stachowiak *et al.*, 2006), *Rlm3* was thought to be useless to control the disease, but it is still found in around 20% of cultivated *B. napus* genotypes (Balesdent & Pinochet, unpublished data). In the light of our results, the presence of *Rlm3* in *B. napus* varieties will now have to be taken into account in varieties management schemes. Even though the never-ending arms race will probably lead to the selection of a3a7 *L. maculans* isolates with increased fitness or with mutations in the *AvrLm3* gene modifying its recognition function but not its biological

function, such strategies will increase the durability of *Rlm7* and hence contribute to a drastic reduction of the *L. maculans* population size.

### **Acknowledgments**

The authors wish to thank Laurent Coudard (INRA BIOGER) for his major contribution to isolate collection management, Hubert and Martin Willigsecker (INRA BIOGER) for plant management. This work was funded by the ANR project CEDRE (ANR-PADD-005), the CASDAR Project “BioViresdur”, the CETIOM (Centre Technique Interprofessionnel des Oléagineux Métropolitains) and a grant from the Santé des Plantes et Environnement (SPE) INRA department. GD was funded by the French Ministère de la Recherche et de l’Enseignement Supérieur.

### **References**

- Attard A, Gourgues M, Gout L, Schmit J, Roux J, Narcy JP, Balesdent MH, Rouxel T. 2001.** Molecular characterisation and polymorphism of *MinLm1*, a minisatellite from the phytopathogenic ascomycete *Leptosphaeria maculans*. *Current Genetics* **40**: 54-64.
- Axtell MJ, Staskawicz BJ. 2003.** Initiation of RPS2-specified disease resistance in *Arabidopsis* is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**: 369-377.
- Balesdent MH, Attard A, Ansan-Melayah D, Delourme R, Renard M, Rouxel T. 2001.** Genetic control and host range of avirulence toward *Brassica napus* cultivars Quinta and Jet Neuf in *Leptosphaeria maculans*. *Phytopathology* **91**: 70-76.
- Balesdent MH, Attard A, Kuhn ML, Rouxel T. 2002.** New avirulence genes in the phytopathogenic fungus *Leptosphaeria maculans*. *Phytopathology* **92**: 1122-1133.

- Balesdent MH, Barbetti MJ, Li H, Sivasithamparam K, Gout L, Rouxel T. 2005.** Analysis of *Leptosphaeria maculans* race structure in a worldwide collection of isolates. *Phytopathology* **95**: 1061-1071.
- Balesdent MH, Louvard K, Pinochet X, Rouxel T. 2006.** A large scale survey of races of *Leptosphaeria maculans* occurring on oilseed rape in France. *European Journal of Plant Pathology* **114**: 53-65.
- Balesdent MH, Daverdin G, Carpezat J, Rouxel T, Pinochet X, Leflon M. 2011.** *AvrLm7* allelic diversity in *Leptosphaeria maculans* nevertheless allows survey of virulent isolates using the HRM (High Melting Resolution) method. *Proceedings of the 13<sup>th</sup> International rapeseed congress, Prague 5-9 June 2011*.
- Bally P, Grandaubert J, Rouxel T, Balesdent MH. 2010.** FONZIE: An optimized pipeline for minisatellite marker discovery and primer design departing from large sequence data sets. *BMC Research Notes* **3**: 322.
- Daverdin G. 2011.** Évolution moléculaire sous pression de sélection et implication dans la reconnaissance AvrLm3/Rlm3 du gène d'avirulence *AvrLm4-7* chez *Leptosphaeria maculans*. PhD thesis, Paris-Sud XI University, Orsay, France
- Daverdin G, Aubertot JN, Gout L, Fudal I, Carpezat J, Rouxel T, Balesdent MH.** Experimental evolution of the *AvrLm4-7* gene of *Leptosphaeria maculans* under *Rlm7* resistance gene selection pressure in the field. *Plos Pathogen* submitted
- Delourme R, Chèvre AM, Brun H, Rouxel T, Balesdent MH, Dias JS, Salisbury P, Renard M, Rimmer SR. 2006.** Major gene and polygenic resistance to *Leptosphaeria maculans* in oilseed rape (*Brassica napus*). *European Journal of Plant Pathology* **114**: 41-52.
- Dilmaghani A, Balesdent MH, Didier JP, Wu C, Davey J, Barbetti MJ, Hua Li, Moreno-Rico O, Phillips D, Despeghel JP, Vincenot L, Gout L, Rouxel T. 2009.** The

*Leptosphaeria maculans*–*Leptosphaeria biglobosa* species complex in the American continent. *Plant Pathology* **58**: 1044-1058.

**Farman MK. 2007.** Telomeres in the rice blast fungus *Magnaporthe oryzae*: the world of the end as we know it. *FEMS Microbiology Letters* **273**: 125-132

**Fitt BDL, Brun H, Barbetti MJ, Rimmer SR. 2006.** World-wide importance of phoma stem canker (*Leptosphaeria maculans* and *L. biglobosa*) on oilseed rape (*Brassica napus*). *European Journal of Plant Pathology* **114**: 3-15.

**Fudal I, Ross S, Gout L, Blaise F, Kuhn ML, Eckert M, Cattolico L, Balesdent MH, Rouxel T. 2007.** Heterochromatin-like regions as ecological niches for avirulence genes in *Leptosphaeria maculans* genome: map-based cloning of *AvrLm6*. *Molecular Plant-Microbe Interaction* **20**: 459-470.

**Gall C, Balesdent MH, Robin P, Rouxel T. 1994.** Tetrad analysis of acid phosphatase, soluble protein patterns, and mating type in *Leptosphaeria maculans*. *Phytopathology* **84**: 1299-1305.

**Gout L, Eckert M, Rouxel T, Balesdent MH. 2006a.** Genetic variability and distribution of mating type alleles in field populations of *Leptosphaeria maculans* from France. *Applied and Environmental Microbiology* **72**: 185-191

**Gout L, Fudal I, Kuhn ML, Blaise F, Cattolico L, Balesdent MH, Rouxel T. 2006b.** Lost in the middle of nowhere: the *AvrLm1* avirulence gene of *Leptosphaeria maculans*. *Molecular Microbiology* **60**: 67-80.

**Houterman PM, Cornelissen BJC, Rep M. 2008.** Suppression of plant resistance gene-based immunity by a fungal effector. *PLoS Pathogen* **4**: e1000061. doi:10.1371/journal.ppat.1000061.



- Huang YJ, Balesdent MH , Li ZQ, Evans N, Rouxel T, Fitt BDL. 2010.** Fitness cost of virulence differs between the AvrLm1 and AvrLm4 loci in *Leptosphaeria maculans* (phoma stem canker of oilseed rape). *European Journal of Plant Pathology* **126**: 279–291.
- Kuhn ML, Gout L, Howlett BJ, Melayah D, Meyer M, Balesdent MH, Rouxel T. 2006.** Genetic linkage maps and genomic organization in *Leptosphaeria maculans*. *European Journal of Plant Pathology* **114**: 17-31.
- Kutcher HR, Balesdent MH, Rimmer SR, Rouxel T, Chèvre AM, Delourme R, Brun H. 2010.** Frequency of avirulence genes in *Leptosphaeria maculans* in western Canada. *Canadian Journal of Plant Pathology* **32**: 77-85.
- Lawrence GJ, Mayo GME, Shepherd KW. 1981.** Interactions between genes controlling pathogenicity in the flax rust fungus. *Phytopathology* **71**: 12-19.
- Leach JE, Vera Cruz CM, Bai J, Leung H. 2001.** Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annual Review of Phytopathology* **39**: 187-224.
- Livak KJ, Schmittgen TD. 2001.** Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_t}$  method. *Methods* **25**: 402-408.
- Orbach MJ, Farrall L, Sweigard JA, Chumley FG, and Valent B. 2000.** A telomeric avirulence gene determines efficacy for the rice blast resistance gene Pi-ta. *Plant Cell* **12**: 2019–2032.
- Parlange P, Daverdin G, Fudal I, Kuhn ML, Balesdent MH, Blaise F, Grezes-Besset B, Rouxel T. 2009.** *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change. *Molecular Microbiology* **71**: 851-863.
- Rouxel T, Balesdent MH. 2010.** Avirulence Genes. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0021267.

- Rouxel T, Grandaubert J, Hane JK, Hoede C, van de Wouw AP, Couloux A, Dominguez V, Anthouard V, Bally P, Bourras S, et al. 2011.** Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat Induced Point mutations. *Nature Communication* in press.
- Rozen S, Skaletsky H. 2000.** Primer3 on the www for general users and for biologist programmers. *Methods in Molecular Biology* **132**: 365-386.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000.** Artemis: sequence visualization and annotation. *Bioinformatics* **16**: 944-945.
- Sprague SJ, Balesdent MH, Brun H, Hayden HL, Marcroft SJ, Pinochet X, Rouxel T, Howlett BJ. 2006.** Major gene resistance in *Brassica napus* (oilseed rape) is overcome by changes in virulence of populations of *Leptosphaeria maculans* in France and Australia. *European Journal of Plant Pathology* **114**: 33-40.
- Stachowiak A, Olechnowicz J, Jedryczka M, Rouxel T, Balesdent MH, Hapstadius I, Gladders P, Latunde-Dada A, Evans N. 2006.** Frequency of avirulence alleles in field populations of *Leptosphaeria maculans* in Europe. *European Journal of Plant Pathology* **114**: 67-75.
- Stergiopoulos I, de Wit PJGM. 2009.** Fungal effector proteins. *Annual Review of Phytopathology* **47**: 233-263.
- Voorrips RE. 2002.** MapChart: Software for the graphical presentation of linkage maps and QTLs. *The Journal of Heredity* **93**: 77-78.
- Van de Wouw AP, Cozijnsen AJ, Hane JK, Brunner PC, McDonald BA, Oliver RP, Howlett BJ. 2010.** Evolution of linked avirulence effectors in *Leptosphaeria maculans* is affected by genomic environment and exposure to resistance genes in host plants. *PLoS Pathogen* **6**: e1001180. doi:10.1371/journal.ppat.1001180.

**Yoshida K, Saitoh H, Fujisawa S, Kanzaki H, Matsumura H, Yoshida K, Tosa Y, Chuma I, Takano Y, Win J. 2009.** Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen *Magnaporthe oryzae*. *The Plant Cell* **21**: 1573-1591.

Table 1 : Characteristics of *Leptosphaeria maculans* parental isolates used in *in vitro* crosses and of their progeny.

Cross # <sup>a</sup>	Parental isolate 1		Parental isolate 2		No of progeny recovered	Segregation ratio (a3 :A3)
	Phenotype <sup>b</sup>	Putative genotype <sup>c</sup>	Phenotype	Putative genotype		
<b>33</b>	a3A4A7	<i>AvrLm3</i>	a3a4a7	<i>avrLm3</i>	238	233 : 5
v23.1.11 x Nz-T4		<i>AvrLm4 AvrLm7</i>		<i>avrLm4 avrLm7</i>		
<b>69</b>	A3a4a7	<i>AvrLm3</i>	a3a4a7	<i>avrLm3</i>	125	60 : 65
G06-E107 x Nz-T4		<i>avrLm4 avrLm7</i>		<i>avrLm4 avrLm7</i>		

<sup>a</sup> Cross id and name of parental isolates 1 and 2, respectively.

<sup>b</sup> ai indicates a virulent interaction phenotype toward the resistance gene *Rlmi*; Ai, indicates an avirulent interaction phenotype toward the resistance gene *Rlmi*.

<sup>c</sup> *AvrLmi* indicates the presence of an avirulent allele at the corresponding locus, *avrLmi* indicates a virulent allele, whatever the mutation event. Putative genotypes at *AvrLm3* locus determined following analysis of segregation in the progeny, not from the phenotype.

Table 2 : List and characteristics of polymorphic molecular markers used for *AvrLm3* mapping in cross #69

Marker	Forward/ Reverse	Primer sequence <sup>a</sup>	Hybridation temperature (°C)	Elongation time (s)	Reference
<i>AvrLm4-7</i>	F R	TATCGCATACCAACATTAGGC GATGGATCAACCCTAACAA	60	60	Parlange <i>et al.</i> 2009
<b>Minisatellite</b>					
MinLm9	F R	GCATTTGTTGGCTGGCATTGGA TGCGTTGGCGTGTGACATGACT	60	60	Parlange <i>et al.</i> 2009
Min12-96	F R	CCAGCACGAAATCGATTGAC AGGGCATCCCTCATGCTCTTT	60	60	This study
<b>CAPS</b>					
124F12_F	F R	AGAGAGCTGCTAATAACCTGC AATCCATCTCCTAAATTATTC	56	60	Parlange <i>et al.</i> 2009
133E3_R	F R	GATCTAAATAGAGCAAGATAGT GTTGCGCTTAGTAGTGGT	48	60	Parlange <i>et al.</i> 2009
46F10_R	F R	TATATCTTTTCTCCTTTTACC AGGAGCTCTAGTAAGGAAGGTC	48	60	Parlange <i>et al.</i> 2009
<b>Transposon junction</b>					
SC12-2	F R	CCCCTAGAAAAGACAATTCACG CGCCTAGTATAGTGTCTTACGC	60	60	This study
SC12-14	F R	GCTGCACGTGGATTCAAAC TCCTCGCACAGTTAATTCCTG	60	60	This study
SC12-16	F R	AACTTTACTAGCCTTCTAGCTTTCCTC TTTAGGTAGGGGTTGCAGGA	60	60	This study
SC12-19	F R	AAACATCTAGTGCCTCCATCTCT TTTCGGCTTAAGGTGATCT	60	60	This study

<sup>a</sup> Sequence of the primers (forward and reverse) used to amplify each marker

Table 3 : Complementation and RNAi silencing of *Leptosphaeria maculans* isolates

Type of experiment	Construction	Isolate transformed	No of transformants analysed
Functional complementation	pPZPnat1-A4A7 <sup>a</sup>	G06-E101	7
		G06-E107	7
	pPZPnat1-a4A7 <sup>a</sup>	G06-E101	10
		G06-E107	5
Silencing	pPZPnat1-sil <i>AvrLm4-7</i> <sup>b</sup>	v23.1.3	29

<sup>a</sup> Constructions from Parlange *et al.* (2009). A4A7, *AvrLm4-7* allele conferring avirulence toward *Rlm4* and *Rlm7*, a4A7, allele conferring avirulence toward *Rlm7* only; <sup>b</sup> this study.

Table 4 : Repartition of *Leptosphaeria maculans* field isolates into the 4 phenotypic classes defined as combination of phenotypes for avirulence toward *Rlm3* and *Rlm4*

Collection of isolates analysed by :	Total no of isolates analysed	Phenotypic class			
		A3A7 <sup>a</sup>	a3a7	a3A7	A3a7
Balesdent <i>et al.</i> , 2005	62	0 (0.0%) <sup>b</sup>	10 (16.1%)	37 (59.7%)	15 (24.2%)
Dilmaghani <i>et al.</i> , 2009	909	10 (1.1%)	35 (3.8%)	578 (63.6%)	286 (31.5%)
Kutcher <i>et al.</i> , 2010	87	1 (1.2%)	55 (63.2%)	23 (26.4%)	8 (9.2%)
Total	1058	11 (1.0%)	100 (9.5%)	638 (60.3%)	309 (29.2%)
This study (Grignon site)	1514	17 (1.1%)	14 (0.9%)	705 (46.6%)	778 (51.4%)

<sup>a</sup> A3A7, isolates avirulent toward *Rlm3* and *Rlm7*; a3a7, isolates virulent toward *Rlm3* and *Rlm7*; a3A7, isolates avirulent toward *Rlm7* only; A3a7, isolates avirulent toward *Rlm3* only.

<sup>b</sup> number of isolates, and percentage (in brackets) in the corresponding collection.

Table 5 : Distribution of molecular events leading to virulence toward *Rlm7* in a collection of *Leptosphaeria maculans* isolates classified as virulent or avirulent toward *Rlm3*

Molecular event	Total No of isolates	Phenotypic class <sup>a</sup>	
		a3a7	A3a7
Complete deletion of <i>AvrLm4-7</i>	341	5 (50.0%) <sup>b</sup>	336 (46.3%)
Partial deletion of, or insertion in, the gene	126	2 (20.0%)	124 (17.1%)
Mutations due to RIP	175	0	175 (24.1%)
Dinucleotidic deletion in <i>AvrLm4-7</i>	54	0	54 (7.4%)
Single nucleotidic deletion in <i>AvrLm4-7</i>	4	0	4 (0.6%)
SNP in <i>AvrLm4-7</i> :			
- with consequence on the protein structure	11	0	11 (1.5%)
- without evident consequence on the protein	19	3 (30.0%)	16 (2.2%)
No mechanism identified (intact gene)	5	0	5 (0.7%)
Total	735	10	725

<sup>a</sup> a3a7, isolates virulent on *Rlm3* and *Rlm7*; A3a7, isolates avirulent toward *Rlm3* only.

<sup>b</sup> number of isolates, and percentage (in brackets) in the corresponding phenotypic class.



Table 6 : Segregation of avirulence of *Leptosphaeria maculans* toward *Rlm3* in cross #69

	A3 (phenotype) <sup>a</sup>	<i>AvrLm7</i> <sup>b</sup>	No of isolates
Parental isolates :			
GO6-E107	A	+	
Nz-T4	V	-	
Phenotypic classes in the progeny	A	+	51
	V	+	7
	V	-	58
	A	-	9

<sup>a</sup> A, avirulent phenotype toward *Rlm3*; V, virulent phenotype.

<sup>b</sup>, + : the *AvrLm7* gene (virulent allele of G06-E107) can be amplified by PCR; - the *AvrLm7* gene can not be amplified.

## FIGURE LEGENDS

### **Figure 1 Genetic and physical maps of the *AvrLm4-7 – AvrLm3* region of *Leptosphaeria maculans***

(a) Genetic linkage group containing *AvrLm3* and *AvrLm4-7* in cross #69 with genetic distances in cM. (b) Marker names. The lines connect the genetic location of the markers to their physical position in the corresponding genomic region (3' end of Super-Contig 12 [SC12] of the sequenced isolate v23.1.3). (c) GC content (%) of the sequence along the region (sliding window of 500 bp); the horizontal line indicates the mean GC content (43.74%) of SC 12. (d) Graphical representation of the location of repeated elements as defined by Rouxel *et al.* (2011) in the region in both strands. Each colour corresponds to a given family of repeat. Red boxes correspond to telomere specific repeats (*LmTelo2*). Numbers indicate the scale in bp from the beginning of SC12. (e) Position of the last two isochores of SC12, GC\_13 and AT\_13, as defined by their homogeneous GC content.

### **Figure 2: Interaction phenotypes on *Rlm3* and *Rlm7* *Brassica napus* lines of *AvrLm3* *Leptosphaeria maculans* isolates complemented with the *AvrLm4-7* gene**

Two natural isolates displaying an avirulent phenotype toward *Rlm3* (G06-E101 [A; C] and G06-E107 [B, D], black bars) were transformed with the a4A7 (grey) or A4A7 (white) alleles of the *AvrLm4-7* gene. Isolates were inoculated on *Brassica napus* lines containing *Rlm7* (A, B) or *Rlm3* (C, D). The symptoms were assessed 14 dpi on a 1-6 scale with 1-3 corresponding to resistant responses of the plant and 4-6 to susceptibility. The two isolates with a star are false positive transformants that did not integrate the *AvrLm4-7* copy.

### **Figure 3: Interaction phenotypes of *L. maculans* isolates silenced for *AvrLm4-7* when inoculated on *Rlm3* and *Rlm7* *Brassica napus* lines**

*AvrLm4-7* was silenced using RNAi in the wild type v23.1.3 isolate (white square; phenotype a3A7). The transformants (black triangles) were evaluated for their interaction phenotypes on

*Rlm3* and *Rlm7*. The symptoms were assessed 16 dpi on a 1-6 scale with 1-3 corresponding to resistant responses of the plant and 4-6 to susceptibility. Two transformants with an altered phenotype on *Rlm3* (T4 and T24) and two control isolates (19.4.24 [A3a7], black square and Nz-T4 [a3a7], white triangle) are also shown.

Table S1 : List and characteristics of additional molecular markers and primers analysed for polymorphism in cross #69 of *Leptospira maculans*

Primer pair names	Forward /Reverse	Primer sequence	Polymorphism in cross #69	Hybridation temperature (°C)	Elongation time (s)	Reference
<b><u>Minisatellite</u></b>						
Min12-91	F	AAATAACGGGCATCAACGAC	-	60	60	This study
	R	AGTCTGTCCTGTTCCGGCATC				
Min12-92	F	GGCGTTATCCTGACTGTGGT	-	60	60	This study
	R	GCCTATCGCTCATGGAAGAA				
Min12-93	F	CCATGCTATCCGGTTATTGAA	-	60	60	This study
	R	AGCGTCTTCCGTAATATCG				
Min12-94	F	GGTGAGAGAGATGTCGAGGC	-	60	60	This study
	R	GGCCTTTACCTGTTGGCTT				
Min12-95	F	TTGAGGGTTTGTGCATCGTA	-	60	60	This study
	R	TAAGCATCGACACCAACTGC				
<b><u>CAPS</u></b>						
1B4-B5 F-A	F	GATCCCCTATTCCTATATTAGAT	-	50	60	Parlange <i>et al.</i> 2009
	R	GCAAACTAGGCTTATAACGT				
1B4-B5 F-B	F	GATCCCCTATTCCTATATTAGAT	-	50	60	Parlange <i>et al.</i> 2009
	R	AGTAAGGAAGGTTACCAA				
35H2-D4	F	GAAGTTTGCCTTAGTGCTA	-	50	60	Parlange <i>et al.</i> 2009
	R	GCTTATTATAGAGGAGGATAGA				
35H2-D9	F	TAGGACTGAGATGCCGACT	-	50	60	Parlange <i>et al.</i> 2009
	R	ACCTTAGCTTGCTACTAATTG				
<b><u>AFLP</u></b>						
TCTX1	F	ACTGTA TCTCCTAGAATAAG	-	60	60	Parlange <i>et al.</i> 2009
	R	TCTAGTAAGGAAGGTTTA				
<b><u>Transposon junction</u></b>						
SC12-1	F	TATCGCGGTAGATAATACACTTG	-	60	60	This study

SC12-3	R	TATAGACTAGCCTTAAGGGAGAATGC ACTAGTACTGCTTCCTGCATAGTAACC	-	60	60	This study
SC12-4	R	TAGGTAGGTAGGGCATCTAAAAGG GAGCCACGTTACCTATCTATCTCC TGCCCTGCCTATTGTTCTATGC	-	60	60	This study
SC12-5	F	AGGCATATAGGAGTGCGGTGC CGCTAGGAGTGGCTAGCGTA	+	60	60	This study
SC12-6	R	AGAGCAGGCATAGGAGCTAGAG GGTGTGTTGCAGGAGGGATA	-	60	60	This study
SC12-7	F	AGAGCTAGGAGGCTGCATAGG TTAATCTATAGACCTGTTGCATTAGCC	-	60	60	This study
SC12-8	R	CCACGGCACGTTACTACTTA AAAGTTATCCAAGCTCTAAAAGCAAA	+	60	60	This study
SC12-9	F	AAAGTTCAAGCTGCTAGTGGT GCTCTACAAGTGCGGGGATA	-	60	60	This study
SC12-10	R	CCCTATTGCTTATTGCGCTCT TGCTCTAATTAGCTACACCACCTACT	-	60	60	This study
SC12-11	F	ATGGAAAATTGGATCACTAAGGTT AACGCACCTAGCTAGGTTTLAGGT	+	60	60	This study
SC12-12	R	GCATATAGGAGTGCGCGTCT GCAACTACTAATCTCTGCCCTGA	-	60	60	This study
SC12-13	F	TTAATATCCCTCCTGCAACATTAGT CCCCGTATAACCCACTATCTAATCA	+	60	60	This study
SC12-15	F	GTAGCAGGCTATCGCGTTTC GGAACGCTTATAGATACGTGTAGGT	-	60	105	This study
SC12-17	F	AGAAAATGCAAGCATAGTAGGG GGTTTTGATTTTCCCTCGCATA	-	60	60	This study
SC12-18	R	CTACGCCTAAGATTACGTTTACCTG ACGATTCACGATGGCTTATG	-	60	60	This study

Table S2 : List of primer pairs used to amplify genes within the *Leptosphaeria maculans* GC-  
isochore GC\_13 of Super-Contig 12.

Primair pair names <sup>a</sup>	Forward / Reverse	Primer sequence	Hybridation temperature (°C)	Elongation time (s)
GD-GC1	F R	ATGTCTGGTCCGAAAGCAC ACGATGCTAATCCCCATTTG	60	60
GD-GC2	F R	CCAGCAACCAGTATCGACAG CAGCTTCCCCGTTCTTCTC	60	90
GD-GC3	F R	CTGGCTTTGTGCTTTCGAGT TCGACAAAGCCCCATAGTA	60	120
GD-GC3-4*	F R	GATCCGATAGCGATGAGTCC TGAAAATGGAACCATGCAAC	60	60
<b>GD-GC4</b>	F R	AGTGAATATCGGACTGCCATC GGACGCCATTATGTGCTTTT	60	90
<b>GD-GC5</b>	F R	AAACTTGCTGTGCTCGATGA CATGTAACAATTCTCTCCCATTGA	60	90
GD-GC6	F R	CCACAAGCTGTTTCAGCACAT GAACAAACTCCGCCATGC	60	90
GD-GC7	F R	GCTGGAGTCGTTCCCTCCTATC GCGTACATTAGAGTTGAGGGTTG	60	120
GD-GC8	F R	TTCCAGGAGCGGCTAAATAA CTCCTTGTCCTTCCTTGACG	60	200
<b>GD-GC9</b>	F R	GGCAAGACCAACCAACAAC TTGCTCTTCACTCTGGTTGC	60	60
GD-GC10	F R	ATGAACATGGTGCATTGAC TCAGGGATTTGAATGGACAA	60	90
GD-GC11	F R	TATCGCTTGCTTTGCATGTC CACTGTTGACATTTAGCCCCT	60	150
GD-GC12	F R	GATCGTTTGTGCGAAGCTTTCT GTCTTGTCAGCCGTTTAGCC	60	30
GD-GC13	F R	GCTTTGCGTCAGGTATGTTG ATGGAGGCCTTGAGGACTTT	60	150
GD-GC14	F R	TGACAGGACAAAGCTCGATG TTCCAGCCAATAATGCCTTC	60	60
GD-GC15	F R	CCTCGTCCTCTTTGCACTCT AACCAAGCTGATACGCCTTC	60	60
GD-GC15-16*	F R	TTGTCGACTCAACGGAATGA GGGCGGTAAGATGTAGAGCA	60	90
GD-GC16	F R	ATGCCGGCTCGTATAGCTC TGCTTGCATTAGCCAGTGAC	60	90
<b>GD-GC17</b>	F R	AAGGTCAAGGGCGAGAATTT TGTTCTTTGTGCGCTTCTTCG	60	150
GD-GC18	F R	CAATTACCCCTCCCTCCCTA GCAGCTTCATCGAACTGTGA	60	90
GD-GC19	F R	GACTACTCACGCCCGACAAC CGTCCTAGCTTCGCAAAAAC	60	120
GD-GC19-	F	GGCTGGGATTTGCAGAGTAT	60	120

20*	R	GCGGATACTTGAGGACGGTA		
<b>GD-GC20</b>	F	GGTACCTCCCTCCGAATGTT	60	60
	R	ACCCTACAAGCTGCCGAGT		

\* Primer pair names followed by a star amplify intergenic regions between two genes; Primer pair names in bold indicate genes that have been sequenced following PCR amplification

Figure 1

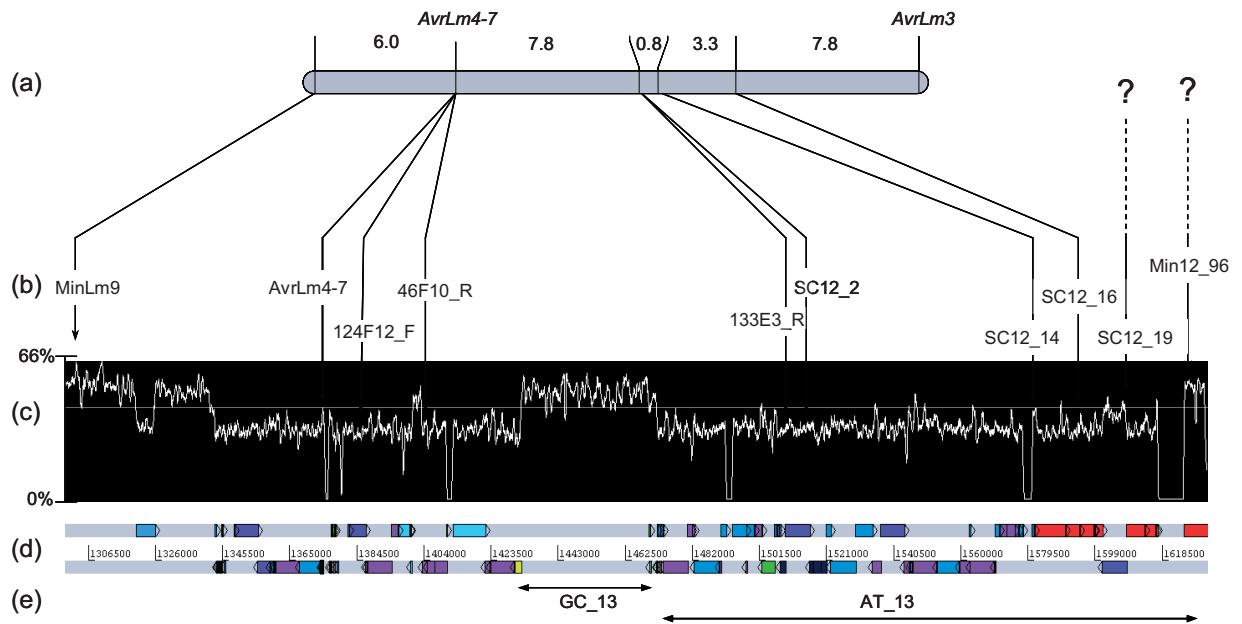




Figure 2

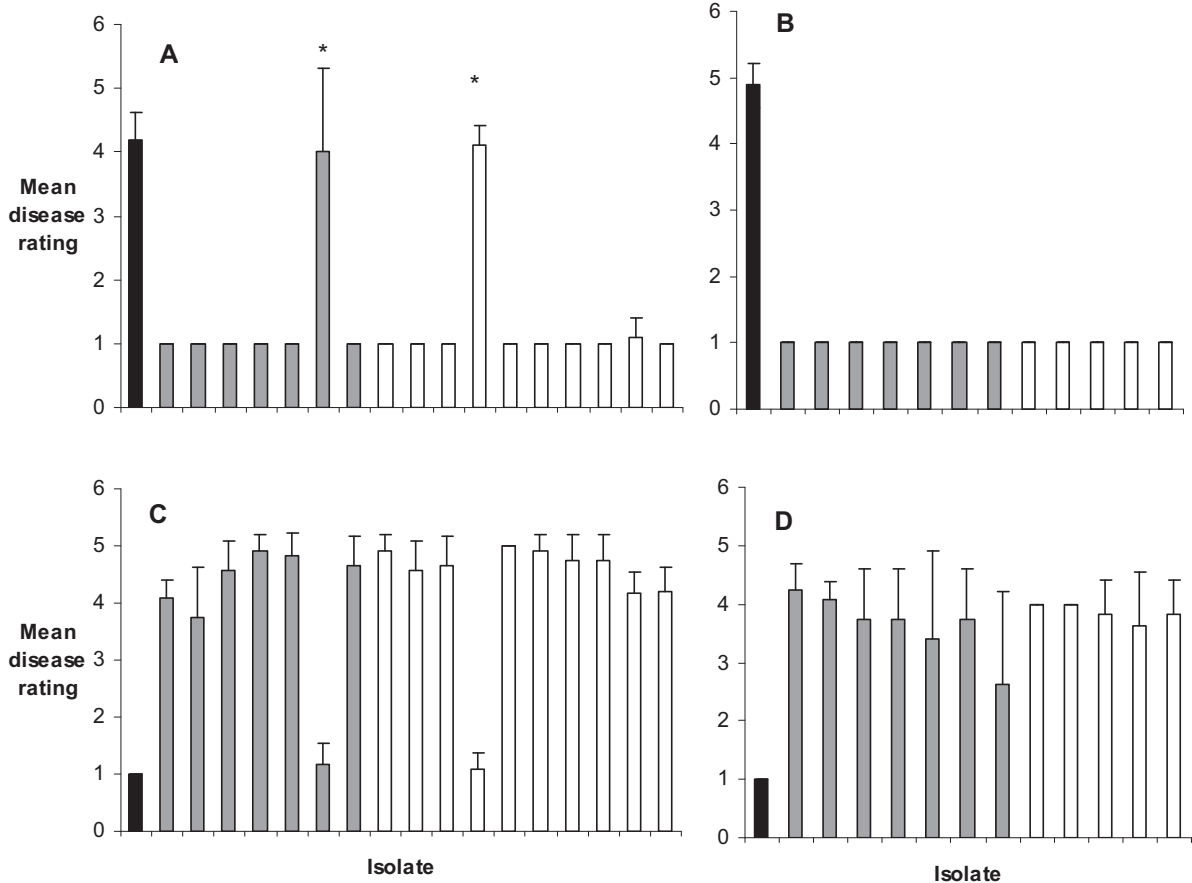
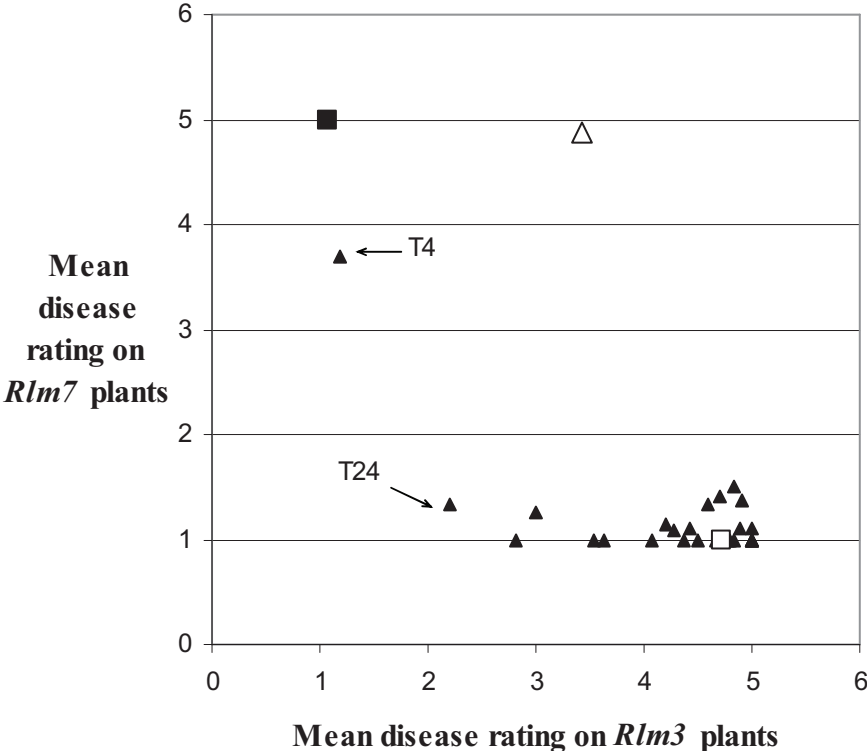


Figure 3





# **Discussion générale**



Au début de ma thèse, le clonage d'*AvrLm4-7* et l'utilisation récente de *Rlm7* en pratique agricole fournissaient un contexte de choix pour étudier en temps réel l'évolution moléculaire d'*AvrLm4-7* sous pression de sélection. L'objectif principal de ma thèse était de réaliser pour la première fois chez un champignon phytopathogène l'étude précoce et détaillée des mécanismes moléculaires à l'origine du contournement d'un gène de résistance. Le travail effectué au cours de cette thèse a permis l'établissement d'une importante collection de souches virulentes et avirulentes vis-à-vis de *Rlm7*, l'évaluation de l'importance de certaines pratiques culturales sur la durabilité de *Rlm7*, la détermination des premiers événements de mutation à l'origine du contournement de *Rlm7* ainsi que de leur évolution dans le temps et enfin d'améliorer la compréhension de l'interaction *AvrLm4-7* – *AvrLm3*. J'ai pu ainsi apporter des éléments de compréhension aux interactions *L. maculans* / *B. napus* à la fois sur des aspects fonctionnels et évolutifs.

### **1. Aspect fonctionnel**

*AvrLm4-7* est à la fois un gène d'avirulence d'ascomycète typique de part sa structure (petite protéine potentiellement sécrétée, 8 cystéines, etc.) et son profil d'expression (Chapitre 1) mais également des gènes d'avirulence et d'une grande partie des effecteurs de *L. maculans* du fait de sa localisation au sein d'un isochoire AT (Gout et al., 2007 ; Fudal et al., 2009 ; Rouxel et al., 2011 ; Chapitre 1). Néanmoins, ce gène est également relativement particulier de part sa double spécificité d'interaction envers les gènes de résistance *Rlm4* et *Rlm7*.

Les travaux de thèse de Francis Parlange ont permis de montrer que le nucléotide présent en position 358 du gène semblait déterminer la virulence de la souche vis-à-vis de *Rlm4* (Parlange, 2006). Par le phénotypage et l'analyse moléculaire d'une importante collection de souches établie à l'occasion de ma thèse j'ai pu confirmer cette observation (Chapitre 2). La présence dans *AvrLm4-7* d'une guanine en position 358 et d'une glycine en position 120 de la protéine est caractéristique des souches A4 tandis que les souches a4 présentent une cytosine en position 358 du gène et une arginine en position 120 de la protéine. L'importance fonctionnelle de ce résidu dans la spécificité d'interaction *AvrLm4-7-Rlm4* a par ailleurs été validée par mutagenèse dirigée, d'une part par la perte de l'avirulence A4 liée à la mutation G<sup>358</sup>>C<sup>358</sup> chez une souche A4A7 (Parlange 2006, Chapitre 1) et

d'autre part par la restauration de l'avirulence A4 constatée suite à la mutation C<sup>358</sup>>G<sup>358</sup> chez une souche a4A7 (Chapitre 1).

La protéine AvrLm4-7 a été produite en système hétérologue (*Pichia pastoris*) (F. Blaise, Com. Pers.) et cristallisée (I. Fudal et H. van Tilbeurgh, non publié), permettant aujourd'hui d'avoir accès à la structure tridimensionnelle des allèles a4A7 de v23.1.2 et A4A7 de v23.1.3. AvrLm4-7 présente une structure constituée de trois feuillets beta, d'une hélice alpha et possède huit cystéines à l'origine de quatre ponts disulfures (Figure 1, discussion générale). L'acide aminé 120 d'AvrLm4-7 est localisé à la surface de la protéine et la comparaison des structures a4 et A4 d'AvrLm4-7 a permis de constater que la mutation de la glycine en arginine à l'origine de la perte d'avirulence vis-à-vis de Rlm4 n'a sans doute aucun impact sur la structure tridimensionnelle de la protéine. Dodds *et al.* (2006) ont montré que la reconnaissance des gènes d'avirulence AvrL567 de *Melampsora lini* par les protéines de résistance L5, L6, L7 chez le lin était liée à des interactions directes entre les protéines d'avirulence et leurs protéines de résistance respectives. Les variants virulents existant présentent des mutations mais gardent la structure et la stabilité de leur protéine. Echapper à la reconnaissance d'un gène de résistance tout en maintenant la présence et la fonction de la protéine d'avirulence serait caractéristique des interactions directes entre protéines d'avirulence et protéines de résistance (Bent et Mackey, 2007). Cette interaction directe entre protéines d'avirulence et protéine de résistance n'a été que rarement démontrée mais l'importance de l'acide aminé 120 dans la virulence des souches vis-à-vis de *Rlm4*, sa localisation à la surface d'AvrLm4-7 et le peu d'impact associé à sa mutation sur sa structure tridimensionnelle laissent penser que l'interaction AvrLm4-7–Rlm4 pourrait également suivre un modèle « ligand – récepteur ». *Rlm4* n'étant pas encore cloné, la validité de cette interaction physique n'a pu être testée à ce jour.

A la différence des souches a4 et A4 dont les allèles *avrLm4* et *AvrLm4* se distinguent par une unique mutation ponctuelle, un grand nombre d'évènements de mutation distincts expliquant la virulence a7 a pu être répertorié dans notre collection (Chapitre 2). Ces évènements de mutation affectent dans la très grande majorité des cas de façon évidente l'intégrité de la protéine, soit par délétion totale ou partielle du gène soit par l'introduction de codons stop dans la séquence codante (Chapitre 2). Ces caractéristiques sont typiques d'un modèle de reconnaissance indirecte dans

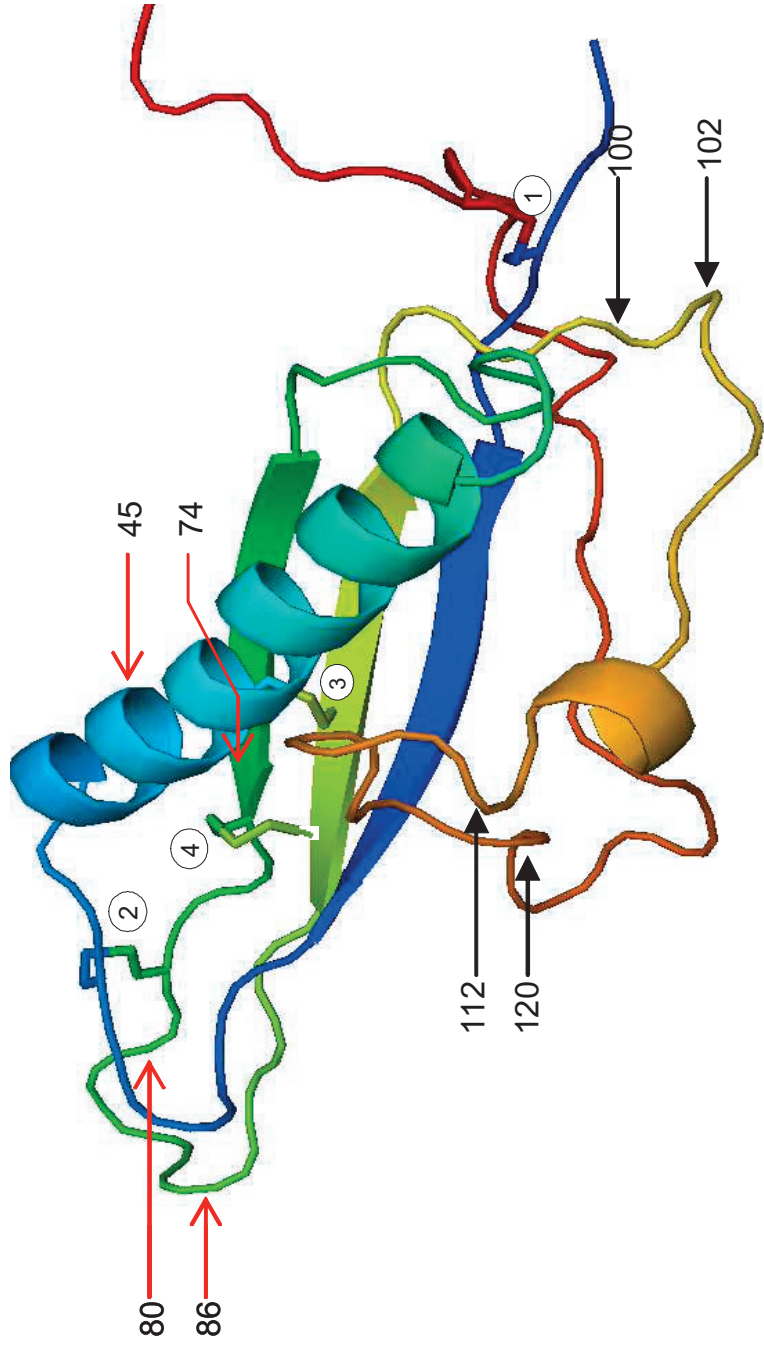
lequel Rlm7 surveillerait non pas la présence d'AvrLm4-7 mais plutôt la cible végétale de cette protéine (Bent et Mackey, 2007).

On a donc ici un gène d'avirulence dont la présence au sein d'une souche serait détectée par deux gènes de résistance utilisant chacun un modèle de reconnaissance distinct.

*AvrLm4-7* étant la cible de deux gènes de résistance, son importance dans le développement de *L. maculans* peut être supposée. Cette hypothèse trouve confirmation avec les études portant sur le déficit de fitness affectant les souches ayant perdue l'avirulence A4 et/ou A7. Huang *et al.*, (2006) ont en effet montré que la perte de l'avirulence A4 liée à la mutation du résidu 120 d'AvrLm4-7 s'accompagne d'un déficit de fitness : les souches avirulentes induisent des lésions en plus grand nombre et de diamètre plus important, une plus grande production d'inoculum secondaire ainsi qu'une meilleure colonisation des tissus végétaux que les souches possédant l'allèle virulent. Dans le modèle de reconnaissance supposé s'appliquer pour l'interaction *AvrLm4-7-Rlm4*, ce déficit de fitness laisserait supposer que l'interaction physique entre les deux protéines se situerait au niveau d'un site fonctionnel favorisant le développement de la souche et au sein duquel l'acide aminé 120 jouerait un rôle primordial.

Les dernières études effectuées dans notre équipe sur le déficit de fitness potentiel lié à la perte de l'avirulence A4 et/ou A7 ont également mis en évidence une différence d'agressivité très importante entre souches possédant ou non différents allèles d'*AvrLm4-7*. Cette agressivité a été déterminée au premier stade de l'infection par la mesure de la surface des macules induites par des souches quasi-isogéniques possédant ou non le gène *AvrLm4-7*, observées sur cotylédon 15 jours après inoculation. Les souches a4A7 et A4A7 induisent des macules significativement plus grandes que les souches a4a7 sur des variétés de colza sensibles, tandis que les différences obtenues entre souches a4A7 et A4A7 (Figure 1, discussion générale), bien qu'allant dans le même sens que décrit par Huang *et al.*, (2006) ne sont pas significatives (Mongin, 2010). La mutation du résidu 120 d'AvrLm4-7 à l'origine de la perte de l'avirulence A4 induit donc un coût de virulence pour les souches affectées mais constitue une solution permettant à la souche de perdre l'avirulence A4 tout en conservant la fonction d'AvrLm7, la souche éviterait ainsi un coût de virulence encore plus élevé lié à la perte totale de la fonction du gène.



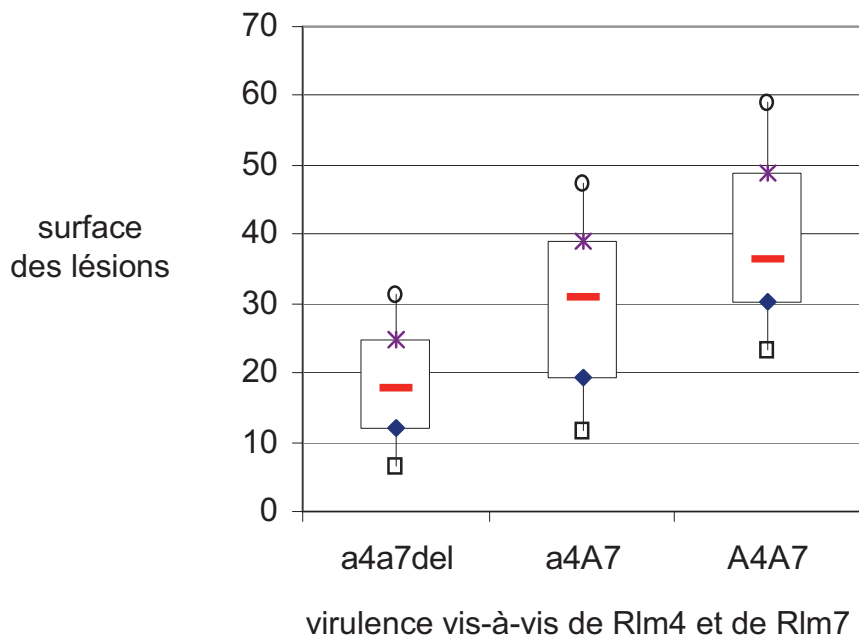


**Figure 1 (discussion générale) : Localisation des mutations ponctuelles décrites chez les souches virulentes et avirulentes vis-à-vis de *Rlm7* sur la représentation tridimensionnelle de l'allèle *a4A7* de la protéine *AvrLm4-7***

Le logiciel PyMOL a été utilisé pour cette représentation. Les flèches rouges indiquent la localisation des mutations ponctuelles répertoriées dans les souches avirulentes A7. Les flèches noires indiquent la localisation des mutations ponctuelles trouvées parmi les souches virulentes a7. Les chiffres présents aux extrémités des flèches indiquent la position exacte de l'acide aminé muté. Les 4 cercles numérotés correspondent aux emplacements des quatre ponts disulfures présents au sein de la protéine AvrLm4-7 (I. Fudal et H. van Tilbeurgh, non publié).

Ajouté aux mutations affectant de façon drastique l'intégrité structurale d'AvrLm4-7, il existe quelques souches pour lesquelles les causes de la virulence a7 ne sont pas intuitives. Ainsi, pour 6 souches a7, la séquence d'AvrLm4-7 ne montre pas de mutation permettant de la distinguer d'allèles avirulents tandis que 15 souches présentent une mutation ponctuelle non synonyme n'introduisant pas de codon stop dans la séquence codante. Mis à part les mutations ponctuelles affectant les cystéines de la protéine qui devraient avoir une répercussion importante sur la structure et donc la fonction de la protéine par la disruption d'un pont disulfure, 13 souches réparties selon 3 mutations n'étant pas supposées affecter de façon trop importante la structure de la protéine ont pu être répertoriées: R<sup>100</sup>>P<sup>100</sup>, F<sup>102</sup>>S<sup>102</sup> et S<sup>112</sup>>R<sup>112</sup>. De la même façon que certaines mutations affectant le gène *Avr4* de *C. fulvum* permettent d'échapper à la reconnaissance par le produit du gène de résistance *Cf-4* tout en conservant sa capacité à se lier à la chitine et donc à protéger le champignon des chitinases de la plante (van den Burg *et al.*, 2003), de telles souches pourraient avoir conservé l'intégrité structurale de la protéine ainsi que sa fonction tout en évitant la reconnaissance par *Rlm7*. Du fait du déficit de fitness associé à la perte de l'avirulence A7, une telle particularité procurerait à ces souches un avantage sélectif certain vis-à-vis des souches virulentes ayant perdu la fonctionnalité d'AvrLm4-7. Ces mutations ponctuelles présentent également la propriété d'être localisées à proximité du résidu 120 (Figure 1, discussion générale), identifiant alors une boucle externe comprenant ce résidu comme une région potentiellement importante pour la fonction de la protéine et pour sa reconnaissance.

Selon les modèles de reconnaissance proposés pour les interactions AvrLm4-7-Rlm4 et AvrLm4-7-Rlm7, les mutations affectant l'intégrité structurale de la protéine AvrLm4-7 dans le cas de la perte d'avirulence A7 affectent par répercussion la zone supposée primordiale à la reconnaissance d'AvrLm4-7 par Rlm4. Dans l'hypothèse où certaines souches virulentes a7 auraient conservé l'intégrité structurale de la protéine, ces souches devraient potentiellement pouvoir être avirulentes vis-à-vis de Rlm4. Cependant, parmi toutes les souches étudiées au cours de ma thèse, aucune ne présente un phénotype A4a7. Du fait de l'utilisation de cultivars *Rlm4* pendant près de 40 ans, la grande majorité des souches isolées en 2000 étaient a4 (Balesdent *et al.*, 2006). La faible fréquence des souches A4 dans la population soumise à sélection *Rlm7*, cumulée à la faible fréquence des souches a7 possédant des allèles *AvrLm4-7* ayant conservé leur structure tridimensionnelle,



**Figure 2, discussion générale : Représentation en « Box Plot » de la surface des lésions induites par des souches de *L. maculans* différents seulement par leur allèle d'*AvrLm4-7* sur une variété de colza sensible**

La ligne rouge représente la médiane des valeurs de surface de macule obtenues 15 jours après inoculation ; les « boîtes » comportent les valeurs comprises entre le premier et le troisième quart de la distribution ; les carrés et ronds représentent respectivement le premier et le neuvième dixième de cette distribution. Une différence significative est mesurable entre les souches a7 et les souches A7 tandis qu'une différence non significative existe entre les souches a4 et A4. Les surfaces sont indiquées en mm<sup>2</sup> (Mongin, 2010).

suffirait à expliquer que nous n'avons pas pu observer dans notre collection de souche de phénotype A4a7.

De façon surprenante, nous avons également pu constater que la perte de l'avirulence A7 s'accompagnait très souvent en France et en Europe de la résurgence du phénotype A3 (Chapitre 3). *AvrLm4-7* ayant une double spécificité d'interaction envers *Rlm4* et *Rlm7*, nous avons tout d'abord supposé qu'*AvrLm3* était une troisième forme allélique d'*AvrLm4-7*. *AvrLm3* s'avère en fait être un gène distinct d'*AvrLm4-7* localisé à moins de 20 cM de celui-ci en région télomérique (Figure 2, Chapitre 3). *AvrLm3*, que nous n'avons pas réussi à cloner ici à cause d'un problème d'assemblage en extrémité chromosomique, est, selon le modèle gène-à-gène, spécifiquement reconnu par le gène de résistance *Rlm3*. Cependant, nous avons pu observer puis démontrer à la fois à l'aide d'analyses de population, par complémentarité fonctionnelle ainsi que par l'extinction de l'expression d'*AvrLm4-7* que la présence d'un allèle fonctionnel d'*AvrLm4-7* dissimulait la présence d'*AvrLm3*, rendant de ce fait la souche virulente vis-à-vis de *Rlm3*. Malgré sa localisation télomérique, nos travaux suggèrent que la majorité des souches françaises possèdent *AvrLm3*, signalant de ce fait un rôle potentiellement important de ce gène dans le cycle vital du champignon.

La relation antagoniste entre les deux phénotypes avirulent A3 et A7 pourrait s'expliquer par des mécanismes de reconnaissance interdépendants. Tout comme la reconnaissance d'*AvrRPM1* par la protéine de résistance *RPM1* et celle d'*AvrRpt2* par *RPS2* dans le pathosystème *A. thaliana* / *P. syringae* dépendent de l'activité des protéines d'avirulence sur la même protéine végétale cible *RIN4* (Figure 6, introduction), *AvrLm4-7* et *AvrLm3* pourraient avoir une même cible végétale (Figure 3, discussion générale). Dans ce modèle, l'antagonisme A3/A7 s'expliquerait par une action d'*AvrLm4-7* sur la protéine cible empêchant l'action d'*AvrLm3* ou une compétition entre les deux protéines qu'*AvrLm7* remporterait de par une meilleure efficacité ou encore une action plus précoce dans le temps. On retrouve cette compétition dans le pathosystème *A. thaliana* / *P. syringae* autour de la protéine *RIN4* (Axtell et Staskawicz, 2003). En effet, son hyperphosphorylation par *AvrRPM1* et son clivage par *AvrRpt2* sont respectivement reconnus par *RPM1* et *RPS2* mais l'action de *AvrRpt2* sur *RIN4* empêche *AvrRPM1* d'agir et prévient donc la reconnaissance de la souche par *RPM1*.

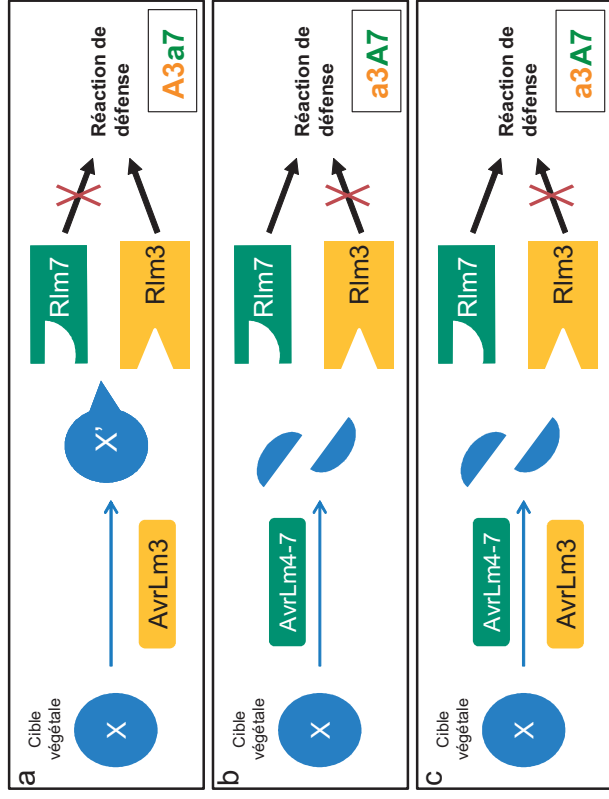
Dans un second modèle possible, la cible d'AvrLm4-7 ne serait pas une protéine végétale mais directement AvrLm3. Dans ce modèle, la protéine AvrLm3 serait reconnue directement ou indirectement par Rlm3 tandis que Rlm7 détecterait l'interaction AvrLm4-7 / AvrLm3 ou les modifications apportées par AvrLm4-7 sur AvrLm3 (Figure 3, discussion générale).

Au cours de cette étude, environ 1% des souches étudiées présentaient un phénotype A3A7. Ce phénotype est paradoxal puisqu'il traduit la présence et la reconnaissance des produits des deux gènes d'avirulence alors que nous avons démontré que l'action d'AvrLm4-7 dissimule la présence d'AvrLm3. L'existence de telles souches pourrait s'expliquer par une sous expression d'AvrLm4-7 ou par des mutations particulières affectant AvrLm3 qui empêcheraient l'action d'AvrLm4-7. A l'opposé des souches A3A7, existent des souches virulentes vis-à-vis de Rlm7 mais aussi vis-à-vis de Rlm3 (a3a7) représentant également moins de 1% des souches étudiées au cours de cette thèse. Ces souches présentent une différence statistique avec les souches A3a7 dans la distribution des événements mutationnels affectant *AvrLm4-7*. Cette différence de distribution s'explique par une proportion significativement plus importante de souches pour lesquelles une mutation ponctuelle expliquerait la virulence a7. Le clonage d'*AvrLm3* n'étant pas encore effectué, on ne peut pour le moment déterminer si la virulence a3 serait due à la délétion d'AvrLm3 ou à des mutations affectant sa séquence mais il est néanmoins intéressant de noter que deux des trois mutations ponctuelles répertoriées chez les souches a3a7 correspondent à la mutation du résidu 112 d'AvrLm4-7 et que cette mutation n'est jamais observée chez les souches A3a7. Nous avons déjà émis l'hypothèse que la mutation de ce résidu permettrait de conserver l'intégrité structurale de la protéine AvrLm4-7 ainsi que sa fonction tout en évitant la reconnaissance par *Rlm7*. Ces souches pourraient donc potentiellement avoir conservé les fonctions biologiques d'AvrLm3 et d'AvrLm4-7 mais sans qu'aucune de ces protéines ne puisse être reconnue par leurs gènes de résistance respectifs.

Afin de poursuivre le travail effectué sur AvrLm4-7 au cours de cette thèse et en particulier la compréhension de l'interaction AvrLm3-AvrLm4-7, il est aujourd'hui important d'étudier plus en profondeur le gène *AvrLm3*.

Pour ce faire, il s'agira dans un premier temps d'achever le clonage d'*AvrLm3*. A la fin de ma thèse, j'avais pu localiser *AvrLm3* au sein d'une région télomérique ou sub-télomérique. Pour finaliser le clonage d'*AvrLm3*, des marqueurs génétiques devront être générés dans les séquences télomériques non assemblées du génome et les marqueurs génétiques liés au groupe de liaison contenant *AvrLm4-7* seront recherchés et assemblés par cartographie génétique. Au cas où cette méthode ne fonctionnerait pas, plusieurs souches de *L. maculans* sont en cours de re-séquençage et permettront une analyse comparative entre souches *AvrLm3* et souches *avrLm3* permettant de trouver des marqueurs télomériques ainsi que des marqueurs de gènes candidats. Le séquençage de la souche NzT4, souche a3a4a7 déjà utilisée comme parent du croisement ayant permis d'entamer la marche chromosomique vers *AvrLm3* et supposée dépourvue du gène *AvrLm3*, est particulièrement attendu à cet effet. Ce clonage permettrait tout d'abord de déterminer si *AvrLm3* partage les mêmes caractéristiques moléculaires que les autres gènes d'avirulence clonés : petite protéine potentiellement secrétée localisée au sein d'un isochore AT, surexpression lors de l'infection, etc. La présence d'*AvrLm3* au sein des populations françaises et européennes pourrait ensuite être confirmée, permettant alors de valider l'intérêt agronomique de l'utilisation des résistances *Rlm3* et *Rlm7* en combinaison ou en alternance. L'étude comparée des cinétiques d'expression d'*AvrLm3* et d'*AvrLm7* apportera des éléments importants pour la validation ou invalidation des modèles d'interaction proposés entre ces gènes (Figure 3, discussion générale).

L'étude des gènes d'avirulence *AvrLm3* et *AvrLm4-7* a également permis d'émettre plusieurs hypothèses quant aux modèles de reconnaissance existant entre ces gènes et leurs gènes de résistance spécifiques. Afin de déterminer si un de ces modèles est effectivement applicable à l'interaction *AvrLm3-AvrLm4-7*, la recherche des différents interactants d'*AvrLm4-7* et d'*AvrLm3* est particulièrement importante. Afin de tester l'hypothèse selon laquelle *AvrLm3* serait directement modifié par *AvrLm4-7*, une approche double hybride est envisageable, l'une des deux protéines étant utilisée comme appât et l'autre comme proie. Dans le cas où cette hypothèse ne semblerait pas concluante ou non réalisable du fait d'un clonage d'*AvrLm3* encore difficile à réaliser, une autre analyse double hybride pourrait être réalisée. Cette fois ci, *AvrLm4-7* serait « l'appât » tandis qu'une banque d'ADNc représentant les gènes exprimés lors d'une phase précoce de l'infection serait utilisée comme « proie ».



## Modèle 1

## Modèle 2

**Figure 3 (discussion générale) : Modèles hypothétiques des mécanismes d'interaction entre les protéines d'avirulence AvrLm3 et AvrLm4-7**

**Modèle 1: AvrLm3 et AvrLm4-7 ont pour cible la même protéine végétale**

- a. En présence d'AvrLm3, la cible végétale (X) sera modifiée (X') de telle façon que seule Rlm3 pourra détecter sa présence.
- b. En présence d'AvrLm4-7, la cible végétale sera modifiée de telle façon que seule Rlm7 pourra détecter sa présence.
- c. En présence d'AvrLm3 et d'AvrLm4-7, la cible végétale est modifiée par l'action directe d'AvrLm4-7. Cette modification (exemple ici avec le clivage de la protéine) empêche AvrLm3 d'agir sur la cible. Seule Rlm7 peut alors détecter la présence de la protéine végétale.

**Modèle 2: la protéine AvrLm3 est prise pour cible par AvrLm4-7**

- d. En l'absence d'AvrLm4-7, AvrLm3 est reconnu directement par la protéine de résistance Rlm3 qui déclenche alors les réactions de défense de la plante. Dans ce cas, Rlm7 ne reconnaît pas AvrLm3.
- e. La modification de la protéine AvrLm3 apportée par AvrLm4-7 ne permet plus à Rlm3 de la détecter. La protéine AvrLm3 modifiée (AvrLm3\*) est en revanche reconnue par Rlm7 qui déclenche alors les réactions de défense de la plante.

Le séquençage d'AvrLm4-7 chez différentes souches a7 a permis de montrer que certaines souches n'étaient affectées que par une mutation ponctuelle d'un des résidus d'AvrLm4-7. Afin d'obtenir plus d'information sur l'aspect fonctionnel d'AvrLm4-7, il serait alors intéressant de valider par mutagenèse dirigée l'importance des acides aminés R<sup>100</sup>, F<sup>102</sup> et S<sup>112</sup> dans la reconnaissance AvrLm4-7-Rlm7. Dans le cas où ces mutations seraient bien à l'origine de la virulence des souches, il serait également intéressant de remplacer dans ces souches par mutagenèse dirigée l'arginine en position 120 d'AvrLm4-7 par une glycine afin de vérifier si une souche A4a7 peut être générée. Nous pourrions ainsi vérifier si l'absence de souche A4a7 est due à une incompatibilité entre les modèles de reconnaissance d'AvrLm4-7 par Rlm4 et par Rlm7 ou simplement à la faible probabilité qu'une telle souche apparaisse dans les populations actuelles de *L. maculans*. Alternativement il serait possible de muter à partir de l'allèle A4A7 d'AvrLm4-7, les sites correspondant aux SNP identifiés au cours de mon travail comme potentiellement impliqués dans l'interaction AvrLm7-Rlm7 (R<sup>100</sup>, F<sup>102</sup> et S<sup>112</sup>). Il faut noter que si une telle approche était fructueuse, *Rlm4* n'étant pas cloné à ce jour, nous aurions ainsi généré une souche A4a7 très utile afin de pouvoir tester la présence de *Rlm4* dans les cultivars *Rlm7*.

## **2. Aspect évolutif**

Le succès des cultivars *Rlm7* observé depuis 2006 s'explique par la présence en France et en Europe de souches de *L. maculans* quasi exclusivement A7 (Balesdent *et al.*, 2006; Stachowiak *et al.*, 2006). Néanmoins, l'utilisation de ce type de cultivar impose une forte pression de sélection sur la population pathogène, ce qui, à la vue de l'historique de contournement de gènes de résistance par *L. maculans* (Rouxel *et al.*, 2003; Sprague *et al.*, 2006), laissait envisager le contournement de *Rlm7* en seulement quelques années.

Tandis qu'à l'automne 2007, moins de 1% des souches prélevées sur cultivars non discriminants à Versailles étaient virulentes vis-à-vis de *Rlm7*, cette proportion dépassait 36% des souches observées sur le site de Grignon en troisième année d'expérimentation. Cet écart est d'autant plus marquant sur cultivar *Rlm7* puisque aucune souche a7 n'a pu être trouvée à l'automne 2007 à Versailles alors qu'à la même époque ainsi que l'année suivante, plus de 94% des souches observées à



Grignon sur cultivar *Rlm7* étaient virulentes vis-à-vis de *Rlm7*. Partant d'une population comportant moins de 1% de souches a7 (Balesdent *et al.*, 2006), on a donc eu une augmentation très rapide de la proportion de ces souches sur le site de Grignon. Les sites expérimentaux de Grignon et de Versailles ne se distinguent que par les méthodes de culture employées, favorisant sur Grignon la reproduction sexuée du champignon. La forte proportion de souches a7 observée sur Grignon démontre que la capacité offerte ou non au champignon de pouvoir effectuer un cycle infectieux complet incluant sa reproduction sexuée influence très fortement l'apparition de souches virulentes. La nature des événements de mutation affectant *AvrLm4-7* et à l'origine de la virulence a7 des souches observées traduit également l'importance de la reproduction sexuée dans l'apparition de ces souches. En effet, la majorité de ces événements sont soit dus à un réarrangement chromosomique (délétion partielle ou totale du gène,...) (63,3%) soit au RIP (23,8 %), deux mécanismes mutationnels favorisés ou exigeant une reproduction sexuée.

McDonald et Linde (2002) considèrent que les agents pathogènes présentant le plus fort risque de contourner la résistance empêchant leur développement présentent un mode de reproduction mixte (sexué + asexué), une forte capacité de dispersion, une population importante et un taux de mutation élevé. Le contraste observé entre les sites de Grignon et de Versailles confirme que le mode de reproduction mixte de *L. maculans* est primordial à l'apparition de souches virulentes et que le passage de ce mode de reproduction mixte à un mode seulement asexué diminue son aptitude à contourner *Rlm7*. La « mutabilité » d'*AvrLm4-7* liée à sa position au sein d'un isochoire AT (possibilité de RIP débordant des séquences répétées adjacentes, réorganisation chromosomique, etc.) semble également primordiale dans la capacité de la population à contourner rapidement la reconnaissance liée à *Rlm7*. En revanche, la capacité de dispersion de *L. maculans* ne semble pas aussi déterminante. En effet, Il est notable qu'aucune souche virulente vis-à-vis de *Rlm7* n'a pu être décelée à 600m de la parcelle expérimentale lorsque l'on était à 18% de souches virulentes au sein de celle-ci (Chapitre 2, figure 2). Nous avons de plus pu montrer que les événements de mutation menant à la perte de l'avirulence A7 apparaissaient au sein même de la parcelle de culture.

Les plantes et leurs agents pathogènes sont constamment au centre d'une course aux armements (Stahl et Bishop, 2000). Avec mon travail de thèse, trois

études portant sur les mécanismes employés par *L. maculans* afin de se défaire de la reconnaissance spécifique d'un gène de résistance de la plante ont désormais été réalisées (Gout *et al.*, 2007; Fudal *et al.*, 2009). La perte de l'avirulence A7 implique une diversité de mécanismes de mutation que l'on retrouve dans la perte de l'avirulence constatée par Fudal *et al.*, (2009). Un grand nombre d'évènements moléculaires à l'origine de la virulence a ainsi pu être répertorié : délétion partielle ou totale du gène, RIP, mutations ponctuelles, etc. En revanche, le contournement de *Rlm1* se distingue du fait de la large prédominance (> 98% des souches étudiées) d'un évènement unique de délétion du gène *AvrLm1* (Gout *et al.*, 2007). *AvrLm4-7* ayant une double spécificité d'interaction, l'analyse moléculaire des populations virulentes et avirulentes vis-à-vis de *Rlm7* a également permis d'étudier les mécanismes à l'origine de la virulence a4. Ainsi, la perte de l'avirulence A4 se distingue également des trois autres contournements moléculaires puisque une unique mutation ponctuelle non synonyme permettant à une souche A7 de perdre l'avirulence A4 a pu être répertorié et validé (Chapitre 1).

On constate donc qu'une grande variabilité existe dans la manière de perdre l'avirulence d'une souche pour un gène d'avirulence donné. Cette variabilité pourrait s'expliquer par plusieurs critères distinguant les gènes d'avirulence étudiés. D'une part, l'étude de Gout *et al.*, (2007) portant sur le contournement de *Rlm1* a été effectuée sur des souches prélevées 10 ans après le début de son contournement (Rouxel *et al.*, 2003) tandis que l'analyse moléculaire des souches virulentes vis-à-vis de *Rlm7* réalisée lors de ma thèse a permis de répertorier les tout premiers évènements de mutation menant à de la perte de l'avirulence A7. Ces deux études portent donc sur deux étapes distinctes du contournement des gènes de résistance. Pour *Rlm1*, le contournement a déjà eu lieu et une stabilisation des populations s'est effectuée tandis que pour *Rlm7*, nous observons les tous premiers évènements menant au contournement du gène. D'autre part, la mise au champ des différents gènes de résistance n'a pas été effectuée dans les mêmes conditions. En effet, tandis que la proportion de souches a7 représentait moins de 1% des populations françaises et européennes lors du déploiement des cultivars *Rlm7*, les souches a1 représentaient déjà entre 5 et 10% des souches lors du déploiement de *Rlm1*. L'évènement de délétion retrouvée dans la très grande majorité des souches a1 avait de plus déjà été retrouvé dès 1990. Cette proportion de souches virulentes a sans doute été un paramètre important dans la vitesse de contournement de *Rlm1*. Le

contournement d'un gène de résistance sera également grandement dépendant du déficit de fitness lié à la perte de l'avirulence correspondante. Ainsi, au contraire du contournement de *Rlm1* pour lequel la perte d'*AvrLm1* peut facilement être envisagé du fait du faible déficit de fitness associé à cette perte (Huang *et al.*, 2010), la perte de l'avirulence A7 est associée à un fort coût de fitness. Ce coût pourrait expliquer pourquoi une mutation ponctuelle permettant de conserver la fonction d'*AvrLm4-7* a été sélectionné pour le contournement de *Rlm4* (Huang *et al.*, 2006; Mongin, 2010). De la même façon, ce coût de virulence pourrait favoriser à l'avenir les évènements de mutation menant à la virulence a7 qui conserveraient la fonction de la protéine afin de ne pas subir ce coût de fitness.

Dans le cas d'*AvrLm4-7*, le suivi sur trois ans de la proportion des différents évènements de mutation à l'origine de la virulence a7 des souches a permis de constater que les mêmes catégories d'évènements étaient retrouvées d'une année sur l'autre mais que la fréquence des souches présentant une délétion d'*AvrLm4-7* augmentait fortement tandis que celle des souches présentant un allèle d'*AvrLm4-7* touché par le RIP diminuait. Cette diminution pourrait s'expliquer par la localisation génomique d'*AvrLm4-7*. En effet, *AvrLm4-7* ainsi que *AvrLm1* et *AvrLm6* sont localisés au sein « d'isochores AT » riches en éléments transposables répétés et dégénérés (Gout *et al.*, 2006 ; Fudal *et al.*, 2007; Chapitre1) qui seraient des cibles potentielles d'appariements non homologues lors de la méiose entre chromatides « sœurs ». Ces appariements résulteraient en un réarrangement chromosomique et en un polymorphisme de la taille des chromosomes (Plummer et Howlett, 1995). La présence des gènes d'avirulence et d'une partie des effecteurs de *L. maculans* permettrait un maintien sélectif de ces structures (Rouxel *et al.*, 2011) mais une fois inactivée (par RIP ou d'autres mutations), cette sélection positive ne s'exercerait plus et la région pourrait alors subir un remaniement sans conséquence pour la souche. La localisation des gènes d'avirulence favoriserait ainsi une réponse évolutive rapide sous pression de sélection (Rouxel *et al.*, 2011). Une telle hypothèse expliquerait la diminution de la proportion de souches affectées par le RIP et laisserait envisager une augmentation progressive de la fréquence des souches ayant subi la délétion d'*AvrLm4-7* dans les années à venir. En parallèle, l'hypothèse selon laquelle une mutation ponctuelle affectant la reconnaissance d'*AvrLm4-7* pourrait conserver la fonction de la protéine (résidus 100, 102, 112) suggère un tout autre scénario. De la même façon que l'avirulence A4 est perdue sans affecter l'avirulence A7 par la

mutation du résidu 120 de la protéine, une mutation ponctuelle affectant la reconnaissance d'AvrLm4-7 par Rlm7 mais conservant la fonction de la protéine pourrait être envisagée. Un tel scénario mènerait alors à une augmentation de la fréquence des mutations ponctuelles dans les souches virulentes.

Il aurait été intéressant de suivre le contournement de *Rlm7* pendant encore quelques années sur la parcelle expérimentale afin de constater si la proportion des différents évènements de mutation dans la population allait ou non évoluer en faveur d'un évènement particulier. Malheureusement, cette étude ne peut être envisagée : devant le niveau de souches virulentes observé au cours de la dernière année expérimentale (>35%), il a été décidé de détruire la parcelle et les résidus pour ne pas risquer de contamination des parcelles agricoles voisines. Un suivi national est actuellement réalisé par le CETIOM et permettra tout de même de voir l'évolution de ces différents évènements de mutation au sein de la population naturelle.

L'étude du contournement moléculaire de *Rlm7* (Chapitre 2) a également permis de montrer que l'allèle fonctionnel d'*AvrLm4-7* permettait à la souche de dissimuler la présence d'*AvrLm3*. Houterman *et al.*, (2008) ont montré dans le pathosystème *Fusarium oxysporum f. sp. lycopersici* / tomate que le gène d'avorulence *Avr1* était à l'origine de la suppression de la reconnaissance des produits des gènes d'avorulence *Avr2* et *Avr3*. *Avr1* serait la réponse évolutive sélectionnée chez le champignon afin de compenser la présence des gènes de résistance *I-2* et *I-3* chez la plante. De façon similaire, *AvrLm4-7* serait alors la réponse évolutive à la mise en place de *Rlm3* chez la plante permettant à la souche de contourner ce gène de résistance. Pour les gènes d'avorulence *AvrLm1*, *AvrLm4-7* et *AvrLm6*, nous avons pu voir que la perte de l'avorulence était liée à la mutation du gène. La mise en place d'un nouvel effecteur permettant à la souche d'éviter la reconnaissance d'un autre gène d'avorulence est inédite chez *L. maculans* et rarement observée chez les autres modèles.

De la même façon que la mutation ponctuelle du résidu 120 d'*AvrLm4-7* permettrait de conserver la fonctionnalité de la protéine tout en évitant la reconnaissance par Rlm4, *AvrLm4-7* permettrait à la souche de ne plus être détectée tout en conservant la fonction d'*AvrLm3*. Cette stratégie d'évitement de la reconnaissance laisserait alors supposer qu'*AvrLm3* aurait un rôle important pour la souche. Cette hypothèse est soutenue par la localisation génomique d'*AvrLm3* et sa

conservation dans les populations pathogènes. En effet, la localisation en région télomérique de certains gènes tel *Avr-Pita* chez *M. oryzae*, semble favoriser la délétion de ceux-ci (Orbach *et al.*, 2000). La présence d'*AvrLm3* en région télomérique aurait donc dû favoriser la perte de ce gène. Le maintien dans la majorité des souches françaises d'*AvrLm3* traduit donc une sélection positive de ce gène suggérant une certaine importance de celui-ci pour l'accomplissement du cycle infectieux de *L. maculans* qui reste encore à caractériser.

Au sein de la course aux armements que se livrent *L. maculans* et son hôte, nous pouvons désormais scénariser la chronologie de la mise en place des gènes *AvrLm3*, *AvrLm4-7* et de leurs gènes de résistance correspondant. En présence d'une forte proportion dans les populations pathogènes du gène *AvrLm3*, *Rlm3* aurait été sélectionné de telle façon à le reconnaître. *L. maculans* a ensuite regagné sa capacité à se développer par l'acquisition d'*AvrLm4-7* qui a lui même été pris pour cible par la suite par les gènes de résistance *Rlm4* et *Rlm7*. *Rlm4* a été en premier sélectionné et utilisé durant de nombreuses années et a conduit à la perte de l'avirulence A4. Le déficit de fitness relatif à la perte d'*AvrLm4-7* aurait conduit à la sélection d'un évènement de mutation permettant de perdre l'avirulence A4 tout en conservant la fonctionnalité de la protéine. Dernièrement (2004), *AvrLm4-7* étant toujours fonctionnel malgré la perte de l'avirulence A4, le gène *Rlm7* a été sélectionné et commercialisé, ma thèse montrant alors l'étape ultérieure dans cette course aux armements avec les différents évènements de mutation permettant de s'affranchir de la résistance *Rlm7*. L'étude de Dilmaghani *et al.*, (2009) a montré que la situation de *L. maculans* au centre du Canada dénotait de la situation française et européenne. En effet, à la différence des populations européennes, les cultivars utilisés dans cette région ont été le plus souvent des cultivars dépourvus de gène de résistance spécifique mais néanmoins, des souches A3a7 sont retrouvées en grande proportion (Dilmaghani *et al.*, 2009). La présence de souches A3a7 au sein de ces populations sans qu'une pression de sélection due à *Rlm7* n'ait eu lieu serait peut-être l'illustration de l'étape de la course aux armements pour laquelle *AvrLm4-7* n'aurait pas encore été mis en place chez ces souches.

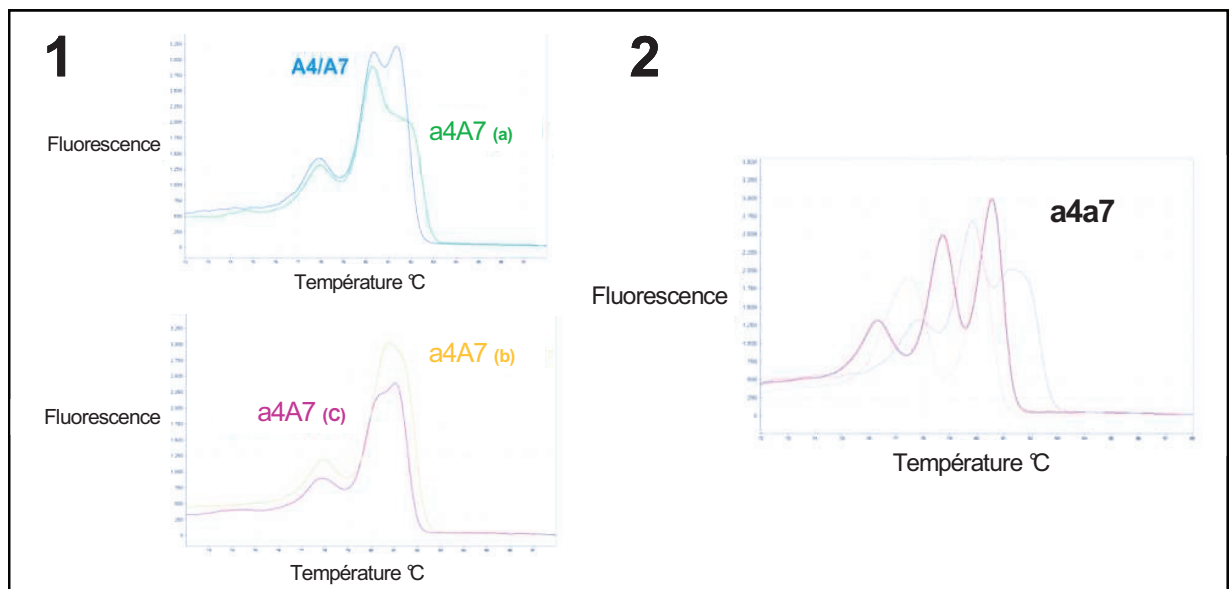
La découverte de nouveaux gènes de résistance, leur sélection ainsi que le développement de cultivars possédant ces résistances est difficile et induit un coût

important pour les sélectionneurs. Il est alors indispensable d'essayer d'augmenter la durabilité des gènes de résistance actuellement connus et utilisés. Ce constat est malheureusement contradictoire avec la tendance actuelle de l'agriculture de privilégier un travail du sol simplifié qui, sans la rotation des cultures encore appliquée, serait relativement proche des conditions de culture du site de Grignon. Les résultats obtenus au cours de cette thèse donnent des éléments objectifs pour proposer des stratégies de gestion de ces résistances.

En effet il a été démontré que les pratiques culturales adoptées au champ et en particulier la gestion des résidus de colza ainsi que l'alternance des cultures sont de première importance dans le contournement de *Rlm7*. Si un tel conseil est formulé depuis de nombreuses années (Alabouvette et Brunin, 1970), nous apportons ici la démonstration de l'effet réel de ces préconisations. De plus, nous proposons une hypothèse explicative de cet effet (limitation de la reproduction sexuée ayant une influence directe sur l'hypermutableté du gène d'avirulence et réduction de la taille des populations par travail du sol).

La gestion raisonnée des résistances passe également par un choix variétal adapté aux populations de l'agent pathogène. En ce sens, la résistance *Rlm3*, bien que présente dans plus de 20% des variétés commercialisées (Pinochet et Balesdent, comm. pers.) était considérée comme une résistance inutile puisque plus de 99% des souches françaises sont capables d'attaquer de telles variétés (Balesdent *et al.*, 2006). Cependant, à la lumière de mes travaux, il semblerait que, bien que masqué par la présence d'*AvrLm4-7*, le gène *AvrLm3* soit présent à fréquence très élevée dans les populations françaises puisque 1,7% seulement des souches virulentes a7 sont également virulentes a3, contre 98,3% de souches avirulentes. Dès lors, plusieurs stratégies de gestion des résistances peuvent être proposées : soit un pyramidage des gènes *Rlm3* et *Rlm7*, soit une alternance spatiale ou temporelle des deux gènes, l'évaluation de la stratégie la plus durable restant encore à déterminer.

La gestion raisonnée des résistances nécessite également de pouvoir surveiller l'évolution de la proportion des souches virulentes dans la population pathogène. Cette surveillance passe « traditionnellement » par la détermination en laboratoire du phénotype des différents isolats trouvés au champ à l'aide de tests d'inoculation sur plantes. Ces tests présentent le désavantage d'être relativement



**Figure 4 (discussion générale) : représentation graphique des courbes de fusion caractéristiques de différentes souches a7 et A7 de *L. maculans* obtenues par HRM**

**1:** Courbes de fusion des quatre allèles A7 les plus fréquemment rencontrés (voir tableau 3 Chapitre 2 : A4A7: v23.1.3 ; (a) : v23.1.2; (b) : G08-C1049 ; (C) : G06-485). **2 :** Courbes de fusion de différents allèles virulent vis-à-vis de *Rlm7*.

longs (1 mois à partir de l'obtention des spores servant à l'inoculation) et nécessitent de disposer d'une infrastructure spécialisée pour le développement des plantes ainsi qu'une collection de cultivars permettant la différenciation des phénotypes recherchés. Afin d'alléger cette procédure, des outils moléculaires se basant sur une connaissance approfondie des différents événements moléculaires à l'origine de la perte de l'avirulence d'une souche peuvent être développés. Ainsi, du fait de l'extrême prédominance d'un événement de délétion à l'origine de la virulence des souches vis-à-vis de *Rlm1*, l'amplification d'*AvrLm1* par PCR permet de déterminer la virulence d'une souche vis-à-vis de *Rlm1* avec une marge d'erreur relativement faible. La validation de l'implication du résidu 120 dans l'interaction AvrLm4-Rlm4 a également débouché sur une détection moléculaire, en combinant amplification du gène et digestion enzymatique ciblant la mutation (CAPS – Chapitre 1).

Dans le cas de la résistance *Rlm7*, le travail effectué au cours de ma thèse a permis de constater que la détermination de la virulence d'une souche vis-à-vis de *Rlm7* à l'aide d'une PCR n'était pas possible puisque celle-ci n'aurait pas permis de détecter une partie importante des souches virulentes (souches touchées par le RIP, par des délétions partielles, par des mutations ponctuelles, etc.). L'amplification PCR simple ne pouvant pas être envisagée, une analyse par « High Resolution Melting » (HRM) a alors été développée par le CETIOM (Centre Technique Interprofessionnel des Oléagineux Métropolitains) sur la base de mes résultats. La HRM est une technique qui permet de différencier des produits PCR selon leur température de fusion. Chaque produit PCR ayant un « profil HRM » (courbe de fusion) qui lui est propre, les différents allèles testés peuvent être comparés et différenciés (Figure 4, discussion générale). Pour valider cette méthode, 100 souches avirulentes et 100 souches virulentes provenant de la collection de souches établie à l'occasion de ma thèse et représentatives de l'éventail d'événements d'acquisition de la virulence ont été analysées par HRM au CETIOM. Il a été déterminé que cette méthode, dans un contexte où la fréquence des souches virulentes a7 atteignait 5%, pouvait être utilisée pour caractériser les isolats avec un résultat fiable dans plus de 99% des cas. Ce travail donnera lieu à une présentation orale lors du 13<sup>ème</sup> "International Rapeseed Congress" (Prague, 5-9 Juin 2011) (Annexe 1). L'utilisation de cette méthode à l'avenir permettra non seulement un suivi plus efficace et plus rapide des populations virulentes, mais devrait également permettre de déceler l'apparition de nouveaux profils HRM, qui en retour pourraient mettre en évidence des mutations



ponctuelles ciblant des sites fonctionnels ou de reconnaissance importants dans le gène *AvrLm4-7*.

# **Annexe**



# Annexe 1 : résumé de la présentation orale portant sur l'utilisation de la HRM pour la détection des souches virulentes a7 de *L. maculans* (Juin 2011, Prague)

## 13th International Rapeseed Congress

---

*AvrLm7* allelic diversity in *Leptosphaeria maculans* nevertheless allows survey of virulent isolates using the HRM (High Melting Resolution) method

Guillaume Daverdin<sup>1</sup>, Julien Carpezat<sup>2</sup>, Martine Leflon<sup>2</sup>, Xavier Pinochet<sup>2</sup>, Thierry Rouxel<sup>1</sup>, Marie-Hélène Balesdent<sup>1</sup>

<sup>1</sup>INRA-UMR1290, BIOGER-CPP, <sup>2</sup>CETIOM; Avenue Lucien Bretignières, 78850, Thiverval Grignon, France.

*Leptosphaeria maculans* is responsible for the damaging stem canker (blackleg) disease of *Brassica napus* (oilseed rape) in Europe, Canada and Australia. This disease is mainly controlled by host genetic resistance. Specific, major gene (*Rlm*) resistance is effective when fungal populations are mainly composed of avirulent isolates, i.e. having the avirulent allele of the corresponding avirulence gene (*AvrLm*). In a couple of agronomic situations, major gene resistance became ineffective after two to three growing seasons due to the rapid increase of virulent allele frequencies in *L. maculans* populations. A rationale deployment and management of *Rlm* genes in commercial cultivars necessitates the survey of virulent isolates in field populations. For that purpose, molecular tools to identify virulent isolates toward a given *Rlm* gene are needed and the first step is to determine which mutational events are responsible for the gain of virulence in the pathogen.

In France, the *Rlm7* gene has been firstly used commercially in 2003 and became widely used after 2004 with the release of *Rlm7* hybrid varieties, in a context where more than 99% of the isolates possessed the avirulent allele *AvrLm7*. To identify the mutations responsible for the gain of virulence toward *Rlm7* (*avrLm7* isolates), a large-scale sampling was performed over 3 years (2006-2008) in a single field located at Grignon (France) sown with the *Rlm7* cv. Exagone. Isolates were recovered either from leaf spots in autumn or from stem residues the following autumn. Around 1500 isolates were obtained and phenotyped as *AvrLm7* or *avrLm7*

following inoculation on a *Rlm7* line. The *AvrLm7* locus was amplified and the PCR product, whenever amplified, was sequenced.

A wide range of mutations were found to be responsible for the gain of virulence toward *Rlm7*, encompassing point mutations leading to changes in the amino-acid sequence of the protein, single or di-nucleotide deletions, insertion of transposons in the coding sequence, partial or whole deletion of the gene, or accumulation of mutations through RIP (Repeat-Induced Point mutation) leading to the introduction of stop codons in the gene. In contrast, a very limited nucleotidic polymorphism was observed for avirulent, *AvrLm7*, alleles.

Because of the diversity of events leading to the virulence toward *Rlm7*, it appeared impossible to characterize isolates at the *AvrLm7* locus using conventional PCR methods. A method, based on High Resolution Melting (HRM) analysis was therefore developed. HRM is a technique that characterizes PCR products based on their melting behaviour during their transition from double-stranded DNA to single-stranded DNA with increasing temperature. At some point during the process of warming, which depends on the sequence, length, GC content or strand complementarity, the melting temperature of the amplicon is reached and the two strands of DNA separate. HRM was assessed on 100 *AvrLm7* and 100 *avrLm7* isolates. In this study, the avirulent isolates showed only four different melting profiles and 87% of virulent isolates showed either no amplification, or melting profiles distinct from those of avirulent isolates. Considering that less than 5% of isolates are currently virulent in French populations, this method could be used to characterise field isolates with a result reliable for more than 99% of isolates.

# **Références** **bibliographiques**



- Agrios, G. N.** (1997). Plant Pathology, fourth edition ed. Academic Press, New York.
- Alabouvette, C. and Brunin, B.** (1970). Recherches sur la maladie du colza due à *Leptosphaeria maculans* (Desm.) CES. et de NOT. 1. Rôle des restes de culture dans la conservation et la dissémination du paasite. *Annu. Rev. Phytopathol.* **2**: 463-475.
- Anderson, J. P., Gleason, C. A., Foley, R. C., Thrall, P. H., Burdon, J. B. and Singh, K. B.** (2010). Plants versus pathogens: an evolutionary arms race. *Functional Plant Biology* **37** (6): 499-512.
- Anonymous** (2010). Légère progression des surfaces en colza. *Semences et Progrès.* **147**: 5.
- Ansan-Melayah, D., Balesdent, M. H., Buée, M. and Rouxel, T.** (1995). Genetic characterization of *AvrLm1*, the first avirulence gene of *Leptosphaeria maculans*. *Phytopathology* **85**: 1525-1529.
- Ansan-Melayah, D., Rouxel, T., Bertrand, J., Letarnec, B., Mendes-Pereira, E. and Balesdent, M. H.** (1997). Field efficiency of Brassica napus specific resistance correlates with *Leptosphaeria maculans* population structure. *European Journal of Plant Pathology* **103**: 835-841.
- Ansan-Melayah, D., Balesdent, M. H., Delourme, R., Pilet, M. L., Tanguy, X., Renard, M. and Rouxel, T.** (1998). Genes for race-specific resistance against blackleg disease in *Brassica napus* L. *Plant Breeding* **117**: 373-378.
- Aubertot, J. N., Pinochet, X. and Dore, T.** (2004). The effects of sowing date and nitrogen availability during vegetative stages on *Leptosphaeria maculans* development on winter oilseed rape. *Crop Protection* **23**: 635-645.
- Axtell, M. J. and Staskawicz, B. J.** (2003). Initiation of RPS2-specified disease resistance in *Arabidopsis* is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**: 369-377.
- Balesdent, M. H., Attard, A., Kuhn, A. L. and Rouxel, T.** (2002). New avirulence genes in the phytopathogenic fungus *Leptosphaeria maculans*. *Phytopathology* **92**: 1122-1133.
- Balesdent, M. H., Barbetti, M. J., Li, H., Sivasithamparam, K., Gout, L. and Rouxel, T.** (2005). Analysis of *Leptosphaeria maculans* race structure in a worldwide collection of isolates. *Phytopathology* **95**: 1061-1071.
- Balesdent, M. H., Louvard, K., Pinochet, X. and Rouxel, T.** (2006). A large-scale survey of races of *Leptosphaeria maculans* occurring on oilseed rape in France. *European Journal of Plant Pathology* **114**: 53-65.
- Bent, A. F. and Mackey, D.** (2007). Elicitors, Effectors, and R Genes: The New Paradigm and a Lifetime Supply of Questions. *Annual Review of Phytopathology* **45**: 399-436.
- Bergelson, J., Dwyer, G. and Emerson, J. J.** (2001). Models and data on plant-enemy coevolution. *Annual Review of Genetics* **35**: 469-499.
- Bhattacharjee, S., Hiller, N. L., Liolios, K., Win, J., Kanneganti, T. D., Young, C., Kamoun, S. and Haldar, K.** (2006). The malarial host-targeting signal is conserved in the Irish potato famine pathogen. *PloS Pathogens* **2**: 453-465.
- Bohman, S., Staal, J., Thomma, B., Wang, M. L. and Dixelius, C.** (2004). Characterisation of an *Arabidopsis-Leptosphaeria maculans* pathosystem: resistance partially requires camalexin biosynthesis and is independent of salicylic acid, ethylene and jasmonic acid signalling. *The Plant Journal* **37**: 9-20.
- Bohnert, H. U., Fudal, I., Dioh, W., Tharreau, D., Notteghem, J. L. and Lebrun, M. H.** (2004). A putative polyketide synthase peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. *Plant Cell* **16**: 2499-2513.
- Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B. and Klein, A.** (2002). The barley stem rust-resistance gene Rpg1



- is a novel disease-resistance gene with homology to receptor kinases. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 9328-9333.
- Brun, H., Chèvre, A.-M., Fitt, B. D., Powers, S., Besnard, A.-L., Ermel, M., Huteau, V., Marquer, B., Eber, F., Renard, M. and Andrivon, D.** (2010). Quantitative resistance increases the durability of qualitative resistance to *Leptosphaeria maculans* in *Brassica napus*. *New Phytologist* **185**: 285-299.
- Buschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., vanDaelen, R., vanderLee, T., Diergaarde, P., Groenendijk, J., Topsch, S., Vos, P., Salamini, F. and Schulze-Lefert, P.** (1997). The barley mlo gene: A novel control element of plant pathogen resistance. *Cell* **88**: 695-705.
- Catanzariti, A. M., Dodds, P. N., Lawrence, G. J., Ayliffe, M. A. and Ellis, J. G.** (2006). Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell* **18**: 243-256.
- Chisholm, S. T., Coaker, G., Day, B. and Staskawicz, B. J.** (2006). Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell* **124**: 803-814.
- da Cunha, L., McFall, A. J. and Mackey, D.** (2006). Innate immunity in plants: a continuum of layered defenses. *Microbes and Infection* **8**: 1372-1381.
- Dangl, J. L. and Jones, J. D. G.** (2001). Plant pathogens and integrated defence responses to infection. *Nature* **411** 826-833.
- Delourme, R., Piel, N., Horvais, R., Pouilly, N., Domin, C., Vallee, P., Falentin, C., Manzanares-Dauleux, M. J. and Renard, M.** (2008). Molecular and phenotypic characterization of near isogenic lines at QTL for quantitative resistance to *Leptosphaeria maculans* in oilseed rape (*Brassica napus* L.). *Theoretical and Applied Genetics* **117**: 1055-1067.
- Deslandes, L., Olivier, J., Theulieres, F., Hirsch, J., Feng, D. X., Bittner-Eddy, P., Beynon, J. and Marco, Y.** (2002). Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by the recessive RRS1-R gene, a member of a novel family of resistance genes. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 2404-2409.
- Deslandes, L., Olivier, J., Peeters, N., Feng, D. X., Khounlotham, M., Boucher, C., Somssich, L., Genin, S. and Marco, Y.** (2003). Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 8024-8029.
- Dilmaghani, A., Balesdent, M. H., Didier, J. P., Wu, C., Davey, J., Barbetti, M. J., Li, H., Moreno-Rico, O., Phillips, D., Despeghel, J. P., Vincenot, L., Gout, L. and Rouxel, T.** (2009). The *Leptosphaeria maculans*-*Leptosphaeria biglobosa* species complex in the American continent. *Plant Pathology* **58**: 1044-1058.
- Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Teh, T., Wang, C. I. A., Ayliffe, M. A., Kobe, B. and Ellis, J. G.** (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 8888-8893.
- Dodds, P. N., Rafiqi, M., Gan, P. H. P., Hardham, A. R., Jones, D. A. and Ellis, J. G.** (2009). Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytologist* **183**: 993-999.
- Donnelly, J. S.** (2001). *The Great Irish Potato Famine*, Sutton Publishing.
- Ellis, J. G., Dodds, P. N. and Lawrence, G. J.** (2007). Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annual Review of Phytopathology* **45**: 289-306.

- Eyre-Walker, A. and Hurst, L. D.** (2001). The evolution of isochores. *Nature Reviews Genetics* **2**: 549-555.
- Farman, M. L.** (2007). Telomeres in the rice blast fungus *Magnaporthe oryzae*: the world of the end as we know it. *Fems Microbiology Letters* **273**: 125-132.
- Felix, G., Duran, J. D., Volko, S. and Boller, T.** (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal* **18**: 265-276.
- Fitt, B. D. L., Hu, B. C., Li, Z. Q., Liu, S. Y., Lange, R. M., Kharbanda, P. D., Butterworth, M. H. and White, R. P.** (2008). Strategies to prevent spread of *Leptosphaeria maculans* (phoma stem canker) onto oilseed rape crops in China; costs and benefits. *Plant Pathology* **57**: 652-664.
- Flor, H. H.** (1971). Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology* **9**: 275-296.
- Friesen, T. L., Meinhardt, S. W. and Faris, J. D.** (2007). The *Stagonospora nodorum*-wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. *The Plant Journal* **51**: 681-692.
- Fudal, I., Bohnert, H. U., Tharreau, D. and Lebrun, M. H.** (2005). Transposition of MINE, a composite retrotransposon, in the avirulence gene ACE1 of the rice blast fungus *Magnaporthe grisea*. *Fungal Genetics and Biology* **42**: 761-772.
- Fudal, I., Ross, S., Gout, L., Blaise, F., Kuhn, M. L., Eckert, M. R., Cattolico, L., Bernard-Samain, S., Balesdent, M. H. and Rouxel, T.** (2007). Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria maculans* Genome: Map-Based Cloning of AvrLm6. *Molecular Plant-Microbe Interactions* **20**: 459-470.
- Fudal, I., Ross, S., Brun, H., Besnard, A. L., Ermel, M., Kuhn, M. L., Balesdent, M. H. and Rouxel, T.** (2009). Repeat-Induced Point Mutation (RIP) as an alternative mechanism of evolution toward virulence in *Leptosphaeria maculans*. *Molecular Plant-Microbe Interactions* **22**: 932-941.
- Galyov, E. E., Hakansson, S. and Wolfwatz, H.** (1994). Characterization of the operon encoding the YPKA SER/THR protein kinase and the YOPJ protein of *Yersinia pseudotuberculosis*. *Journal of Bacteriology* **176**: 4543-4548.
- Gassmann, W., Dahlbeck, D., Chesnokova, O., Minsavage, G. V., Jones, J. B. and Staskawicz, B. J.** (2000). Molecular evolution of virulence in natural field strains of *Xanthomonas campestris* pv. *vesicatoria*. *Journal of Bacteriology* **182**: 7053-7059.
- Gautier, C.** (2000). Compositional bias in DNA. *Current Opinion in Genetics & Development* **10**: 656-661.
- Glazebrook, J.** (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology* **43**: 205-227.
- Gomez-Gomez, L. and Boller, T.** (2000). FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Molecular Cell* **5**: 1003-1011.
- Gomez-Gomez, L. and Boller, T.** (2002). Flagellin perception: a paradigm for innate immunity. *Trends in Plant Science* **7**: 251-256.
- Gout, L., Fudal, I., Kuhn, M. L., Blaise, F., Eckert, M., Cattolico, L., Balesdent, M. H. and Rouxel, T.** (2006). Lost in the middle of nowhere: the *AvrLm1* avirulence gene of the Dothideomycete *Leptosphaeria maculans*. *Molecular Microbiology* **60**: 67-80.
- Gout, L., Kuhn, M. L., Vincenot, L., Bernard-Samain, S., Cattolico, L., Barbetti, M., Moreno-Rico, O., Balesdent, M. H. and Rouxel, T.** (2007). Genome structure

- impacts molecular evolution at the *AvrLm1* avirulence locus of the plant pathogen *Leptosphaeria maculans*. *Environmental Microbiology* **9**: 2978-2992.
- Gu, K. Y., Yang, B., Tian, D. S., Wu, L. F., Wang, D. J., Sreekala, C., Yang, F., Chu, Z. Q., Wang, G. L., White, F. F. and Yin, Z. C.** (2005). R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**: 1122-1125.
- Hammond-Kosack, K. E. and Parker, J. E.** (2003). Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. *Current Opinion in Biotechnology* **14**: 177-193.
- Heath, M. C.** (2000). Nonhost resistance and nonspecific plant defenses. *Current Opinion in Plant Biology* **3**: 315-319.
- Hogenhout, S. A., Van der Hoorn, R. A. L., Terauchi, R. and Kamoun, S.** (2009). Emerging Concepts in Effector Biology of Plant-Associated Organisms. *Molecular Plant-Microbe Interactions* **22**: 115-122.
- Houterman, P. M., Cornelissen, B. J. C. and Rep, M.** (2008). Suppression of plant resistance gene-based immunity by a fungal effector. *PloS Pathogens* **4** (5): 6.
- Huang, Y. J., Li, Z. Q., Evans, N., Rouxel, T., Fitt, B. D. L. and Balesdent, M. H.** (2006). Fitness cost associated with loss of the *AvrLm4* avirulence function in *Leptosphaeria maculans* (Phoma stem canker of oilseed rape). *European Journal of Plant Pathology* **114**: 77-89.
- Huang, Y. J., Balesdent, M. H., Li, Z. Q., Evans, N., Rouxel, T. and Fitt, B. D. L.** (2010). Fitness cost of virulence differs between the *AvrLm1* and *AvrLm4* loci in *Leptosphaeria maculans* (phoma stem canker of oilseed rape). *European Journal of Plant Pathology* **126**: 279-291.
- Idnurm, A. and Howlett, B. J.** (2003). Analysis of loss of pathogenicity mutants reveals that repeat-induced point mutations can occur in the Dothideomycete *Leptosphaeria maculans*. *Fungal Genetics and Biology* **39** (1): 31-37.
- Jia, Y., McAdams, S. A., Bryan, G. T., Hershey, H. P. and Valent, B.** (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *Embo Journal* **19**: 4004-4014.
- Johnson, R.** (1984). A critical analysis of durable resistance. *Annual Review of Phytopathology* **22**: 309-330.
- Jones, J. D. G. and Dangl, J. L.** (2006). The plant immune system. *Nature* **444**: 323-329.
- Joosten, M. H. A. J. and de Wit, P. J. G. M.** (1999). The tomato - *Cladosporium fulvum* interaction: A versatile experimental system to study plant-pathogen interactions. *Annual Review of Phytopathology* **37**: 335-367.
- Kale, S. D., Gu, B. A., Capelluto, D. G. S., Dou, D. L., Feldman, E., Rumore, A., Arredondo, F. D., Hanlon, R., Fudal, I., Rouxel, T., Lawrence, C. B., Shan, W. X. and Tyler, B. M.** (2010). External Lipid PI3P Mediates Entry of Eukaryotic Pathogen Effectors into Plant and Animal Host Cells. *Cell* **142**: 284-295.
- Kang, S., Lebrun, M. H., Farrall, L. and Valent, B.** (2001). Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. *Molecular Plant-Microbe Interactions* **14**: 671-674.
- Kay, S., Hahn, S., Marois, E., Hause, G. and Bonas, U.** (2007). A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**: 648-651.
- Keen, N. T.** (1982). Specific recognition in gene-for-gene host-parasite systems. *Advances in Plant Pathology* **1**: 35-82.
- Knogge, W.** (1996). Fungal infection of plants. *Plant Cell* **8** (10): 1711-1722.
- Lahaye, T. and Bonas, U.** (2001). Molecular secrets of bacterial type III effector proteins. *Trends in Plant Science* **6**: 479-485.

- Leach, J. E., Cruz, C. M. V., Bai, J. F. and Leung, H.** (2001). Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annual Review of Phytopathology* **39**: 187-224.
- Lee, S. W., Han, S. W., Bartley, L. E. and Ronald, P. C.** (2006). Unique characteristics of *Xanthomonas oryzae* pv. *oryzae* AvrXa21 and implications for plant innate immunity. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 18395-18400.
- Luderer, R., Rivas, S., Nurnberger, T., Mattei, B., Van den Hooven, H. W., Van der Hoorn, R. A. L., Romeis, T., Wehrfritz, J. M., Blume, B., Nennstiel, D., Zuidema, D., Vervoort, J., De Lorenzo, G., Jones, J. D. G., de Wit, P. J. G. M. and Joosten, M. H. A. J.** (2001). No evidence for binding between resistance gene product Cf-9 of tomato and avirulence gene product AVR9 of *Cladosporium fulvum*. *Molecular Plant-Microbe Interactions* **14**: 867-876.
- Luderer, R., Takken, F. L. W., de Wit, P. J. G. M. and Joosten, M. H. A. J.** (2002). *Cladosporium fulvum* overcomes Cf-2-mediated resistance by producing truncated AVR2 elicitor proteins. *Molecular Microbiology* **45**: 875-884.
- Mackey, D., Holt, B. F., Wiig, A. and Dangl, J. L.** (2002). RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell* **108**: 743-754.
- Mackey, D., Belkhadir, Y., Alonso, J. M., Ecker, J. R. and Dangl, J. L.** (2003). *Arabidopsis* RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell* **112**: 379-389.
- Marti, M., Good, R. T., Rug, M., Knuepfer, E. and Cowman, A. F.** (2004). Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**: 1930-1933.
- Martin, G. B., Bogdanove, A. J. and Sessa, G.** (2003). Understanding the functions of plant disease resistance proteins. *Annual Review of Plant Biology* **54**: 23-61.
- McDonald, B. A. and Linde, C.** (2002). Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**: 349-379.
- McIntosh, R. A. and Brown, G. N.** (1997). Anticipatory breeding for resistance to rust diseases in wheat. *Annual Review of Phytopathology* **35**: 311-326.
- Mendes-Pereira, E., Balesdent, M. H., Brun, H. and Rouxel, T.** (2003). Molecular phylogeny of the *Leptosphaeria maculans*-*L. biglobosa* species complex. *Mycological Research* **107**: 1287-1304.
- Mongin, T.** (2010). Etude des gènes impliqués dans l'agressivité de *Leptosphaeria maculans*, champignon pathogène du colza : rôle du gène d'avirulence *AvrLm4-7* et recherche de QTLs de l'agressivité. Rapport de Master 2 « Ecosystemes Terrestres et Action de l'homme », Orléans, France, Université d'Orléans.
- Morgan, W. and Kamoun, S.** (2007). RXLR effectors of plant pathogenic oomycetes. *Current Opinion in Microbiology* **10**: 332-338.
- Mucyn, T. S., Clemente, A., Andriotis, V. M. E., Balmuth, A. L., Oldroyd, G. E. D., Staskawicz, B. J. and Rathjen, J. P.** (2006). The tomato NBARC-LRR protein Prf interacts with Pto kinase in vivo to regulate specific plant immunity. *Plant Cell* **18**: 2792-2806.
- Mudgett, M. B.** (2005). New insights to the function of phytopathogenic bacterial type III effectors in plants. *Annual Review of Plant Biology* **56**: 509-531.
- Mundt, C. C.** (2002). Use of multiline cultivars and cultivar mixtures for disease management. *Annual Review of Phytopathology* **40**: 381-410.
- Mysore, K. S. and Ryu, C. M.** (2004). Nonhost resistance: how much do we know? *Trends in Plant Science* **9**: 97-104.

- Niks, R. E.** (1987). Nonhost Plant-species as donors for resistance to pathogens with narrow host range .1. Determination of nonhost status *Euphytica* **36**: 841-852.
- Nurnberger, T., Brunner, F., Kemmerling, B. and Piater, L.** (2004). Innate immunity in plants and animals: striking similarities and obvious differences. *Immunological Reviews* **198**: 249-266.
- Oliver, R. P. and Solomon, P. S.** (2010). New developments in pathogenicity and virulence of necrotrophs. *Current Opinion in Plant Biology* **13**: 415-419.
- Orbach, M. J., Farrall, L., Sweigard, J. A., Chumley, F. G. and Valent, B.** (2000). A telomeric avirulence gene determines efficacy for the rice blast resistance gene Pi-ta. *Plant Cell* **12**: 2019-2032.
- Orth, K., Xu, Z. H., Mudgett, M. B., Bao, Z. Q., Palmer, L. E., Bliska, J. B., Mangel, W. F., Staskawicz, B. and Dixon, J. E.** (2000). Disruption of signaling by *Yersinia* effector YopJ, a ubiquitin-like protein protease. *Science* **290**: 1594-1597.
- Parisi, L., Lespinasse, Y., Guillaumes, J. and Kruger, J.** (1993). A new race of *Venturia inaequalis* virulent to apples with resistance due to the Vf gene. *Phytopathology* **83**: 533-537.
- Parker, J. E., Coleman, M. J., Szabo, V., Frost, L. N., Schmidt, R., vanderBiezen, E. A., Moores, T., Dean, C., Daniels, M. J. and Jones, J. D. G.** (1997). The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6. *Plant Cell* **9**: 879-894.
- Parlange, F.** (2006). Analyse de l'environnement et clonage du gène *AvrLm4-7* chez l'ascomycète *Leptosphaeria maculans*. Thèse de doctorat, Orsay, France : Université Paris-Sud XI.
- Pinochet, X., Mestries, E., Penaud, A., Delourme, R., Chèvre, A. M., Renard, M., Brun, H., Bousset, L., Balesdent, M. H., Rouxel, T. and Aubertot, J. N.** (2003). Towards a durable management of genetic resistances to *Leptosphaeria maculans*. *Ocl-Oleagineux Corps Gras Lipides* **10**: 208-211.
- Pitman, A. R., Jackson, R. W., Mansfield, J. W., Kaitell, V., Thwaites, R. and Arnold, D. L.** (2005). Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. *Current Biology* **15**: 2230-2235.
- Poland, J. A., Balint-Kurti, P. J., Wissler, R. J., Pratt, R. C. and Nelson, R. J.** (2009). Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science* **14**: 21-29.
- Rapilly, F.** (1998). Évolution historique des concepts de résistance des plantes aux agents pathogènes. *Cahiers Agricultures* **7**: 223-8.
- Rep, M., van der Does, H. C., Meijer, M., van Wijk, R., Houterman, P. M., Dekker, H. L., de Koster, C. G. and Cornelissen, B. J. C.** (2004). A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for I-3-mediated resistance in tomato. *Molecular Microbiology* **53**: 1373-1383.
- Rep, M.** (2005). Small proteins of plant-pathogenic fungi secreted during host colonization. *Fems Microbiology Letters* **253** (1): 19-27.
- Rep, M., Meijer, M., Houterman, P. M., van der Does, H. C. and Cornelissen, B. J. C.** (2005). *Fusarium oxysporum* evades I-3-mediated resistance without altering the matching avirulence gene. *Molecular Plant-Microbe Interactions* **18**: 15-23.
- Ridout, C. J., Skamnioti, P., Porritt, O., Sacristan, S., Jones, J. D. G. and Brown, J. K. M.** (2006). Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance. *Plant Cell* **18**: 2402-2414.
- Rimmer, S. R.** (2006). Resistance genes to *Leptosphaeria maculans* in *Brassica napus*. *Canadian Journal of Plant Pathology* **28**: S288-S297.

- Ritter, C. and Dangl, J. L. (1995). The AVR/RPM1 gene of *Pseudomonas syringae* pv *maculicola* is required for virulence on *Arabidopsis*. *Molecular Plant-Microbe Interactions* **8**: 444-453.
- Rohe, M., Gierlich, A., Hermann, H., Hahn, M., Schmidt, B., Rosahl, S. and Knogge, W. (1995). The race specific elicitor NIP1 from the barley pathogen *Rhynchosporium secalis* determines avirulence on host plants of the RRS1 resistance genotype. *Embo Journal* **14**: 4168-4177.
- Rooney, H. C. E., van 't Klooster, J. W., van der Hoorn, R. A. L., Joosten, M. H. A. J., Jones, J. D. G. and de Wit, P. J. G. M. (2005). *Cladosporium* Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science* **308**: 1783-1786.
- Rouxel, T., Penaud, A., Pinochet, X., Brun, H., Gout, L., Delourme, R., Schmit, J. and Balesdent, M. H. (2003). A 10-year survey of populations of *Leptosphaeria maculans* in France indicates a rapid adaptation towards the Rlm1 resistance gene of oilseed rape. *European Journal of Plant Pathology* **109**: 871-881.
- Rouxel, T., Mendes-Pereira, E., Brun, H. and Balesdent, M. H. (2004). Species complex of fungal phytopathogens: the *Leptosphaeria maculans*-*L. biglobosa* case study, *Plant Genome: Biodiversity and Evolution*. **2**: 33-75.
- Rouxel, T. and Balesdent, M. H. (2005). The stem canker (blackleg) fungus, *Leptosphaeria maculans*, enters the genomic era. *Molecular Plant Pathology* **6**: 225-241.
- Rouxel, T. and Balesdent, M. H. (2010). Avirulence Genes, John Wiley & Sons, Ltd.
- Rouxel, T., Grandaubert, J., Hane, J. K., Hoede, C., van de Wouw, A. P., Couloux, A., Dominguez, V., Anthouard, V., Bally, P., Bourras, S., Cozijnsen, A. J., Ciuffetti, L. M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin, S. B., Gout, L., Glaser, N., Linglin, J., Kema, G. H. J., Lapalu, N., Lawrence, C. B., May, K., Meyer, M., Ollivier, B., Poulain, J., Schoch, C. L., Simon, A., Spatafora, J. W., Stachowiak, A., Turgeon, B. G., Tyler, B. M., Vincent, D., Weissenbach, J., Amselem, J., Quesneville, H., Oliver, R. P., Wincker, P., Balesdent, M. H. and Howlett, B. J. (2011). Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat Induced Point mutations. *Nature Communication* (in press).
- Schoch, C. L., Crous, P. W., Groenewald, J. Z., Boehm, E. W. A., Burgess, T. I., de Gruyter, J., de Hoog, G. S., Dixon, L. J., Grube, M., Gueidan, C., Harada, Y., Hatakeyama, S., Hirayama, K., Hosoya, T., Huhndorf, S. M., Hyde, K. D., Jones, E. B. G., Kohlmeyer, J., Kruys, A., Li, Y. M., Lucking, R., Lumbsch, H. T., Marvanova, L., Mbatchou, J. S., McVay, A. H., Miller, A. N., Mugambi, G. K., Muggia, L., Nelsen, M. P., Nelson, P., Owensby, C. A., Phillips, A. J. L., Phongpaichit, S., Pointing, S. B., Pujade-Renaud, V., Raja, H. A., Plata, E. R., Robbertse, B., Ruibal, C., Sakayaroj, J., Sano, T., Selbmann, L., Shearer, C. A., Shirouzu, T., Slippers, B., Suetrong, S., Tanaka, K., Volkmann-Kohlmeyer, B., Wingfield, M. J., Wood, A. R., Woudenberg, J. H. C., Yonezawa, H., Zhang, Y. and Spatafora, J. W. (2009). A class-wide phylogenetic assessment of Dothideomycetes. *Studies in Mycology* **64**: 1-15.
- Schurch, S., Linde, C. C., Knogge, W., Jackson, L. F. and McDonald, B. A. (2004). Molecular population genetic analysis differentiates two virulence mechanisms of the fungal avirulence gene NIP1. *Molecular Plant-Microbe Interactions* **17**: 1114-1125.
- Selker, E. U. (1990). Premeiotic instability of repeated sequences in *Neurospora crassa*. *Annual Review of Genetics* **24**: 579-613.
- Shabab, M., Shindo, T., Gu, C., Kaschani, F., Pansuriya, T., Chintha, R., Harzen, A., Colby, T., Kamoun, S. and van der Hoorn, R. A. L. (2008). Fungal effector protein

- AVR2 targets diversifying defense-related Cys proteases of tomato. *Plant Cell* **20**: 1169-1183.
- Shan, L. B., Thara, V. K., Martin, G. B., Zhou, J. M. and Tang, X. Y.** (2000). The *Pseudomonas* AvrPto protein is differentially recognized by tomato and tobacco and is localized to the plant plasma membrane. *Plant Cell* **12**: 2323-2337.
- Shan, W. X., Cao, M., Dan, L. U. and Tyler, B. M.** (2004). The Avr1b locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene Rps1b. *Molecular Plant-Microbe Interactions* **17**: 394-403.
- Shoemaker, R. A. and Brun, H.** (2001). The teleomorph of the weakly aggressive segregate of *Leptosphaeria maculans*. *Canadian Journal of Botany* **79**: 412-419.
- Sprague, S. J., Marcroft, S. J., Hayden, H. L. and Howlett, B. J.** (2006). Major gene resistance to blackleg in *Brassica napus* overcome within three years of commercial production in southeastern Australia. *Plant Disease* **90**: 190-198.
- Stachowiak, A., Olechnowicz, J., Jedryczka, M., Rouxel, T., Balesdent, M. H., Happstadius, I., Gladders, P., Latunde-Dada, A. and Evans, N.** (2006). Frequency of avirulence alleles in field populations of *Leptosphaeria maculans* in Europe. *European Journal of Plant Pathology* **114**: 67-75.
- Stahl, E. A. and Bishop, J. G.** (2000). Plant-pathogen arms races at the molecular level. *Current Opinion in Plant Biology* **3**: 299-304.
- Staskawicz, B. J., Dahlbeck, D. and Keen, N. T.** (1984). Cloned avirulence gene of *Pseudomonas syringae* pv *Glycinea* determines race-specific incompatibility on glycine-max (L) MERR. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **81**: 6024-6028.
- Thomma, B., Van Esse, H. P., Crous, P. W. and de Wit, P. J. G. M.** (2005). *Cladosporium fulvum* (syn. *Passalora fulva*), a highly specialized plant pathogen as a model for functional studies on plant pathogenic Mycosphaerellaceae. *Molecular Plant Pathology* **6**: 379-393.
- Thordal-Christensen, H.** (2003). Fresh insights into processes of nonhost resistance. *Current Opinion in Plant Biology* **6**: 351-357.
- Travadon, R., Bousset, L., Saint-Jean, S., Brun, H. and Sache, I.** (2007). Splash dispersal of *Leptosphaeria maculans* pycnidiospores and the spread of blackleg on oilseed rape. *Plant Pathology* **56**: 595-603.
- Tsuba, M., Katagiri, C., Takeuchi, Y., Takada, Y. and Yamaoka, N.** (2002). Chemical factors of the leaf surface involved in the morphogenesis of *Blumeria graminis*. *Physiological and Molecular Plant Pathology* **60**: 51-57.
- U, N.** (1935). Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Japanese Journal of Botany* **7**: 389-452.
- van den Burg, H. A., Westerink, N., Francoijs, K. J., Roth, R., Woestenenk, E., Boeren, S., de Wit, P. J. G. M., Joosten, M. H. A. J. and Vervoort, J.** (2003). Natural disulfide bond-disrupted mutants of AVR4 of the tomato pathogen *Cladosporium fulvum* are sensitive to proteolysis, circumvent Cf-4-mediated resistance, but retain their chitin binding ability. *Journal of Biological Chemistry* **278**: 27340-27346.
- van der Biezen, E. A. and Jones, J. D. G.** (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences* **23**: 454-456.
- van der Hoorn, R. A. L. and Kamoun, S.** (2008). From Guard to Decoy: A new model for perception of plant pathogen effectors. *Plant Cell* **20**: 2009-2017.
- van Kan, J. A. L., van den Ackerveken, G. and de Wit, P. J. G. M.** (1991). Cloning and characterization of cDNA of avirulence gene AVR9 of the fungal pathogen

- Cladosporium fulvum*, causal agent of tomato leaf mold. *Molecular Plant-Microbe Interactions* **4**: 52-59.
- Vincenot, L., Balesdent, M. H., Li, H., Barbetti, M. J., Sivasithamparam, K., Gout, L. and Rouxel, T.** (2008). Occurrence of a new subclade of *Leptosphaeria biglobosa* in Western Australia. *Phytopathology* **98**: 321-329.
- West, J. S., Kharbanda, P. D., Barbetti, M. J. and Fitt, B. D. L.** (2001). Epidemiology and management of *Leptosphaeria maculans* (phoma stem canker) on oilseed rape in Australia, Canada and Europe. *Plant Pathology* **50**: 10-27.
- West, J. S., Balesdent, M. H., Rouxel, T., Narcy, J. P., Huang, Y. J., Roux, J., Steed, J. M., Fitt, B. D. L. and Schmit, J.** (2002). Colonization of winter oilseed rape tissues by A/Tox(+) and B/Tox(0) *Leptosphaeria maculans* (phoma stem canker) in France and England. *Plant Pathology* **51**: 311-321.
- Whalen, M. C., Wang, J. F., Carland, F. M., Heiskell, M. E., Dahlbeck, D., Minsavage, G. V., Jones, J. B., Scott, J. W., Stall, R. E. and Staskawicz, B. J.** (1993). Avirulence gene AVRXXV from *Xanthomonas campestris* pv *vesicatoria* specifies resistance on tomato line HAWAII-7998. *Molecular Plant-Microbe Interactions* **6**: 616-627.
- Williams, R. H. and Fitt, B. D. L.** (1999). Differentiating A and B groups of *Leptosphaeria maculans*, causal agent of stem canker (blackleg) of oilseed rape. *Plant Pathology* **48**: 161-175.
- Young, J. and Holland, I. B.** (1999). ABC transporters: bacterial exporters-revisited five years on. *Biochimica Et Biophysica Acta-Biomembranes* **1461**: 177-200.
- Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E. J., Jones, J. D. G., Felix, G. and Boller, T.** (2004). Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* **428**: 764-767.
- Zhou, E. X., Jia, Y. L., Singh, P., Correll, J. C. and Lee, F. N.** (2007). Instability of the *Magnaporthe oryzae* avirulence gene *AVR-Pita* alters virulence. *Fungal Genetics and Biology* **44**: 1024-1034.





## **Abstract :**

*Leptosphaeria maculans* is a filamentous ascomycete causing stem canker of oilseed rape (*Brassica napus*). This disease is often controlled by the use of *B. napus* cultivars harbouring major resistance genes (*Rlm*). Direct or indirect recognition of the corresponding avirulence protein (*AvrLm*) in the pathogen triggers plant defence reactions. Several resistances have been massively deployed in France and worldwide, they initially showed commercial success due to the protection provided and in a second time, a very fast decrease of efficiency (resistance breakdown).

Prior to this thesis, field studies of resistance gene breakdown mechanisms were rare, especially for fungi. The purpose of this PhD thesis was to study the molecular evolution of the avirulence gene *AvrLm4-7* under selection pressure, by exploiting our knowledge of the gene, and the recent release of *Rlm7* cultivars, to obtain an early and detailed study of the molecular mechanisms involved in a resistance gene breakdown.

*AvrLm4-7* induces resistance responses in plant harbouring either *Rlm4* or *Rlm7* and I validated by targeted point mutagenesis the central role of the amino acid 120 in the *avrLm4-7-Rlm4* interaction. Its mutation prevents *AvrLm4-7* recognition by *Rlm4* without affecting *avrLm4-7-Rlm7* recognition.

Loss of avirulence towards *Rlm7* was then studied by the analysis of an important isolate collection originating from two independent French experimental fields (Grignon; Versailles) over three years. In the first field was cropped *Rlm7* cultivars in monoculture with low tillage agronomical practices whereas crop rotation and ploughing were done in the second field.

In contrast to *AvrLm4-Rlm4* evolution, a great number of mutations were found to explain the “gain” of virulence towards *Rlm7*. Seven mutational event categories were found. The great majority of these categories involve *AvrLm4-7* deletion but mutation due to RIP and several other mutational events causing premature apparition of stop codons in the coding sequence of the gene were observed too. The majority of these events are linked to the sexual reproduction of the fungus and occurs in the experimental field.

In addition, our work showed the importance of the cultural practices in preserving *Rlm7* efficacy. Indeed, after three years using *Rlm7* cultivars, a7 frequency was below 1% whereas representing around 30% of the isolates observed in Grignon.

Finally, phenotyping of the isolate collection also showed the resurgence of an A3 phenotype linked with the loss of *AvrLm7* avirulence in more than 98% of the isolates. Genetic analysis and collection phenotyping showed that *AvrLm3* is not a new *AvrLm4-7* allele but a distinct gene located in a telomeric region at 19.3 cM of *AvrLm4-7*. I also demonstrated that an antagonistic interaction between *AvrLm4-7* and *AvrLm3* exists: the presence of *AvrLm4-7* prevents *Rlm3* to detect *AvrLm3* and explains the surge of the *AvrLm3* avirulence along with the loss of the *AvrLm7* avirulence.

By an original association of molecular biology, population genetic and agronomy, this work provided a new illustration of the plant-pathogen arms race, *AvrLm3* and *AvrLm4-7* using two different strategies to escape their respective resistance genes.

## Résumé :

*Leptosphaeria maculans*, agent de la nécrose du collet des crucifères, est un agent pathogène majeur du colza (*Brassica napus*). La lutte génétique est aujourd'hui le procédé le plus utilisé afin de protéger les cultures des attaques de ce champignon. Cette méthode se base principalement sur l'utilisation de cultivars possédant des gènes de résistance spécifique (*Rlm*) qui permettent le déclenchement des réactions de défense de la plante par la reconnaissance directe ou indirecte des produits des gènes d'avirulence correspondants (*AvrLm*) présents dans la population pathogène. Plusieurs de ces résistances ont déjà été massivement déployées en France et dans le monde, connaissant dans un premier temps un fort succès commercial grâce à la protection fournie, suivie d'une perte d'efficacité très rapide.

Avant cette thèse, le nombre d'études au champ des processus impliqués dans le contournement d'un gène de résistance était très limité, en particulier chez les champignons. L'objectif de cette thèse était d'étudier l'évolution moléculaire du gène d'avirulence *AvrLm4-7* sous pression de sélection, en profitant de son clonage et de la commercialisation récente de cultivars *Rlm7*, afin d'obtenir une étude précoce et détaillée des mécanismes moléculaires à l'origine du contournement d'une résistance spécifique.

Le gène *AvrLm4-7* présente l'originalité de coder pour une protéine responsable d'une double spécificité d'interaction vis-à-vis des gènes *Rlm4* et *Rlm7*. Dans un premier temps, j'ai pu valider par mutagenèse dirigée le rôle primordial de l'acide aminé 120 dont la mutation affecte la reconnaissance d'*AvrLm4* par *Rlm4* sans toutefois altérer la reconnaissance d'*AvrLm7* par *Rlm7*.

Le contournement de la résistance *Rlm7* a été ensuite analysé à l'aide d'une importante collection de souches prélevée sur deux sites expérimentaux indépendants (Grignon ; Versailles) sur une période de trois ans. Sur le premier site était cultivé une variété *Rlm7* en monoculture avec un travail du sol simplifié tandis que sur le second site, le mode de culture incluait rotation culturale et enfouissement par labour des résidus de cultures.

Il a ensuite été montré que, au contraire de la reconnaissance *AvrLm4/Rlm4*, un grand nombre d'évènements de mutation peuvent être à l'origine de la virulence d'une souche vis-à-vis de *Rlm7*. L'analyse moléculaire des souches virulentes et avirulentes de cette collection a ainsi permis de répertorier sept catégories d'évènements de mutation. La grande majorité des cas concerne la délétion d'*AvrLm4-7* mais des mutations dues au RIP et plusieurs autres évènements de mutation provoquant l'introduction prématurée de codons stop dans la séquence codante du gène sont aussi observés. La majorité de ces évènements de mutation sont liés à la reproduction sexuée du champignon et ont lieu au sein même de la parcelle d'étude.

Le phénotypage de cette collection a par ailleurs révélé un fort contraste entre les deux sites expérimentaux, démontrant ainsi l'importance des pratiques culturales dans le maintien de l'efficacité de la résistance *Rlm7* dans le temps. En effet, après trois années de culture de cultivars *Rlm7*, la fréquence des souches virulentes a7 dans les populations du site de Versailles reste inférieure à 1 % contre environ 30 % sur le site de Grignon.

Finalement, le phénotypage de la collection de souches a également montré que le contournement de *Rlm7* s'accompagnait dans plus de 98% des souches de la résurgence de l'avirulence *AvrLm3*. Par l'étude de cette collection et par croisements génétiques, j'ai pu montrer que *AvrLm3* n'était pas un nouvel allèle d'*avrLm4-7* mais un second gène situé en région télomérique à 19.3 cM d'*AvrLm4-7*. J'ai également démontré une interaction fonctionnelle antagoniste entre *AvrLm4-7* et *AvrLm3* qui empêche la reconnaissance *Rlm3* / *AvrLm3* en présence d'*AvrLm4-7* et explique la restauration de l'avirulence *AvrLm3* lors de la perte de l'avirulence *AvrLm7*.

Par une association originale de biologie moléculaire, de génétique des populations et d'agronomie, j'ai ainsi pu apporter une nouvelle illustration à la course aux armements entre un agent pathogène et sa plante hôte, les gènes *AvrLm3* et *AvrLm4-7* utilisant deux stratégies distinctes afin d'échapper à la reconnaissance de leurs gènes de résistance spécifiques.