Study of the molecular basis of recognition of the fungal effector AVR-Pia by the rice immune receptor RGA5

Diana Vallejo Ortiz

Study of the molecular basis of recognition of the fungal effector AVR-Pia by the rice immune receptor RGA5

Diana Ortiz

THÈSE

Pour obtenir le grade de

Docteur

Délivré par

Montpellier SupAgro

Préparée au sein de l'école doctorale GAIA

Et de l'unité de recherche BGPI

Spécialité : **Biologie Des Interactions**

Présentée par **Diana ORTIZ-VALLEJO**

Étude des bases moléculaires de la reconnaissance de l'effecteur fongique AVR-Pia par le récepteur immunitaire du riz RGA5

Soutenue le 07 novembre 2016 devant le jury composé de

M. Thomas KROJ, DR, INRA	Directeur
M. Laurent NOËL, DR, CNRS	Rapporteur
M. Marc-Henri LEBRUN, DR, CNRS	Rapporteur
Mme Claire NEEMA, Professeur, SupAgro	Examinatrice
M. Harald KELLER, DR, INRA	Examinateur

Dedicada a

Baltazar, por tu magia en mi vida

REMERCIEMENTS

« Un ciel, la nuit. La voie lactée scintille au loin. La Terre tourne sur elle-même, poursuivant sa course à travers l'espace, parcourant une parcelle infime d'un univers sans fin. Un cosmos décoré de planètes, d'étoiles, de soleils et de galaxies, qui se déploie, jour après jour. De ses origines nous ne savons rien, ni de son devenir, pourtant nous savons qu'il est habité. La Terre a emporté à travers l'espace et le temps le foisonnement changeant des innombrables engendrements du vivant avec les formes les plus belles et les plus merveilleuses qui ont été composée et recomposée depuis l'origine. Chaque créature vivante peut être considérée comme un microcosme, un infime univers, constitué d'une inimaginable multitude de petits organismes qui se reproduisent sans fin et que sont aussi nombreux que les étoiles dans le ciel ». *La sculpture du vivant*

Au mystère de l'origine du cosmos qui nous entoure répond le mystère de l'origine du vivant. Tous les deux enlacent pour un fil invisible de molécules qui sont le cœur de ces univers parallèles. Eh bien… c'était l'irrémédiable pulsion de plonger dans l'univers moléculaire qui m'a propulsé à faire une thèse. C'est pour cela qu'avant tout je remercie la vie. C'est grâce à elle que, pour une succession d'imprévisibles accidents terribles et merveilleux je suis arrivée ici, un coin charmant de la France auprès de gens et paysages en qui j'ai énormément appris. Je les garderai pour toujours dans ma mémoire et mon cœur.

A quelques kilomètres d'ici se trouvent les gens à qui je dois la force que j'ai eu quand l'aventure de la thèse a été difficile à traverser. Padresito, alguna vez escuche que los hijos son la prolongación de la vida y yo lo siento así, es por eso que en cada uno de los pasos que he dado lejos de ti he intentado ser feliz, esta es mi única manera de agradecerte y hacerte parte de un sueño que considero mutuo. A mi madre le agradezco por todos los ángeles que pone cada día en mi destino, son muchos!!! Y sé que todos vienen enviados por ti. Mis hermanos, sobrinos, amigos y demás miembros anexos a la familia, quisiera nombrarlos uno a uno pero no con todo lo que podría decirles no acabaría nunca… en estas cortas líneas solo me queda mencionar que ustedes representan la alegría y la luz de esta loca travesía, los amo profundamente. Camilito, gracias por todo el amor compartido, tu de alguna forma fugaz paso por mi vida se ha inscrito en mi como el paso de una ráfaga de dulzura y ternura inolvidable, tu sabes que nos llevaremos por siempre en la eternidad sin tiempo ni formas.

Thomas je te remercie énormément pour m'avoir donné l'opportunité de faire une thèse sous ton encadrement. Ça fait déjà plus de cinq ans qu'on a commencé à travailler ensemble pour faire d'un rêve une réalité. Tu m'as soutenue toujours, même quand j'ai fait des erreurs avec la candidature à Colciencias! Je ne vais jamais oublier que même à ce moment-là tu m'as fait confiance. Ce sont ces beaux actes qui transforment les gens, alors tout comme Jean Valjean à qui la confiance de quelqu'un a transformé, moi je peux te dire que l'encouragement de ta part

a été vital tout au long de cette thèse et que sans cela ce rêve serai une réalité sans saveur. Merci aussi pour toute ta patience et pour tout ce que tu m'as appris et tu m'a poussé à faire. Trois congrès internationaux, un super stage en Angleterre et deux beaux papiers, sont un petit résumé de ton encadrement! En plus je suis très contente d'être ta première thésarde officielle. A Stella d'avoir fait un super travail et faire mon chemin plus facile. Aux membres de mon comité de thèse pour les conseils toujours très constructifs.

Jean Benoit merci beaucoup de m'avoir accueillie dans ton équipe et pour tous tes conseils scientifiques et personnels. Saches que sans le savoir, tes mots d'encouragement sont toujours arrivés dans les moments où j'avais le plus besoin. Merci aussi pour tes gestes qui m'ont fait beaucoup de bien comme par exemple le manioc ou les lumières pour le vélo dans un obscur hiver.

A tout l'ensemble de l'équipe 4 mille et mille mercis pour toute l'aide lors de manips et aussi pour les moments partagés. Véro merci pour toute ton aide avec mes expériences, j'ai trop adoré travailler avec toi. Merci aussi pour les livres et toutes nos conversations « philosophiques » tu vas me manquer. Coco, c'était très agréable recevoir tes câlins et partager notre passion pour le tango, j'espère que tu ne t'arrêteras jamais de danser. Aurélie j'espère que tous les souhaites qu'on a envoyé aux étoiles filantes deviendront réalité. Przemyslaw merci pour ton hospitalité en Pologne et pour les vins partagés. Emilie, quelles vagues ont a surmonté ensemble!!!!!! On est souvent passé de l'euphorie aux larmes… mais bon on est toujours là et tu vois tout va bien… particulière époque de nos vies hein ? Je te remercie pour m'avoir fait découvrir de très belles musiques, je suis contente de te voir rapprocher de plus en plus de ton coté vivant et paisible. Celui que j'ai eu l'opportunité de découvrir et apprécier dans les montagnes Polonaises lorsque tu chantes et joues de la guitare. J'espère que la musique et la science continueront à te procurer autant de bonheur. Elsa, Isabelle, Audrey, Loïc, Huichuan, Bastien, Jing-Jing merci beaucoup pour votre courtoisie, votre aide et vos sourires. Paola, c'était cool d'avoir una compatriota au labo! A ton tour de continuer à faire vibrer la Colombie à BGPI.

Romain Ferdinand qu'est-ce que j'aurais fait à BGPI sans toi !!! Je n'ai pas des mots pour te dire combien je suis contente de t'avoir croisé dans ce coin de la terre. Merci énormément pour tout ton soutien. Je sais que je t'ai embêté un peu trop mais tu sais, j'en avais besoin. Merci aussi pour m'avoir appris le français, m'amener quand il « pleuvait » et enfin d'être toujours là. La facture je te la payerai quand je serai une chercheuse milliardaire. Que la lumière de chaque couché de soleil continue à t'émerveiller. Charlotte, à toi aussi je te remercie profondément pour tes justes mots remplis de beauté et de sincérité. Avoir ton amitié c'est un grand cadeau venu du ciel. En fait ce tour en avion c'était une expérience qui représente très bien tout ce que tu m'as apporté. J'espère qu'on fera d'autres voyages ensemble.

Katia, merci pour être à l'écoute et t'investir pour rendre les choses plus simples pour moi et tous les thésards. Merci aussi pour ta joie de vivre, elle est contagieuse!!! Henri, grâce à toi j'ai eu le courage de venir en vélo pendant l'hiver, dans une route sans trop de lumière et avec la menace de sangliers hehe.. quelle aventure !!! mais tu sais ?, pendant que je roulais, je jetais un œil à la lune et les étoiles et ça devenait un expérience agréable. Bon, quand-même je dois avouer que je suis rassurée de pas devoir le faire à nouveau mais mille merci pour me sauver des heures passées dans les transportes publiques. Merci aussi de nous avoir fait découvrir fillols et le Pyrénées-Orientales c'était top. Sonia, Marlène, Romain Gallet, Agnés, Pierre, Didier, merci beaucoup pour être là dans cette dernière ligne droite. De manière plus générale je remercie vraiment toutes les personnes que j'ai croisées à BGPI et qui ont fait partie de cette belle aventure. A Marie Noel pour m'avoir amenée de retour à ma vraie maison et me rendre mes ailes.

A tous mes amis de Montpellier, Boris mon lien Colombie-France, merci pour m'avoir fait rêver d'étudier dans ton pays. Cela a été sans aucun doute le meilleur choix. On pourrait même dire que t'avoir croisé a changé ma vie! Je t'apprécie très fortement. A la bande de intensos Jesus, Nico, Ponchis, Reini, Raul, Nico, Sonia, Thierry, Miguel, Luis, Sheena, Jean-Philippe, Lili, Ryan et Alvaro, j'ai vécu des très bons moments avec vous et je suis sûre qu'on se retrouvera à nouveau parce que les bons amis restent tout au long d'une vie , vous êtes dans mon cœur. A Elisabeth pour toute son hospitalité. A Dani et à Michel mes « parents » en France, quelle chance j'ai de vous avoir rencontré.

A la nouvelle bande de thésardes avec qui j'ai eu la chance de partager ma dernière année de thèse Coralie, Edwige et Maelle trois filles magnifiques et courageuses, toute mon admiration pour vous! Vous m'avez énormément appris et votre amitié c'est l'une des choses les plus jolies que j'ai eu avant partir de ces murs. Je suis sûre que toutes les trois feront de super thèses ! A mes anciens collègues thésardes Pauline, Deborah, Anne, Souhir merci aussi pour les bons moments et les pauses du vendredi. Votre exemple c'était un moteur de motivation. A Claire Neema pour son inconditionnel soutien aux thésards. Merci à Marie-Agnès, Pierre, Léa et Vincent de m'avoir accueillie avec autant de tendresse dans votre famille.

Et bon… moi qui attendais patiemment la voie de retour à mon pays, comme dans les contes de fée, c'est à BGPI que je suis tombée avec un lobo rêveur qui m'a persuadée d'allonger mon chemin pour découvrir d'autres univers. Je te remercie de m'avoir regardée et de m'inviter à faire partie d'une aventure commune, je suis sûre qu'elle sera magnifique.

LIST OF ABBREVIATIONS

ABA: Abscisic acid

BAS: Biotrophy-Associated Secreted

BIC: Biotrophic Interfacial Complex

CC: Coiled-Coil

CEBiP: Chitin Elicitor-Binding Protein

CERK1: Chitin Elicitor Receptor Kinase-1

CK: Cytokinin

DAMP: Damage-Associated Molecular Patters

EGF: Epidermal Growth Factor

ER: Endoplasmic Reticulum

ETI: Effector-Triggered Immunity

ETS: Effector-Triggered Susceptibility

HMA: Heavy Metal-Associated

HR: Hypersensitive Response

IP: Invasion Pattern

IPR: Invasion Pattern Receptor

IPTR: Invasion Pattern Triggered Response

JA: Jasmonic Acid

LRR: Leucine Rich-Repeats

LYK: Lysin motif receptor kinase

LysM: Lysin Motif

MAMP: Microbial Associated Molecular Patterns

MHD: Methionine-Histidine-Aspartic acid motif

MPSS: Massively Parallel Signature Sequencing

NBS: Nucleotide-Binding Site

NLP: Necrosis- and Ethylene-inducing Peptide1-Like Protein

NLR: Nucleotide-binding and Leucine-rich repeat protein

NO: Nitric Oxide

NPR1: Non-expresser of PR proteins 1

PAMP: Pathogen Associated Molecular Patterns

PI3P: Phosphatidylinositol-3-Phosphate

P-loop: Phosphate-binding loop

PR: Pathogenesis-related proteins

PRR: Pattern Recognition Receptor

PTI: Pattern-Triggered Immunity

RATX1: Related to ATX1

RLK: Receptor-Like Kinases

RLP: Receptor-Like Kinases and Receptor-Like Proteins

ROS: Reactive Oxygen Species

SA: Salicylic Acid

SBS: Sequencing By Synthesis

SNP: Single-Nucleotide Polymorphisms

SPD: Suppresors of Plant Cell Death

STAND: Signal Transduction ATPase with Numerous Domains

T3SS : Type 3 Secretion System

TIR: Toll Interleukin-1-like Receptor

GLOSSARY

NLRs:

Intracellular receptor proteins that monitories pathogen attack in the cytosol and play central roles in the innate immune systems of plants

PRRs:

Membrane receptor proteins that perceive pathogens in the extracellular space and have important roles in plant immunity

Effector :

Pathogen secreted molecules that modulates plant immunity to facilitate infection in compatible interactions

AVR-Effector :

Secreted proteins recognized by plant immune receptors in an incompatible interaction

Compatible interaction:

Pathogens can develop and reproduce due to an absence of an effective immune receptor in the plant host

Incompatible interaction:

Pathogens are not able to develop because they carry AVR-effector proteins that are recognized by plant immune receptors

Direct recognition:

Plant immune receptors recognize pathogens by the interaction with the effector proteins

Indirect recognition:

Plant immune receptors recognize pathogens by monitoring the integrity of host proteins target by effectors

Integrated domain:

Unconventional domains integrated into the classical NLR protein structure

Integrated decoy:

NLR-integrated domain involved in effector perception

***NLR* pairs:**

Two *NLR*s genes genetically linked that work together in the recognition of effector proteins. One function as sensor and the other as executor

Helper *NLRs*:

NLRs acting as signalling components of one or more NLR sensors. Helpers working in pairs are less specific compare to *NLR* pairs

TABLE OF CONTENTS

INTRODUCTION

1. Plant immunity

1.1 Plant defenses and pathogen infection

Plants are essential organisms of any live food chain and are constantly exposed to abiotic and biotic factors that can positively or negatively affect their development. For instance, during their life cycle plants are in intimate association with microorganisms as diverse as bacteria, fungi, nematodes, oomycetes and viruses and interact with them in different ways. In some cases, this association can lead to beneficial interactions in which both organisms benefit to each other, e.g. the interaction between legumes and rhizobium, a bacteria that, after becoming established inside rood nodules, helps the plant to fix nitrogen (Hirsch et al., 2001). However, in other cases microorganism act as pathogens and represent a danger for the plant since their interaction lead to disease. To face this challenge, plants possess **preformed defenses** to arrest pathogen infection, such as physical barriers and constitutive production of chemical compounds with antimicrobial properties (Göhre and Robatzek, 2008). In addition, plants have evolved a sophisticated **innate immune system** that confers to individual plant cells the capacity to sense and respond to pathogen attack. Indeed, after pathogen recognition, plants induce complex and multilayered defense responses (Figure 1). At sites with contact to the pathogen, the cell wall is reinforced by the production of various cell wall components such as glycoproteins, lignin, callose or suberin. ROS (Reactive Oxygen Species) are produced and act directly by their antimicrobial activities and the crosslinking of cell wall components as well as indirectly by their role as signaling molecules. The production of pathogenesis-related proteins (PR) and secondary metabolites are induced by complex gene regulatory networks and limit pathogen development by their antimicrobial activities (Hammond-kosack & Jones 1996; VanLoon 1997; Torres et al. 2006; Dong & Kahmann 2009; Göhre & Robatzek 2008; Lehmann et al. 2015)**.**

These plant immune responses lead to the establishment of either **nonhost** or **host** resistance. **Nonhost resistance** refers to the broad-spectrum plant defense that provides immunity to all

members of a plant species against all isolates of a microorganism that is pathogenic to other plant species whereas host resistance is cultivar or accession specific and is activated only in response to some isolates or races of adapted pathogen species.

Pathogens use specialized structures and molecular arms to overcome plant defenses, colonize plant cells and acquire nutrients. For instance, fungi and oomycetes develop specialized cells named appressoria to break the leaf surface and invasive hyphae or haustoria to feed on living plant cells. Nematodes and insects use styletes to penetrate and feed on plant cells. Bacteria possess macromolecular structures like the type III or IV secretion systems to deliver virulence factors inside host cells. Regardless of life style, infection structures and reproduction mode, all plant pathogens use molecular weapons during plant infection corresponding in particular to secondary metabolites or **secreted proteins**.

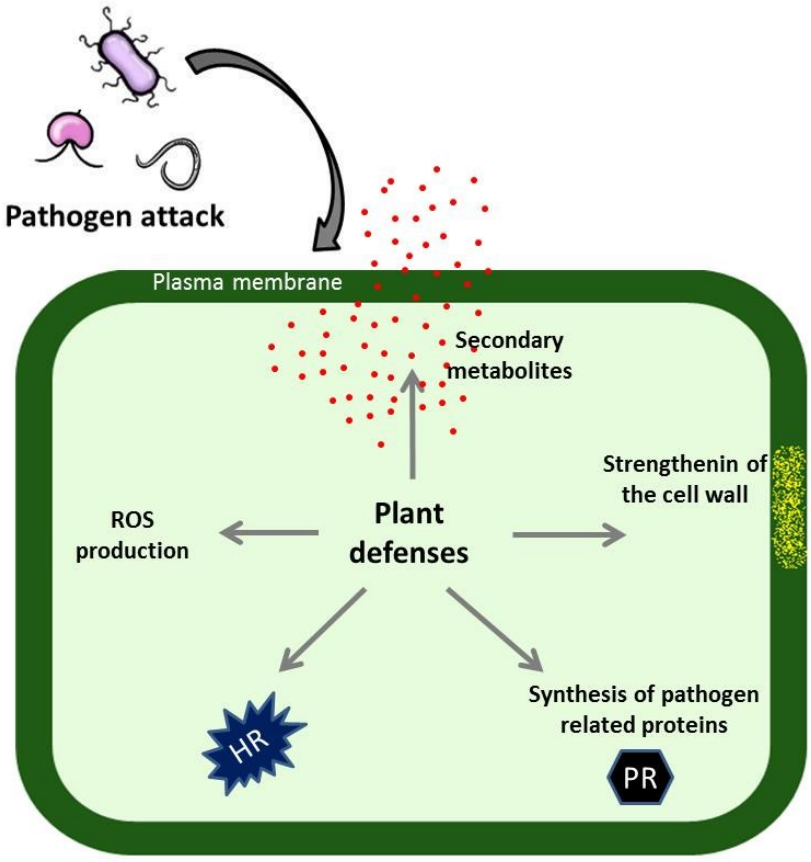


Figure 1. Induction of plant defense responses upon pathogen attack.

Adapted from (Hammond-kosack & Jones 1996)

1.2 Models to explain plant immunity

In all biological systems, immunity relies on the induction of defense responses upon pathogen attack and this is mediated by the recognition of two types of signals, non-self and modified self either inside cells or on the cell surface. In plants, pathogen perception inside cells is exclusively mediated by NLRs (Nucleotide-binding and Leucine-rich repeat proteins) while two different types of immune receptors perceive signals on the cell surface, RLKs (Receptor-Like Kinases) and RLPs (Receptor-Like Proteins) (Figure 2). A detailed description of NLRs is provided in section 3, while basic knowledge on RLKs and RLPs is summarized in Figure 8.

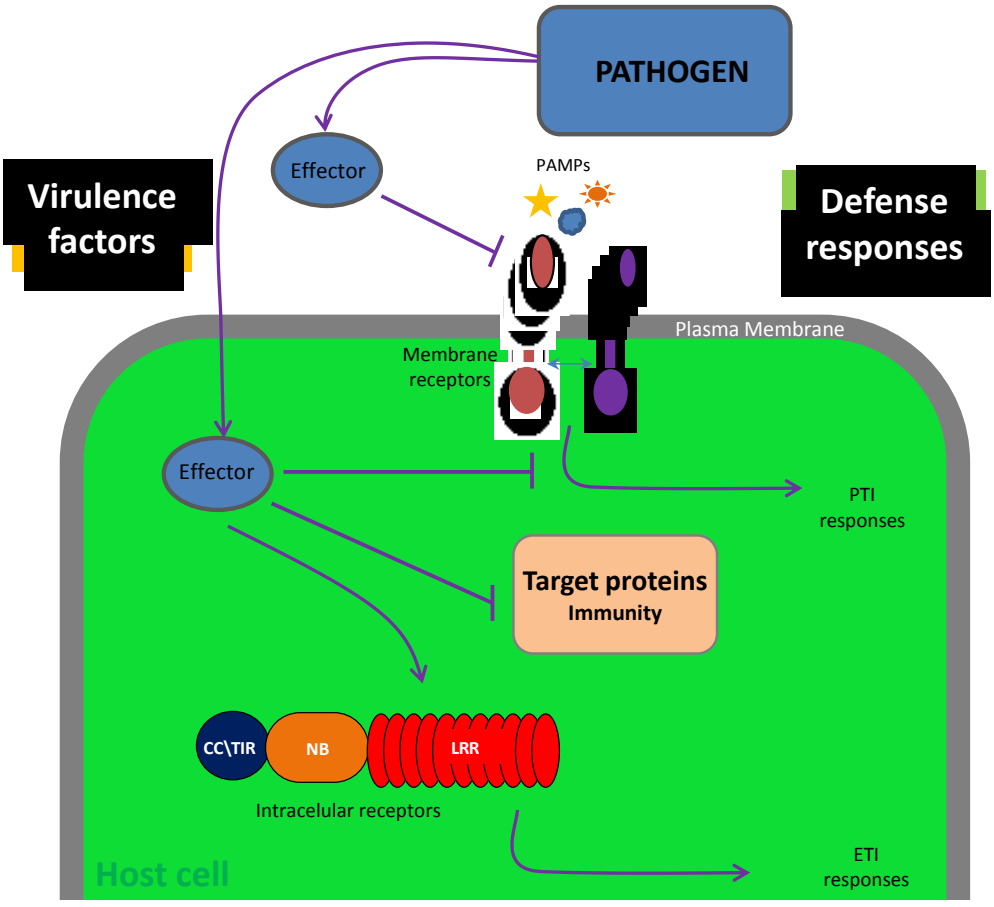


Figure 2. Schematic view of plant immunity. Plant immunity relays in two different classes of immune receptors located in the plasma membrane and cytoplasm where they detect virulence factors and modified-self. Adapted from Dodds & Rathjen 2010.

The Zig-zag model

Over the last 10 years, the understanding and interpretation of plant immunity and the action and evolution of the plant immune receptors and their corresponding ligands has been very much influenced by the zig-zag model (Jones & Dangl 2006; Dodds & Rathjen 2010). This model (Figure 3) divides the plant immune systems in two layers.

A first layer relies on the recognition of **microbial or pathogen associated molecular patterns (MAMPs or PAMPs)** such as bacterial flagellin or chitin from fungi. MAMPs and PAMPs are defined as conserved microbial molecules with low variability and wide distribution in a phylogenetically large range of organisms. They allow thus the recognition of wide ranges of potentially deleterious microorganisms by activating **pattern-triggered immunity (PTI)**. This provides basal protection against non-adapted pathogens and potentially damaging microorganisms and attenuates the virulence of adapted pathogens. The recognition of MAMPs or PAMPs in plants is mediated by **pattern recognition receptors (PRRs)** that are transmembrane receptor proteins of the RLK and RLP type. However, adapted pathogens are able to overcome PTI by the deployment of effectors that suppress or circumvent this first layer of immunity. They are therefore able to infect their host plant and to establish **effector-triggered susceptibility (ETS)**.

A second layer of immunity that is based on the recognition of effectors by NLR-type immune receptors and therefore named **effector-triggered immunity (ETI)** allows the control of these adapted pathogens. The ETI response has been described as an amplification of PTI responses and leads to complete disease resistance. It is frequently associated with a localized programmed cell death called the hypersensitive response (HR). Sometimes, pathogens can adapt to ETI by evolving novel effectors that suppress ETI and establish again ETS.

The merit of the Zig-zag model was to provide a first unified view of plant immunity that integrates the phenomena of basal resistance and specific (gene-for-gene type) plant resistance and links it to pathogen effector action. In addition, it provided an evolutionary

explanation for the complex interplay between host resistance and pathogen virulence by nicely illustrating the co-evolution between pathogens and plants that both continually modify their weapons to either avoid or induce resistance. In addition, the zig-zag model also provides a good framework for the understanding of the infection process by phytopathogenic bacteria, in particular *Pseudomonas sryingae* and many *Xanthomonas* bacteria under artificial conditions. In infection experiments, these bacteria are generally infiltrated into plant leafs where they instantaneously liberate huge quantities of PAMPs and MAMPs and induce rapid PTI responses (Katagiri et al., 2002; Buell, 2002). Since the type 3 secretion system (T3SS) and T3SS effectors are under transcriptional control and only produced in planta, effector delivery occurs only several hours after infiltration. PTI suppression and ETS establishment is therefore only visible in a second step, typically 3-4 hours after infiltration. Therefore, recognition of effectors and ETI induction also only occurs in a second step (Katagiri et al., 2002; Buell, 2002).

Limitations of zigzag model

Despite its wide popularity and the historical importance of the zig-zag model, research over the last decade showed that major assumptions and conclusions of the zig-zag model miss precision or are erroneous and that there is a need for a novel more unifying model (Thomma et al., 2011). The fundamental conceptual error of the zig-zag model is the separation of plant immunity in two separated layers. This fits badly with experimental data and has led to many misconceptions and wrong assumptions about the evolution of immune receptors and their ligands. In addition, it generates numerous problems in the classification of virulence factors PAMPs vs effector as well as the immune response either PTI or ETI triggered by immune receptors.

Problems become e.g. evident when effectors that are defined in the Zig-zag model as highly variable, rapidly evolving and species or lineage specific proteins behave more like, PAMPs meaning they are broadly distributed and present in large ranges of organisms. It is also problematic, when PAMPs that are defined as pathogen molecules involved in general cellular functions behave like effectors since they are highly polymorphic and trigger differential

resistance responses (Cook et al., 2014). Indeed, research in multiple pathosystems has shown that not only effectors but also MAMPs/PAMPs evolve to escape recognition and are under selection pressure (Michelmore et al., 2013; Vinatzer et al., 2014). The rate of evolution of both types of ligands can vary considerably and their phylogenetic distribution can be more or less wide. For example, there is strong variation in the amino acid sequences of the most-studied PAMPs, the flagellin protein, in the bacterial pathogen *Xanthomonas campestris* pv. *campestris* resulting in variation of defense responses on *Arabidopsis thaliana* possessing the flagellin receptor FLS2 (Sun et al., 2006). Similarly, the significant variation in the residues of the flg22 epitope of *Ralstonia solanacearum* K60 and *Pseudomonas cannabina* pv. *alisalensis* ES4326 resulted in a variation in the resistance response triggered by their corresponding receptor proteins in the host (Pfund et al. 2004; Clarke et al. 2013).

Additionally, the discovery of a second epitope, the flagellin flgII-28, that is sufficient to trigger immune responses in tomato, confirmed the dynamic nature of flagellin-host perception and highlight the importance of PAMP diversification as a virulence strategy of the pathogen (Clarke et al., 2013). In addition, PAMPs can have a patchy distribution. An example for this is the case of the sulfated RaxX peptide recognized by the Xa21 RLK in rice. RaxX is present in a limited number of *Xanthomonas* species and shows presence absence polymorphism in *Xoo* isolates (Pruitt et al., 2015).

On the contrary, NLPs (Necrosis- and Ethylene-inducing Peptide1-Like Proteins) are an example of effectors with an extremely wide distribution in prokaryotic and eukaryotic phytopathogens that are recognized in *Arabidopsis thaliana* by a conserved protein pattern (Bohm et al. 2014; Oome & Van den Ackerveken 2014; Oome et al. 2014). Another example of widely distributed effectors, are the LysM domain effector proteins that occur in a broad range of fungal pathogens and that were frequently miss-qualified as PAMPs in the frame of the zig-zag model (Thomma et al., 2011).

Another problem of the zig-zag model is that RLPs and RLKS are categorized as PRRs dedicated to MAMP/PAMP detection. They are predicted to be conserved throughout and between plant species and to evolve at slow rates. NLRs, on the contrary, are viewed as receptors of isolate or pathogen-specific effectors and to be present in only limited numbers of plant accessions. This view is again very biased since RLPs may also recognize isolate-specific effectors and act as full, gene-for-gene type R proteins. For example the race-specific Cf-2 receptor from tomato monitors the presence of the effector protein Avr2 from *Cladosporim fulvum* strains by guarding the target Rcr3 protease (Rooney et al., 2005). RLKs, show less diversity but possess marked presence-absence polymorphism within and between species. Indeed most MAMP receptors are redistricted to certain species or phylogenetic classes such as EFR that is limited to Brassicaceae (Kunze, 2004; Boller and Felix, 2009). As a consequence, contrary to the view given by the zig-zag model that describes PTI as a stable character, there is clear variability in the detection of MAMPs within and between species (Robatzek et al., 2007; Gómez-Gómez and Boller, 2000). Finally, the view that ETI responses are stronger and more intense than PTI responses is not always verified since various PAMPs, including flagellin induce an HR.

The invasion model

A recent attempt to describe plant immunity in a more unbiased way, to incorporate a large range of mechanistically diverse interactions and to avoid strict divisions is the invasion model (Figure 3). This model states that plant immunity relies on the recognition of **invasion patterns (IP)** that can be two different types of ligands, microbe-derived molecules or modified-self ligands by **invasion pattern receptors (IPRs**) that may be either PRRs or NLRs. Any molecule that can be recognized by an **IPR** can act as an **IP** and induce **IP-triggered responses (IPTRs).** An important distinction from the zig-zag model is that IPs are only defined with respect to their function in host perception without any restriction or prediction of their physiological or biochemical function. Thus, in the invasion model the probability to activate **IPTR** depends on several factors such as the ligand's molecular constrains to retain function and the variability of IPs and IPRs across organisms. Thereby, the model incorporates on one hand the diversity of ligands that are important in plant immunity and that englobe MAMPs,

damage-associated molecular patters (DAMPs) and effectors and on the other hand since all types of immune receptors fall in the same category, the specificities of recognition are explained in terms of the nature of the response e.g., weak or strong, narrow or broad, specific or common etc., (Figure 3).

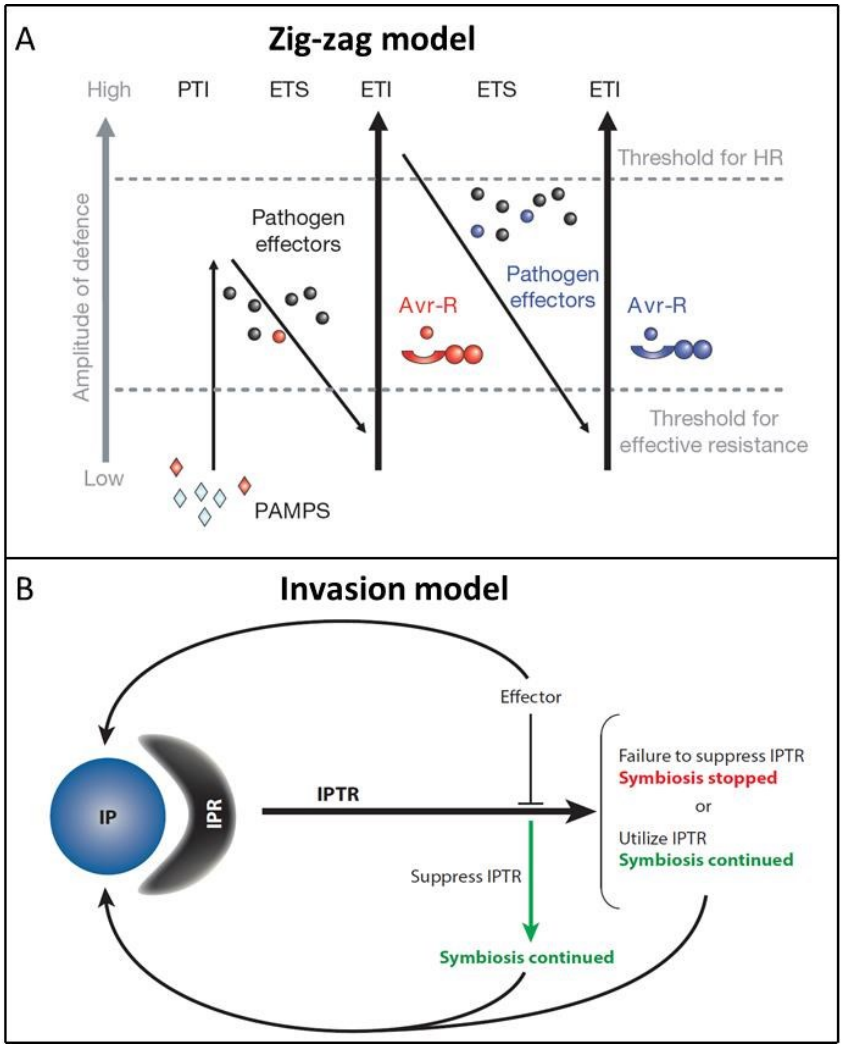


Figure 3. Plant immunity explained by Zig-zag model and Invasion model. In the Zig-zag model plant immunity is divided in two separates branches depending on the biochemical function of the ligand recognized by immune receptors in the host **(A)**. In the invasion model plant immunity result after the stimulation of immune receptors by ligands with any biochemical function **(B).** The invasion model proposes a more wide view of plant immunity and incorporates diverse interactions that are strictly separate in the Zig-zag model. Adapted from Jones & Dangl 2006; Cook et al. 2014.

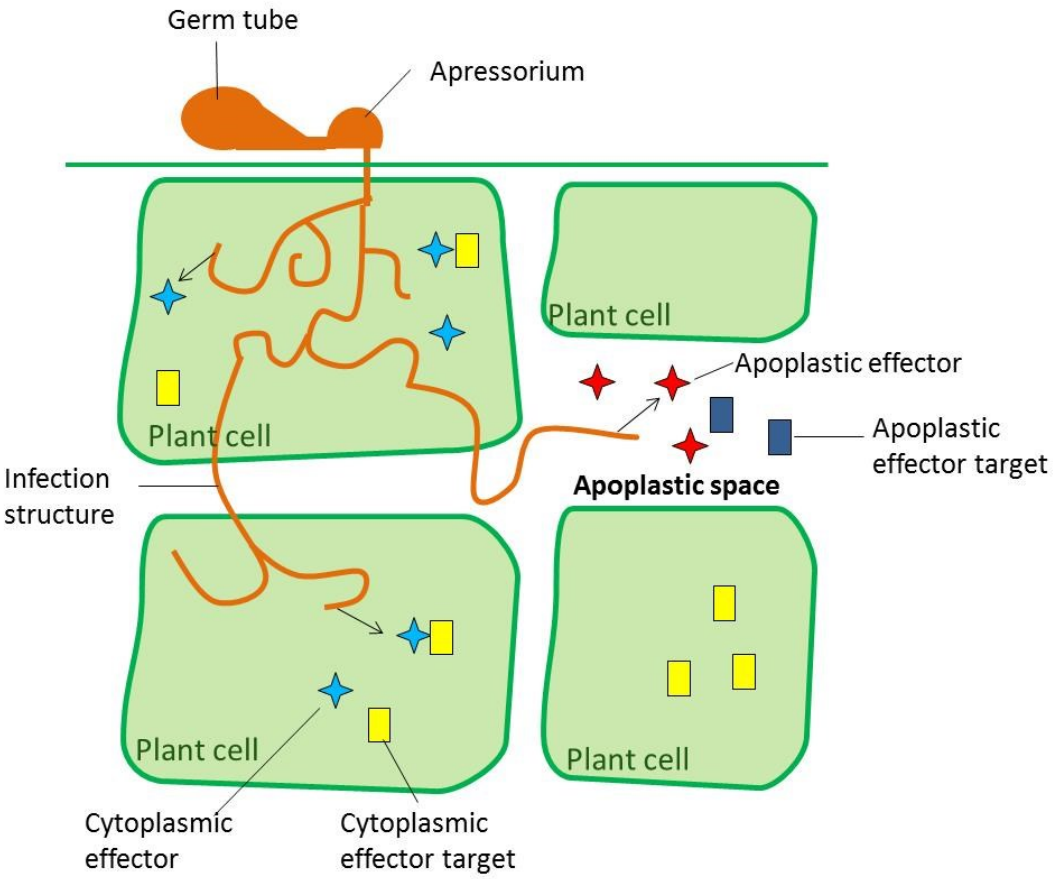
2. Fungal effectors

2.1 The importance of fungal plant diseases

Fungi are one of the most important groups of plant pathogens. They cause devastating diseases and are a permanent threat for food production. The use of resistant crops represents the economically and ecologically most sustainable solution for disease control. However, the durability of crop resistance is frequently broken by the rapid evolution of fungal populations (Brown, 2015). The current knowledge on the molecular bases of fungal pathogenicity and its evolution is still too limited to obtain durably resistant plants by knowledge-driven breeding or engineering. Indeed, resistance break-down frequently occurs only after few cultural cycles reducing the available genetic resistance resources and limiting the utility of natural resistance. Achieving durable crop resistance is a major goal for a sustainable, knowledge-based agriculture and it cannot be achieved without an improved and detailed understanding of the molecular mechanisms underlying plant–fungal interactions (Dangl et al., 2013). Since plant immune receptors and effectors molecules are pivotal components of plant defense responses, in this study our main objective is to better understand the molecular bases of the **recognition of fungal effector proteins by plant NLRs.**

2.2 Fungal effectors act either in the plant apoplast or in the cytoplasm

Fungi release an arsenal of highly diverse effectors, here defined as any secreted protein that modulates plant immunity to facilitate infection. Fungal effectors can act in the **apoplast** or in the **cytoplasm** (Figure 4) of plant cells and are very variable in sequence and distribution.



Adapted from Kamoun 2006

Figure 4. Plant pathogenic fungi secrete apoplastic effectors into the plant extracellular space and cytoplasmic effectors inside the plant cell

While most effectors are species or linege-specific, some, in particular apoplastic ones are widespread and have conserved domains. The already mentioned NLP proteins that induce cell death, necrosis and ethylene production when applied to plant leaves are e.g. present throughout a large range of phytopathogenic bacteria, fungi and oomycetes. Effectors containing a sugar-binding lysin motifs (LysMs) occur in many pathogenic and non-pathogenic fungi (Oome et al., 2014; Kombrink and Thomma, 2013). Some of these LysM effectors such as Ecp6 from the fungal tomato pathogen *Cladosporium fulvum* and Slp1 from the rice blast fungus *Magnaporthe oryzae* bind chitin with high affinity and sequester by this cell wall-derived chitin fragments to interfere with chitin detection by immune receptors and chitin-

triggered immunity in the host (Figure 7) (Sánchez-Vallet et al. 2013; Mentlak et al. 2012). Other LysM effectors such as Avr4 from *C. fulvum* protect fungal hyphae against degradation by hydrolytic enzymes secreted by the host (van Esse et al., 2007) showing that widespread effectors with conserved domains can have versatile functions.

Among the effectors that act in the host cytoplasm, only few show conserved domains that give clues about their function. An example is the secreted chorismate mutase Cmu1 from the maize smut fungus *Ustilago maydis* that acts as an effector during infection. Chorismate mutase is a key enzyme of the shikimate pathway and catalyses the conversion of chorismate to prephenate the precursor for tyrosine and phenylalanine synthesis. Cmu1 has been shown to be translocated in the host cytoplasm and to reduce levels of the plant defence hormone SA (Salicylic acid) by removing the SA precursor chorismate (Djamei et al., 2011). Secreted chorismate mutases are found in many fungal and nematode plant pathogens suggesting that this virulence mechanism to reduce plant SA levels is widespread (Bekal et al., 2003). Similarly, the effector proteins Psls1 and Vdlsc1 from the two evolutionary distant filamentous pathogens, *Phytophthora sojae* and *Verticillium dahlia*, are secreted isochorismatases required for full virulence of the pathogens and acting by reducing the amount of SA in the plant cells. Another example of a cytoplasmic effector with conserved domains is AVR-Pita from *M. oryzae* which shows homology to fungal zinc-dependent metalloproteases (Jia et al. 2000; Orbach 2000). However, the molecular function and role in virulence of AVR-Pita has not been reported.

2.3 Effector delivery and translocation

Effector delivery from the pathogen is required to allow effector functions either at the interface with the host or inside plant cells. For this, most fungal effectors harbor an N-terminal secretion signal that directs them in the eukaryotic secretory pathway which involves passage through the Endoplasmic Reticulum (ER) and exocytosis of Golgi-derived secretory vesicles. In this way, effectors are secreted and get into the apoplast (Panstruga and Dodds, 2009). However, effectors acting in the host cytoplasm have to enter into plant cells and to

pass the plasma-membrane barrier. To date, this step remains poorly understood. In oomycetes, the conserved motifs RxLR and LxLFLAK are found in a large number of effectors proteins and were shown to be required for host cell translocation (Dou et al. 2008; Schornack et al. 2010). It has been proposed that that RxLR motifs enable oomycete effectors to bind phosphatidylinositol-3-phosphate (PI3P) at the host cell surface and subsequently enter host cells through vesicle-mediated endocytosis (Kale et al., 2010). However, this was largely based on results from lipid binding assays and effector uptake assays that have been shown to be very unspecific. Therefore, this model and the corresponding data are largely debated and it is widely considered that the mechanism of RXLR effector uptake remain to be discovered. In fungi, similar conserved sequence motifs have not been identified with the exception of the powdery mildew fungi where 80% of the effector candidates share a conserved Y/F/WxC motif in the N-terminus. A role of this motif in translocation has however not been demonstrated (Godfrey et al., 2010). A clear limitation for a better understanding of effector translocation is the lack of reliable effector uptake assays. Their development represents a crucial next step and challenge in fungal and oomycete effector research (Lo Presti et al. 2015).

2.4 Evolution of fungal effector proteins

Effectors evolve to optimize the virulence function of the pathogen and to escape from plant recognition. In turn, plants evolve immune receptors that frequently act as resistance proteins to recognize effector and impair pathogen development. This continuous interaction between effectors and plant immune receptors results in certain cases in a boom-and-bust cycle (Brown and Tellier, 2011). On the one hand Avr effectors are selected when resistant genes are in low proportion but are eliminated when resistance genes are abundant, on the other hand resistance genes are selected when avr effectors are common and eliminated when avr effectors are rare. As consequence of this boom-and-bust cycle, co-evolution between effectors and resistance genes may be explained by basically two different models (Lo Presti et al. 2015). The **arms race** model in which both pathogens and plants will continuously develop new proteins, effectors and immune receptors respectively, that will temporarily fixed in the population (Dawkins & Krebs, 1979) (Figure 5). On the contrary, in the **trench warfare model**

effectors and immune receptors are maintained in populations and their frequencies oscillate over the time (Stahl et al. 1999) (Figure 5). One of the main signs of an arms race way of evolution is the high variability observed in the interacting proteins. To date, the majority of fungal pathosystems studied in agricultural context support the arms race model what may in part be explained by the constant human intervention on host varieties (Brown & Tellier 2011; Lo Presti et al. 2015).

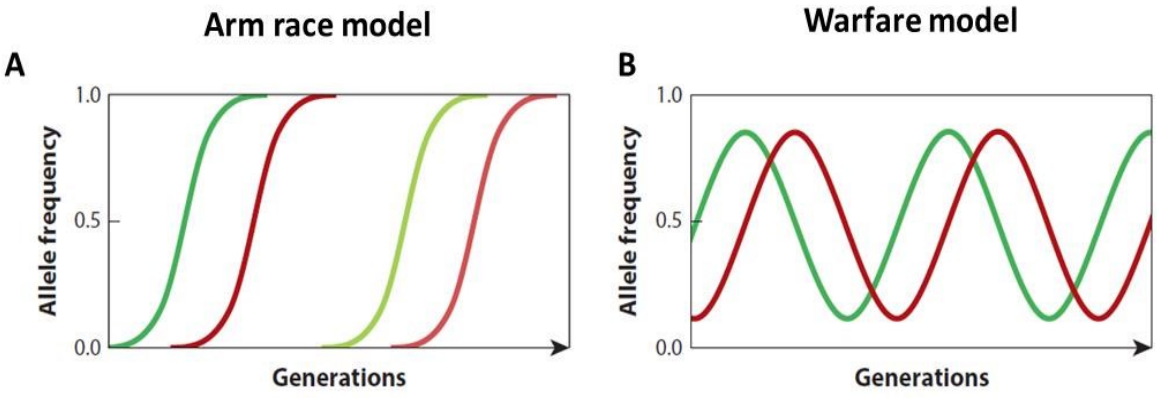


Figure 5. Co-evolutionary principles driving effector and plant target evolution. Population-wide allele frequencies of a pathogen-derived effector molecule and a host-derived interactor are represented by red and green lines respectively. Allele fixation and recurrent development of new alleles are indicated by a light-colored line in the arms model **(A)** and contrast with the fluctuation of allele frequencies in the trench warfare model **(B)**. Adapted from Lo Presti et al. 2015.

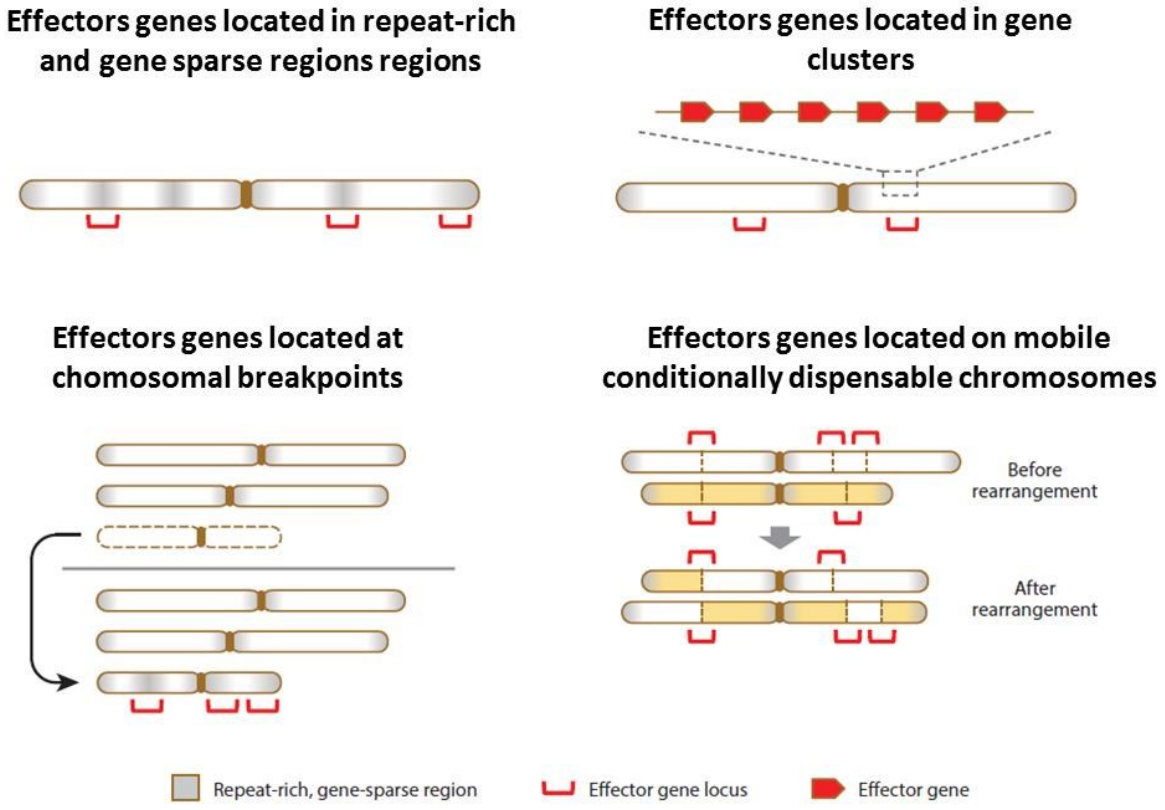
Fungal genome-wide analyses have unrevealed a huge repertoire of effector molecules highly polymorphic in their presence or absence and sequence variation e.g., single-nucleotide polymorphisms (SNPs) or insertions and deletions (indels). Fungal effectors have also been shown to exhibit signs of positive selection occurring when the ratio between the number of nonsynonymous substitutions and synonymous substitutions (dN/dS) is greater than one (Joly et al. 2010; Thines 2014; Hacquard et al. 2013). Many effectors belong to multigene families that have diversified from a common ancestor (Nemri et al., 2014; Pendleton et al., 2014; Stergiopoulos et al., 2012). These multigene families are frequently lineage specific but some of them have also been shown to be widespread across the fungal kingdom following multiple expansions (Stergiopoulos et al., 2012). Evolution of pathogen effectors by diversifying

multigene families is according to **the birth-and-death** model that proposes that new genes are created by gene duplication, and some duplicated genes are maintained in the genome for a long time, whereas others are deleted or become nonfunctional through deleterious mutations (Nei and Rooney, 2006). An example of effector diversification followed of specialization and host adaptation has been recently shown for the oomycetes pathogens *Phytophtrora infestans* and *Phytophthora mirabilis* infecting potato and *Mirabilis jalapa* respectively. In this work they showed that diversification of an effector belonging to the cystatin-like protease family changed the specificity of the pathogen toward it associated cysteine protease in the host plant leading to a host jump (Dong et al., 2014).

2.5 Localization of effector proteins in the genome

Many putative fungal effectors have been found to co-localize in distinct genome compartments (Figure 6), such as gene-sparse genomic regions with high genome plasticity and enriched with repetitive elements, accessory chromosomes and AT-blocks (Orbach 2000; Hogenhout et al. 2009). The association of effectors with plastic genomic loci constitute an important strategy to increase the genetic variation and allow an accelerated evolution and adaptation to host resistance without affecting the evolution of essential or housekeeping genes evolving a slow rates (Guttman et al., 2014). The co-localization of effectors with transposons rich regions also promotes the horizontal gene transfer, gene losses and the chimeras making of this genomic context an important hot spot for evolution of virulence traits (Rafaelle & Kamoun, 2012; Lo Presti et al. 2015). The genome of the fungal pathogen *Leptosphaeria maculans* represents one of the examples of genome organization and effector compartmentalization. In this genome e.g. the GC- rich blocks are enriched with housekeeping genes whereas the AT-blocks are gene-sparse, harbors mosaics of transposable elements and are enriched in putative effector genes (Rouxel et al., 2011; Soyer et al., 2014). Effector compartmentalization has also evidenced in the genome of *Fusarium oxysporum.* In this case it has been found that all known effector genes localize in one of the four *F. oxysporum* dispensable chromosomes which also contains lineage-specific (LS) genomic regions and transposons. Transferring of LS cromosomes between different *F. oxysporum* strains conferred

pathogenicity to non-pathogenic strains showing that gene horizontal transfer may also be one of the strategies of fungal pathogens to exchange virulence factors and accelerates evolution (Ma et al., 2010).



Adapted from Lo Presti et al. 2016

Figure 6. Effector genes reside in distinct genome compartments

2.6 Identification of effectors in fungal genomes

Small secreted proteins (<300 amino acids) are considered the most important class of fungal effectors because almost all effectors recognized by either cytoplasmic or cell surface plant immune receptors belong to this class (Hogenhout et al., 2009). The identification of these Avr-effector was achieved in most cases by genetic, map-based cloning strategies and has been accelerated over the last years by the availability of fungal genome sequences and transcriptome data.

With the advent of genome sequencing over the last 10 years, the genomes of many fungal pathogens became available allowing genome-wide searches for effector candidates that resemble these Avr-effectors. The criteria for such effector candidate searches were generally small size, frequently with a cut-off of <300 aa, presence of a secretion signal and lack of homology to other proteins (Wouw and Howlett, 2011). With these criteria, hundreds of effector candidates were identified in the genomes of individual fungal phytopathogens. In the powdery mildew fungus *Blumeria graminis,* 491 effector candidates were identified (Pedersen et al., 2012) while in the genome of *P. graminis f. sp. Tritici,* 1106 effector candidates were predicted (Duplessis, 2011). An important additional criterion to refine the search for effector candidates is *in planta* expression data since effectors show infection specific expression. Indeed, in several cases effector expression is specific to particular infection stages or to the invasion of particular plant tissues. This has e.g. been documented for *Colletotrichum higgensianum* and *C. graminearum* where different waves of stage specific effectors have been identified (O'Connell et al., 2012) as well as in *U. maydis* where host tissue specific effector repertoires were described (Schilling et al., 2014).

The recent development of dedicated bioinformatic tools allows now more precise and more reliable identification of fungal effectors. Pipelines using Markov clustering of similar sequences and hierarchical clustering with a set of at least 8 defining features for effectors allowed to identify restricted sets of high confidence effector candidates in different rust fungi and *Sclerotinia sclerotiorum* (Saunders, et al. 2012*;* Nemri et al. 2014; Guyon et al. 2014). A machine learning approach based on the extraction of sequence-derived properties characteristic for experimentally validated effectors led to the development of the EFFECTORP pipeline that predicts effector candidates in fungal secretome with high accuracy (Sperschneider, et al. 2015). Features that discriminate fungal effectors from secreted non-effectors are predominantly protein length, weight and net charge, as well as cysteine, serine and tryptophan content. EFFECTORP is particularly powerful when combined with *in planta* expression data.

Bioinformatic analysis of fungal genome sequences has also provided valuable information e.g. about important features of effector repertoires. The secretome of 84 plant fungi with different lifestyles was stablished from their genome sequence essentially based on the presence of an N-terminal signal peptide, the absence of transmembrane domains and the assignation of functional domains (Lo Presti et al. 2015). Further, fungi were grouped according with their feeding strategies and individual secretomes were classified into three different categories; plant cell wall-degrading enzyme (PCWDE), secreted proteins with functional annotation except PCWDE and secreted proteins without functional annotation. This study led to infer that the effector repertoire was determined by the life fungal style. For example, the proportion of PCWDE was higher in necrotrophs and hemibiotrophs than in biotrophs. This is consistent with the fact that both necrotrophic and hemibiotrophic fungus have to induce plant cell death to colonize plant cells, growth and reproduction. Consequently, hemibiothophs had the largest amount of secreted proteins because they included PCWDEs and putative effectors proteins whereas both biotrophs and necrotrophs had in proportion either putative effectors or PCWDEs respectively

2.7 Validation of fungal effectors

The most straightforward strategy to validate candidate effectors is the creation of loss of function mutants, e.g. by gene disruption or RNAi, and further confirmation that such pathogen mutants have reduced virulence. However, in several cases knock-out of an individual effector has little or no impact on virulence probably due to redundancies in effector function. Indeed, it is believed that multiple effectors in the huge effector arsenals of phytopathogenic fungi target redundantly host cellular pathways important for infection. Loss of individual effectors has therefore in most cases only a small impact on virulence (Guttman et al., 2014). This is e.g. illustrated in a study in *M. oryzae* where among mutants for 78 different effector candidates expressed during early infection only one, Δmc69, had an infection phenotype (Saitoh et al., 2012).

Other common strategies for the validation and functional characterization of effectors are protein-protein interaction analyses, screening for inducers and suppressors of plant cell death or PTI responses and subcellular localization (Sharpee & Dean 2016; Lo Presti et al. 2015). This have led to the validation of a steadily increasing number of fungal effectors and the elucidation of their role in virulence (reviewed in Jonge et al. 2011; Rovenich et al. 2014; Giraldo & Valent 2013; Wit et al. 2009; Stergiopoulos & de Wit 2009; Lo Presti et al. 2015a; Sharpee & Dean 2016; Selin et al. 2016). In particular large scale screens for effectors in heterologous or homologous experimental systems as cell death inducers or suppressors were quite successful and identified potential cell death inducers or suppressors. However, their localization (apoplasm or host cytoplasm) and their precise molecular activity remain unknown. For example, the transient expression of 70 candidate effector proteins from the necrotrophic apple canker fungus *Valsa mali* in *N. benthamiana* identified seven effectors suppressing plant cell death induced by BAX, a pro-apoptotic mouse protein that induces a cell death response similar to the defense-related HR (Lacomme and Santa Cruz, 1999; Li et al., 2015). Knock-out of one of these 7 effectors, VmEP1, resulted in a significant reduction in virulence of the pathogen. This suggests that the suppression of cell death usually associated with ETI, is also important in the interaction of this necrotrophic fungus with its host plant (Li et al, 2015).

Similar large-scale effector analysis revealed that effectors typically act at specific stages during infection. *Colletotrichum higginsianum* effectors involved in cell death suppression are e.g. specifically expressed during the early stages of disease whereas effectors involved in cell death induction are expressed in the late stages of disease development (Kleemann et al., 2012). Other effectors like Pep1 from *U. maydis* that accumulates in the apoplastic space, at sites of fungal cell-to-cell passages are needed during penetration into the host tissue (Doehlemann et al., 2009). Indeed, Δpep1 mutant strains were unable to penetrate epidermal cells and elicited a strong plant defense response revealing an important role for Pep1 in *U. maydis* virulence.

2.8 Function of fungal effectors

Overall, only for a limited number of fungal effectors, we have a detailed understanding of their molecular activity and their role during infection (Lo Presti et al. 2015; Selin et al. 2016; Sharpee & Dean 2016). Further optimization of high-throughput approaches and development of new methods for effector analysis are therefore now a crucial challenge to functionally characterize the large effector repertoires identified in plant pathogen fungi and to get a better understanding of their action (Sharpee and Dean, 2016). Our present knowledge on fungal effectors indicates that they employ highly diverse modes of action and that many of them target common effector host targets such as plant proteases, the ubiquitin-proteasome system or phytohormone signaling and homeostasis.

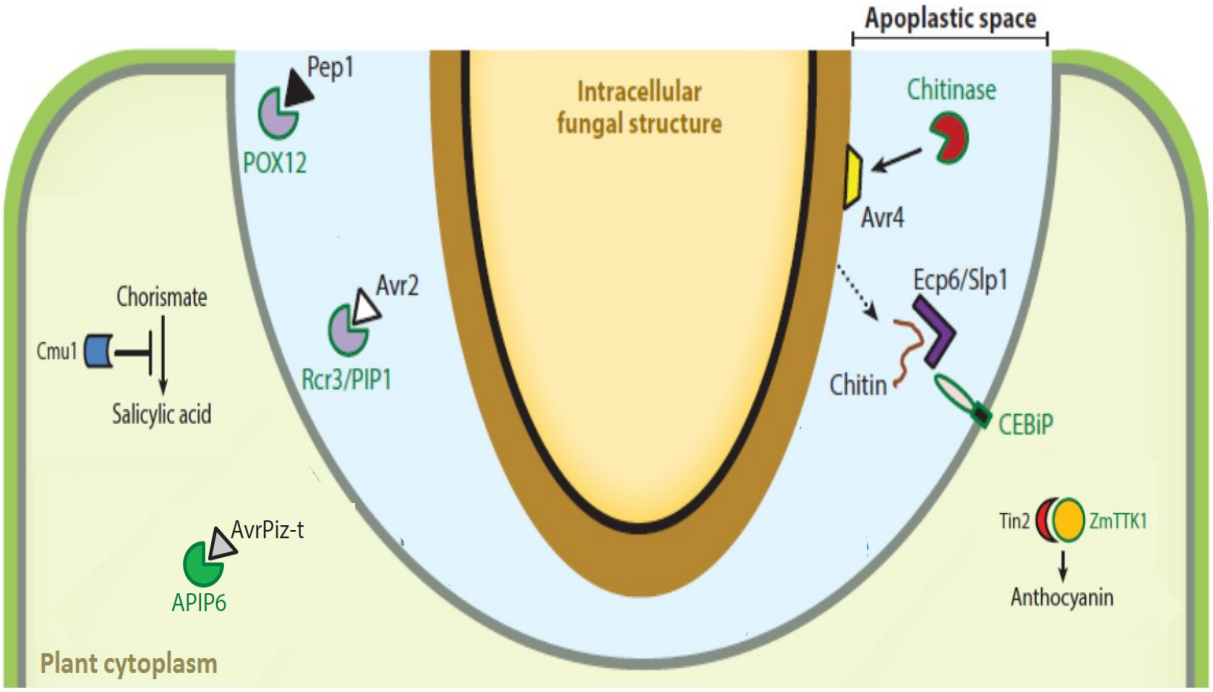
Apoplastic effectors

Most likely the functions of apoplastic effectors are related to enzyme inhibitors as well as prevention of PTI (Lo Presti et al. 2015; Selin et al. 2016; Sharpee & Dean 2016). For example, *C. fulvum* secretes Ecp6, an effector containing a LysM chitin-binding domain which binds selectively to the chitin oligosaccharides, preventing their recognition by host PRRs (de Jonge et al., 2010). To limit the release of chitin oligosaccharides, *C. fulvum* also secretes the Avr4 effector which binds to chitin in the intact fungal cell wall, preventing its hydrolysis by host chitinases (Figure 7) (van Esse et al., 2007). In other cases, apoplastic effectors act in a more direct manner in the inactivation of plant proteins involved in defense. This is e.g. the case of Avr2 also from *C. fulvum* that interacts with the plant proteases PIP1 and Rcr3 and inhibits their functions (Figure 7) (Van Esse et al. 2008). Similarly, Pit2, an effector from *U. maydis* required for virulence directly inhibits a group of apoplastic maize cysteine proteases that act downstream of SA signaling in the activation of maize defense responses (Mueller et al., 2013) and Pep1 also from *U. maydis* inhibits POX12, a peroxidase protein secreted from maize to counteract pathogen attack (Figure 7) (Doehlemann et al., 2009).

Cytoplasmic effectors

Cytoplasmic effectors employ different strategies to manipulate host cellular functions. An example is the effector protein Tin2 from *U. maydis* that targets an important secondary

metabolite pathway in plant defense. Tin2 interacts with the cytoplasmic host protein ZmTTK1 that is involved in anthocyanin biosynthesis. Tin2-ZmTTK1 interaction stabilizes ZmTTK1 and increases anthocyanin production provoking a reduction of *p*-coumaric acid levels required for the lignification of plant cell walls to form physical barriers and avoid pathogen spread (Figure 7) (Tanaka et al., 2014). Cmu1, another effector form *U. maydis* is a secreted and translocated chorismate mutase that reduces host SA levels by reducing the pool of the biosynthetic precursor chorismate as was mentioned in the previous section (Figure 7). AvrPiz-t an avr-effector from *M. oryzae* is recognized by the rice resistance gene *Piz-t* (Figure 7) (Li et al., 2009). In absence of the resistance gene, AvrPiz-t interacts with APIP6, a RING E3 ubiquitin ligase involved in plant immunity, resulting in an APIP6 loss of function (Park et al., 2012).



Adapted from Lo Presti et al. 2015

Figure 7. Effector proteins have different function and different targets in the extra and intracellular space of plant cells.

3. Plant immune receptors

3.1 Plant surface immune receptors

Plants can detect pathogen invasion by the recognition of self or nonself-derived patterns either on the cell surface or in the cytoplasm. Plant cell surface immune receptors, frequently named **Pattern Recognition Receptors (PRR)** typically recognize small epitopes from invasion patterns present in the plant apoplast (Boller and Felix, 2009; Zipfel, 2008). The two main classes of PRRs are the **Receptor Like Kinases (RLKs)** composed of an ectodomain, a transmembrane domain and an intracellular kinase domain and the **Receptor Like Proteins (RLPs)** that are similar to RLKs in their structural organization but do not possess an intracellular kinase domain (Macho and Zipfel, 2014; Böhm et al., 2014). The extracellular domains of PRRs are very diverse and can contain e.g. Leucine Rich-Repeats (LRR), lysine motifs (LysMs), lectin motifs or and epidermal growth factor (EGF)-like domains (Figure 8). LRR-type PRRs usually bind to peptides such as flg22, a fragment of bacterial flagellin whereas LysM and EGF-type PRRs recognize carbohydrate-containing molecules such as chitin, bacterial peptidoglycans, extracellular ATP, or plant-cell-wall-derived oligogalacturonides (Brutus et al., 2010; Choi et al., 2014; Kaku et al., 2006; Miya et al., 2007; Willmann et al., 2011). PRRs ligand perception occurs via ectodomains and induces the formation of PRR homo- or hetero-complexes, the activation of intracellular kinase domains and the phosphorylation of substrates that contribute to intracellular signal transduction and activation of plant defense responses.

For instance, in the model plant *Arabidopsis thaliana*, LYK5 (Lysin motif receptor kinase 5) and CERK1 (Chitin Elicitor Receptor Kinase-1) cooperate in the perception of chitin-oligomers (Miya et al. 2007; Wan et al. 2008; Cao et al. 2014). On the contrary to CERK1 that only possess moderate chitin-binding affinity, LYK5 binds chitin oligomers with very high affinity and acts as the primary chitin receptor. Chitin-binding by LYK5 induces formation of a LYK5/CERK1 hetero-complex resulting in phosphorylation of CERK1 and activation of immune signaling. Interestingly, LYK5 like other LysM RLKs involved in Nod factor receptor (NFR) perception lacks intracellular kinase activity.

In monocotyledonous plants, chitin perception seems to occur by a slightly different mechanism. In rice e.g., CERK1 is recruited by the LysM-RLP CEBiP (Chitin elicitor-binding protein) and LYP4-LYP6 receptors (Couto and Zipfel, 2016). However, in this case CEBiP appears to be the primary high affinity chitin-binding receptor(Kaku et al. 2006; Liu et al. 2012). After chitin binding, CEBiP forms a hetero-oligomeric receptor complex with OsCERK1, the rice ortholog of AtCERK1 which only possess one extracellular LysM domain and does not bind chitin. Subsequently, OsCERK1 activates chitin-mediated signaling and triggers immunity. These two examples illustrate that pathogen ligand perception mediated by PRRs can be different between plants (Shimizu et al., 2010).

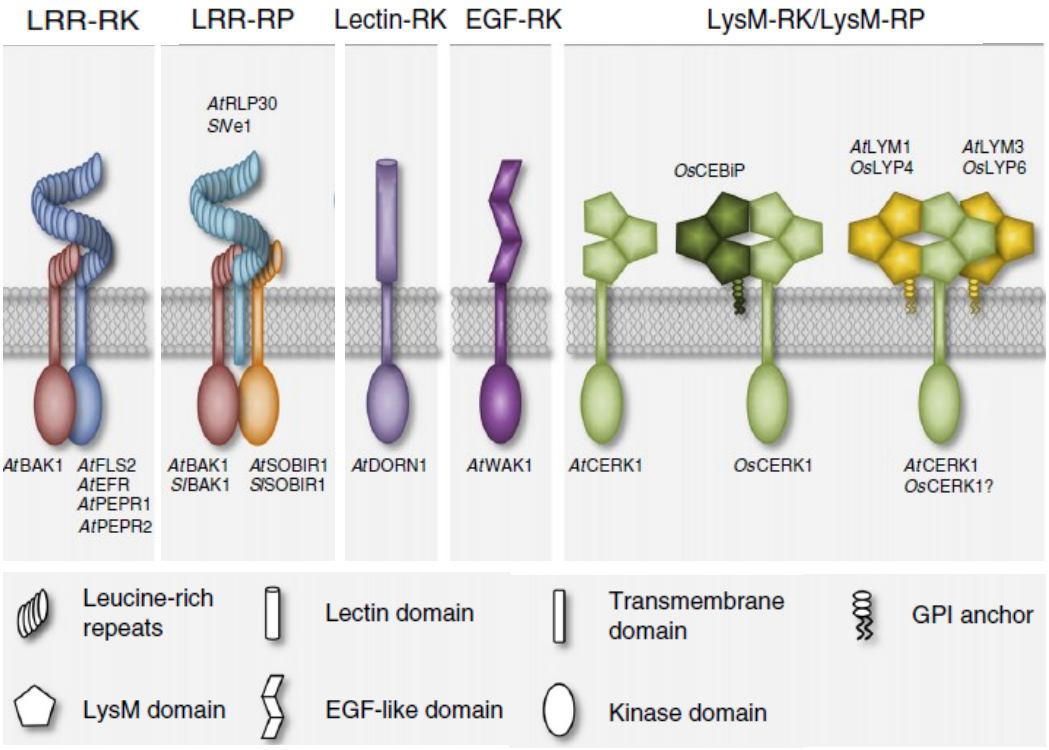


Figure 8. Plant surface immune receptors belong to two main different classes: receptor like kinases (RLK) and receptor like proteins (RLP). RLKs are composed of an ectodomain, a transmembrane domain and an intracellular kinase domain. RLP also have an extracellular and transmembrane domain but lack intracellular kinase domain. After invasion pattern perception both RLKs and RLPs hetero-complexes with downstream acting RLKs. This leads to the

activation of intracellular signal transduction and triggers plant immunity. Adapted from Bohm et al. 2014

3.2 Intracellular immune receptors

Plants possess intracellular immune receptors that mediate recognition of both modified 'host-self' and invasion patterns in the cytoplasm (Jones and Dangl, 2006; Dodds and Rathjen, 2010c; Cook et al., 2014)**.** The largest family of plant intracellular immune receptor proteins is the **Nucleotide-binding domain and leucine-rich repeat proteins (NLRs).** NLRs belong to the Signal Transduction ATPase with Numerous Domains (STAND) super family of proteins that are regulated by nucleotide-binding and intramolecular domain interactions (Lukasik and Takken, 2009; Danot, 2015). NLRs are present in all eukaryotic organisms and act as molecular switches in the regulation of various processes such as activation of immune responses, regulation of abiotic stresses and apoptosis (Goverse 2012; Jacob et al. 2013; Collier & Moffett 2009; Bernoux et al. 2016; Takken &).

NLRs present a modular architecture with a central nucleotide-binding domain (NB-ARC), a Leucine-rich repeat domain (LRR) at the C-terminus and a coiled-coil (CC) or a Toll Interleukin-1-like receptor (TIR) domain at the N-terminus (Jacob et al., 2013). Typically, NLRs are arranged in TNL or CNL (TIR (T), CC (C), NB-ARC (N), and LRR (L)) configuration but alternative configurations such as "truncate" forms, TNTNL, TNLT, CNNL, TCNL can be also found (Figure 9) (Meyers et al., 2003, 2002; Jacob et al., 2013). In solanaceas an N-terminal domain different to CC and TIR has been frequently found. This domain is called SD (solanacea domain) because is restrict to *Solanaceae* and is usually found in SDCNL configuration (Lukasik-Shreepaathy et al., 2012).

NLRs can also carry non-canonical domains integrated at low frequencies (Césari, et al. 2014; Sarris et al. 2016; Kroj et al. 2016). These integrated domains correspond to a wide range of molecular and functional categories such as signal transduction proteins, transcription factors or metabolic enzymes. Interestingly, many of them are present in regulators or actors of plant

immunity and/or in targets of pathogen effectors. For instance, RIN4 one of the most studied effector targets which interacts with many NLRs including RPS2 and RPM1 has been found to be fused to different NLRs in different plants such as rice, barley and apple (Sarris et al., 2016). In addition, many of the integrated domains fused to plant NLRs have been found to interact with effectors in targeted studies or effector interactome screens (Sarris et al., 2016). This suggests that effector recognition is a general feature of integrated domains similar to what has been demonstrated experimentally for Pik-1 and RRS1 (Maqbool et al., 2015a; Le Roux et al., 2015; Sarris et al., 2015).

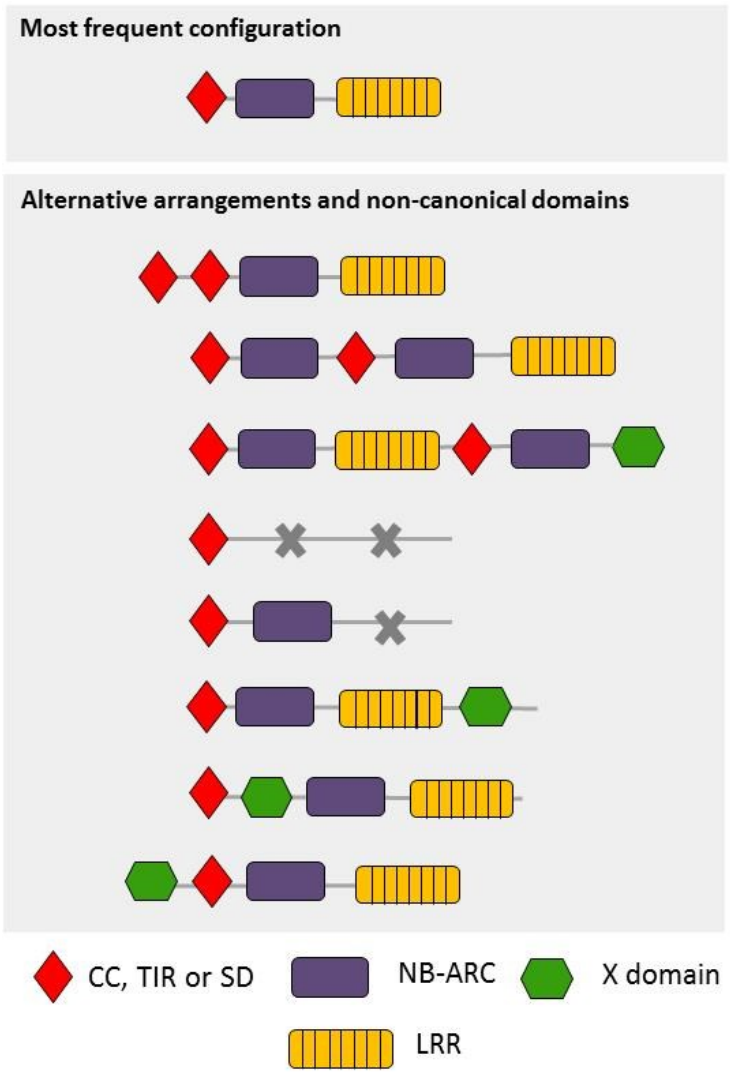


Figure 9. Intracellular immune receptors present flexible structures with different arrangements of conserved domains.

The majority of NLRs present a TIR-NB-ARC-LRR or CC-NB-ARC-LRR configuration but alternative organization of these domain are also found. Many different non-canonical domains can also been fused to NLRs mainly in the C-terminus but also in different other positions such as the N-terminus or between the CC or TIR and the NB-ARC. Adapted from Jacob et al 2013

3.3 Function and structure of canonical NLR domains

CC and TIR – the N-terminal domains

Functional analysis of **CC** and **TIR** domains suggest that both domains are involved in downstream signaling and cell death triggering (Qi and Innes, 2013). Transient expression of the TIR_{1-248} fragment from the flax resistance proteins L10, L6, L2 and L7 triggered effector-independent cell death in flax leaves (Frost et al., 2004; Bernoux et al., 2011b). Point mutations in conserved amino acids of the L6- TIR_{1-248} domain abolished cell death induction but did not affect the interaction between L6 and its cognate avirulence effector protein AvrL567 from *Melampsora lini* showing that the TIR domain from L6 is not required for effector binding but is necessary and sufficient for immune signaling (Bernoux et al., 2011b). Furthermore, it was shown that the L6-TIR domain self-associates and forms homodimers that are required for signaling since L6-TIR mutants impaired in self-association lost the ability to trigger cell death (Bernoux, et al. 2011).

Similarly, the CC domain of the barley resistance protein MLA is sufficient to induce cell death and self-associates *in vivo*, even in the absence of the Avr effector. As with the L6-TIR domain, self-association of the MLA-CC domain is required for immune signaling (Maekawa, et al. 2011). However, isolated CC or TIR domains are not always able to trigger cell death and, in certain cases, others NLR domains seem to activate immune signaling. For example, the overexpression of the of the CC domain of the resistance proteins Rx or RPS5 from potato and *A. thaliana* respectively did not induce cell death but the overexpression of the Rx NB domain did suggesting that in some cases the NB domain is engaging downstream signaling components and triggers immune signaling (Rairdan et al., 2008; Ade et al., 2007). These

different results may reflect important differences about the signaling mechanisms used by NLRs to trigger immunity.

Furthermore, many effector targets such as Pto, RIN4 and PBS1 has been shown to interact with the N-terminal domain of resistance proteins (Mackey et al., 2003; Mucyn et al., 2006; Ade et al., 2007) indicating that this domain is involved in effector recognition. For instance, N-terminal domain SD of the tomato resistance protein Prf interacts with the kinase effector target protein Pto and mediate recognition to the effector proteins AvrPto and AvrPtoB from *Peudomonas syringae* (Mucyn et al. 2006; Balmuth & Rathjen 2007; Saur et al. 2015). Similarly, the CC domain of the resistance gene RPS5 has been shown to associate with PBS1 prior to PBS1 cleavage by the effector protein AvrPphB (Ade et al., 2007; Qi et al., 2012).

To date only the crystal structures of the isolated MLA10 and Rx1 CC domains have been resolved (Maekawa, et al. 2011; Hao et al. 2013). Despite the sequence similarity between these domains, structural analyses revealed that both domains adopt different topologies. The MLA10-CC domain is formed by three different α-helices connected by loops and forms homodimers (Figure 10) whereas CC from Rx1 consists of a more compact fold composed of four different α-helices in a helix bundle (Figure 10) that does not dimerize but instead interact with the conserved domain WPP (Trp-pro-pro) of the RanGAP2 protein which is required for Rx1 function.

The crystal structures of the TIR domain from the plant NLRs L6, RPS4 and RRS1 has been determined (Bernoux, et al. 2011; Williams et al. 2014). These domains consist of a flavodoxin-like fold formed by a five-stranded parallel β-sheet surrounded by five α-helices (Figure 10). Surfaces involved in the formation of L6 and RPS4 TIR domain homodimers and RRS1/RPS4 TIR domain heterodimers has been also identified (Bernoux, et al. 2011; Williams et al. 2014).

NB-ARC the central switch domain

The **NB-ARC domain** is the most conserved domain in plant and animal NLRs and acts as a molecular switch that translates the recognition of effector proteins into signal initiation through intra and intermolecular interactions and nucleotide binding (Hu et al. 2013; Takken & Goverse 2012). Indeed, the NB-ARC domain consists of a nucleotide-binding pocket. In the inactive or "off" state, the nucleotide-binding pocket adopts a "closed" configuration where ADP is preferentially bound and stabilizes the closed structure by mediating intramolecular interactions. The activation or "on" state of the NB-ARC is mediated by the release of the ADP which is replaced by ATP (Bernoux et al., 2011a; Williams et al., 2011b). After ATP-binding the intramolecular interactions of the protein are modified and the pocket adopts an "open" configuration that is required to mediate defense responses. Thus, the central NB-ARC domain functions as a molecular switch that fluctuate between an "off" and "on" state depending on ADP- or ATP-binding respectively. The most conserved part of the NB domain is the Walker-A motif, also called phosphate-binding loop (P-loop) which is a glycine-rich flexible loop that is crucial for ATP-binding. Indeed, the p-loop coordinates together with other amino acids a magnesium cation that binds the β and ϒ phosphates and thereby properly positions ATP. In addition, a highly conserved lysine residue in the P-loop interacts directly with the β and ϒ phosphate groups of the nucleotide and is indispensable for its binding (Walker et al., 1982).

Mutations in the NB-ARC domains that weaken ATP binding or stabilize the fixation of ADP result in loss-off function while mutations that weaken ADP-binding or strengthen ATP binding lead to gain of function, autoactive NLR mutants. One of the conserved structural motifs that can be distinguished in the NB-ARC domain is the MHD (Met-His-Asp) motif. Direct mutagenesis in MHD motif of many NLRs result in a spontaneous induction of the defense and cell death responses (Bendahmane et al. 2002; Howles et al. 2005; van Ooijen et al. 2008; Williams, et al. 2011). For instance, site-directed mutation of key residues within the MHD motif of the resistance protein M resulted in autoactivation and ATP binding whereas mutations in important residues within the P-loop motif of M resulted in a loss of nucleotide binding and the inactivation of the resistance protein (Williams, et al. 2011)

To date, the three dimensional structure of an NB-ARC domain from a plant NLRs has not been determined but different models have been proposed based in the crystal structures of the NB-ARC domains from the human and *Caenorhabditis elegans* proteins Apaf-1 and CED-4 respectively, that are both involved in apoptosis (Zou et al. 1997; Yan et al. 2005; Shiozaki et al. 2002). These structures show that the NB-ARC domain consists of three structural subunits, notably NB, ARC1 and ARC2 that together form the nucleotide-binding pocket (Figure 10).

LRR the C-terminal domain

The LRR domain is defined by a repeated sequence motif in which hydrophobic residues that are often leucine alternate with hydrophilic residues in a fixed pattern (consensus motif LxxLxLxxNxL). It is the most polymorphic domain of plant NLR proteins and has been demonstrated to play a crucial role in the recognition of effector proteins in several NLR resistance proteins such as RPP1, RPS5 and L (Dodds et al. 2006;Qi & Innes 2013; Steinbrenner et al. 2015). For example, seven variants of the effector AvrL567 from the flax rust fungus *Melampsora lini* are differentially recognized by the extremely similar NLR proteins L5, L6 and L7 from flax. The specificity of this recognition was shown to be determined by polymorphic amino acids in the LRR domains of the different L variants. This matching specificity between AVRL567 alleles and the LRR domain of L alleles may be due to direct, receptor-ligand type interaction since physical interaction between the LRR domain from L and AvrL567 variants was demonstrated by yeast two hybrid assays (Dodds et al. 2006). Similarly, the recognition of different alleles of the ATR1 effector from the downy mildew pathogen *Hyaloperonospora arabidopsidis* was explained by polymorphic amino-acids and the number of leucine repetitions in the LRR domain of RPP1 alleles from different *A. thaliana* accessions. The specificity of this recognition was recapitulated by co-immunoprecipitation experiments (Rehmany 2005; Krasileva et al. 2010).

In addition to their role in signal perception, the LRR domain plays also an important role in keeping NLR proteins in a “off” state since deletions or mutations in LRR domains frequently

result in auto-activation (Qi et al., 2012; Ade et al., 2007; Tameling et al., 2010). This dual role was very convincingly shown by systematic swapping of polymorphic sites between the potato immune receptors Gpa2 and Rx1 that possess high sequence homology but different recognition specificity. A small region in the ARC2 and N-terminal repeats of the LRR domain was found to control the activation state of the proteins whereas the C-terminal LRR repeats were found to determine the specificity of the recognition of the nematode *Globodera pallida* by Gpa2 and of Potato virus X by Rx1 (Slootweg et al., 2013). Furthermore, *in silico* modeling combined with structure-informed functional analysis revealed that the opposite charge distribution between the N and C terminal part of the LRR domain determines the dual function of the LRR as an effector sensor and regulator of NLR activity and that intramolecular interactions between the N-terminus of the LRR and the NB-ARC from Gpa2 and Rx1 are mediated by conserved basic residues in the LRR (Slootweg et al., 2013). In RPS5, truncations in the N-terminal part of the LRR showed that the first four repeats are sufficient to inhibit autoactivation but that a full length LRR domain was required to mediate effector recognition. In addition, swapping of the LRR domain between RPS5 and the closely related NLR RPS2 resulted in autoactivation of RPS5, demonstrating a fine tuned co-evolution between the LRR and NB-ARC domains (Qi et al., 2012).

The 3 dimensional structure of LRR domains from plant NLRs has not yet been determined but those from other LRR domain proteins and in particular from animal NLRs are available (Hu et al. 2013). Basically, LRRs display an arc-shape or horse-shoe-like structure where the repeated leucine-rich motifs form a parallel β-sheet at the concave side (Figure 10). The leucine form the hydrophobic core of the β-sheet and the other residues are exposed to the surface. This surface of the concave side is supposed to be engaged in the intramolecular interactions with the NB-ARC domain in the N-terminus of the LRR and to establish protein-protein interactions with ligands that mediate effector recognition in the C-terminus of the LRR. These different functions of the 2 different parts of the LRR are reflected, as already mentioned, by the dual structure of the LRR with an N-terminal part dominated by positively charged amino acids and a C-terminal part enriched in aromatic residues probably involved in hydrophobic interactions

(Takken & Goverse 2012). The sequences between the LRR repeat motifs form the convex side of the arc and can form different secondary structures. Compared to other LRR proteins, LRRs in plant NLRs are diverse and irregular presenting strong variation in the number and the length of repeats.

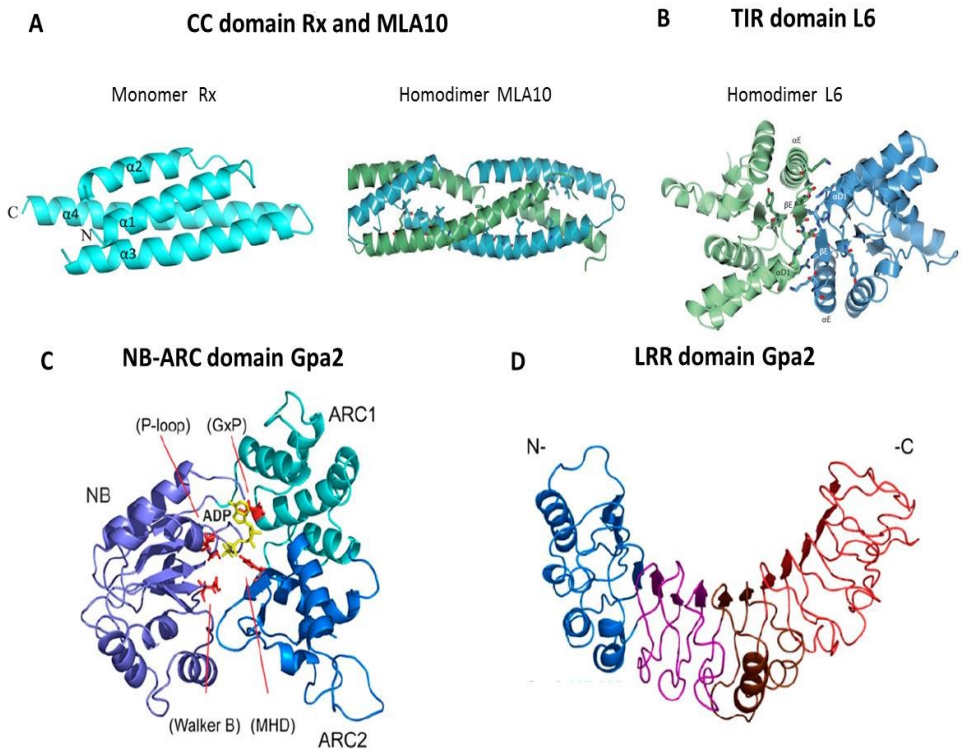


Figure 10. Structure of NLR domains. Cartoon representing the crystal structure of the CC domain of RX and MLA10 **(A)** and the TIR domain of L6 **(B)**. MLA10 and L6 forms homodimers that are represented in green and blue. The residues that have been shown to be important for stability and/or dimer formation in each protein are highlight with sticks. Three-dimensional models are represented for the structure of the NB-ARC **(C)** and LRR domains **(D)** of Gpa2 generated by homology modeling based on the Apaf1 crystal structure. The subdomains NB, ARC1 and ARC2 are represented by different shades of blue and the main residues contributing to the nucleotide-binding pocket are indicated in red **(C)**. The NB-ARC domain is represented in a closed conformation bound to a yellow ADP molecule. The N and C terminal parts of LRR domain from Gpa2 involved in NB-ARC interaction and effector recognition respectively are represented by different colors **(D)**. Adapted from Slootweg et al. 2013; Hao et al. 2013; Maekawa, et al. 2011

3.4 Recognition of pathogen attack

NLR receptor proteins recognize pathogen effectors by two different mechanisms (Figure 11). Either, they directly bind the effector protein alone, or in complex with a host protein that serves as a co-factor or they recognize host target protein modifications induced by effector proteins (Chisholm et al., 2006; Jones and Dangl, 2006; van der Hoorn and Kamoun, 2008). In the first case, effector recognition occurs in a receptor-ligand-like mode and in the second case, recognition is indirect and follows the guard or decoy model.

Receptor-ligand model

Direct interaction between effectors and resistance proteins has been demonstrated in few cases. The best documented examples are L6 and Pi-ta from flax and rice respectively (Figure 11) (Dodds, et al. 2006; Yulin Jia, McAdams, et al. 2000) L6 has been shown to recognize the AvrL567 effector protein in an allele-specific manner and the specificity of this recognition has been recapitulated by the physical interaction of the effector protein with the corresponding resistance protein over a spectrum of alleles and mutants in yeast two hybrids assays (Dodds, et al. 2006). For ATR1 recognition by RPP1, a similar direct interaction model has been proposed since the recognition of ATR1 alleles by matching RPP1 alleles correlates with co-immunoprecipitation of the protein (Krasileva et al. 2010). However, an important role of additional host proteins e.g. co-factors that are required to mediate ATR1 and RPP1 interaction cannot be excluded by these experiments. Involvement of a host-cofactor in effector recognition has been shown in the case of the NLR N from tobacco that recognizes the p50 protein of tobacco mosaic virus. P50 forms a protein complex with N as demonstrated by co-immunoprecipitation assays. However, this association and the resulting immune activation requires the host protein NRIP1 that binds on the one hand p50 and on the other the TIR domain of N (Peart et al. 2005; Ueda et al. 2006; Burch-Smith et al. 2007; Caplan et al. 2008).

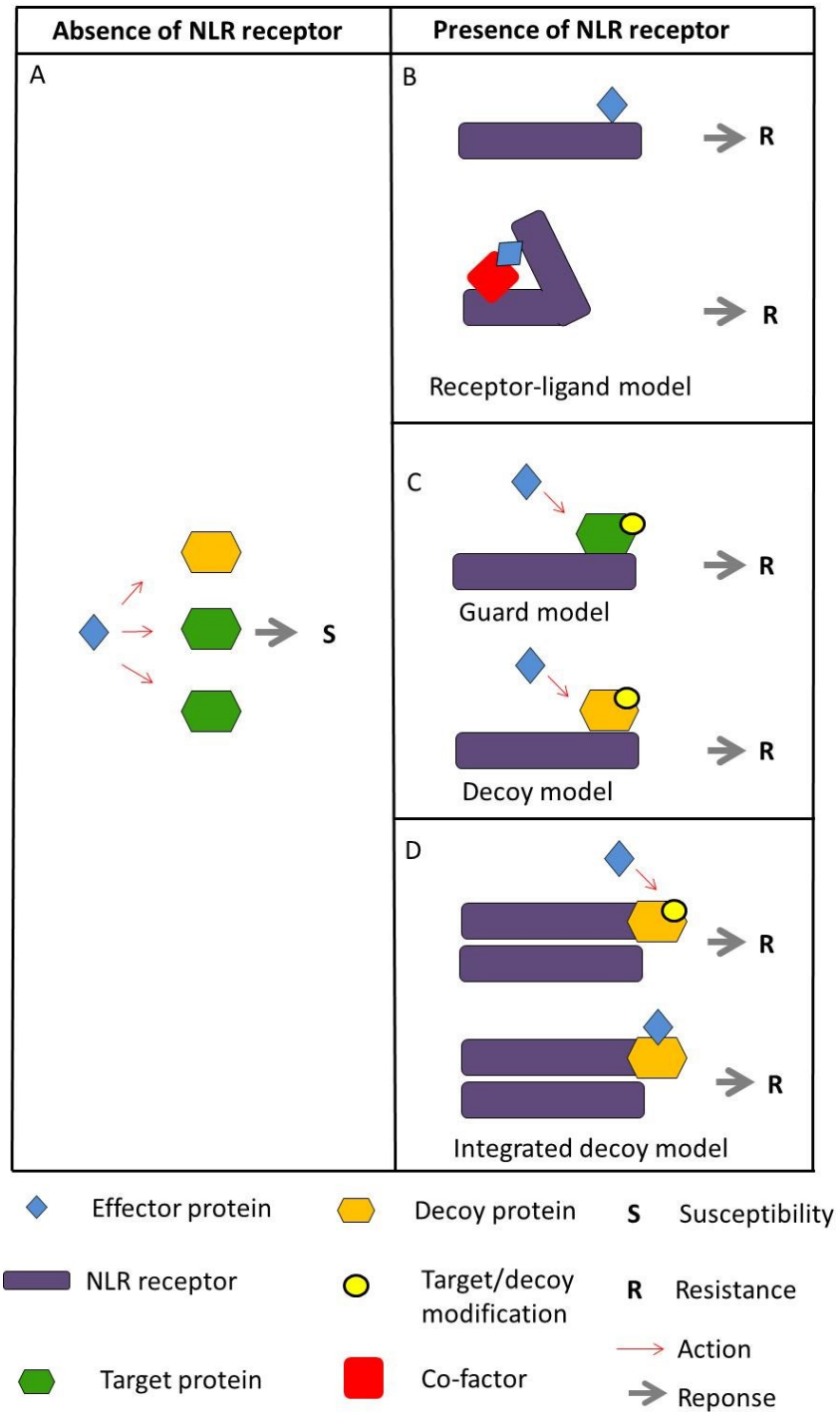


Figure 11. Modes of effector protein recognition by NLRs. In the absence of NLRs acting as resistance proteins, effectors target host proteins to cause disease **(A)**. Resistance plants possess NLR immune receptors that recognize effector proteins by direct association **(B)** or by sensing modifications such as degradation or phosphorylation of target or decoy proteins **(C)**. Decoys can also be fused to NLRs to allow efficient effector recognition **(D)**.

Adapted from Khan et al. 2016

The guard/decoy model

Alternatively, NLRs can recognize effectors in an indirect manner that does not involve direct association but relies on the sensing of effector-triggered modifications in host proteins.These effector-modified and NLR-guarded host proteins may be either the true effector targets frequently involved in plant immunity or mimics of the real target that solely monitor effector activity . In the first case, the host proteins surveyed by the NLR and modified by the effector are called guardees, in the second case they are called a decoys (Dangl and Jones, 2001; Dangl and McDowell, 2006; van der Hoorn and Kamoun, 2008). Decoy proteins are thought to evolve from the real targets to avoid the conflict of having two opposite selection pressures acting in the same protein (Van der Hoorn & Kamoun 2008). The corresponding modes of recognition are described by the so-called guard model or the decoy model.

One of the best characterized examples of the **guard model** is the indirect recognition of the effector proteins AvrRpt2 by RPS2 and AvrRpm1/AvrB by RPM1 (Chung et al., 2011; Day, 2005; Kim et al., 2005) (Figure 12). In this case, both immune receptors, RPS2 and RPM1 recognize the presence of their cognate effector proteins by sensing modifications in the host target protein RIN4. RIN4 is involved in plant immunity and contributes to pathogen fitness in absence of RPS2 and RPM1. As a consequence, the guarded effector target RIN4 is exposed to opposing natural selection forces. On the one hand, in the absence of a functional resistance protein, it is favorable to change to evade modification by the effector. On the other hand, when a guarding NLR is present, interaction of the guarded effector target RIN4 with the effector protein is beneficial to enable pathogen perception.

One of the best characterized systems operating according to the decoy model is the tomato NLR Prf that recognizes the effector proteins AvrPto and AvrPtoB from *Pseudomonas syringae* by sensing the modifications in the protein kinase Pto (Figure 12). Pto does not have a role in immunity and seems to act as a decoy for the true AvrPto and AvrPtoB protein kinase targets. Indeed, both AvrPto and AvrPtoB target multiple RLKs including CERK1, EFR1 and FLS2 that

have an established role in immunity and thereby contribute to virulence of *Pseudomonas syringe* (Mucyn et al., 2006; Gutierrez et al., 2010; Ntoukakis et al., 2014). Similarly, the NLRs RPS5 and ZAR1 recognize the effector proteins AvrPphB and AvrAC respectively, by sensing perturbations in the decoy kinase proteins PBS1 and PBL2 (Shao et al., 2003; Ade et al., 2007; Wang et al., 2015)

Integrated decoy model

Recently, a third, intermediate mode of action has been discovered in NLRs that carry unconventional integrated domains (Césari et al. 2014). It was shown that the unusual domains that are found in these NLRs mediate effector recognition by either binding them directly or being modified by them. From this, it was concluded that they act as decoy proteins that are integrated into the immune receptor to mediate pathogen perception (Ellis, 2016). An example that supports the integrated decoy model is the recognition of the effector protein PopP2 by the NLRs pair RRS1/RPS4 (Figure 12). PopP2 is an acetyltransferase that inactivates host WRKY transcription factors involved in plant immunity by acetylating a conserved lysine in their DNA-binding domain (Le Roux et al., 2015; Sarris et al., 2015). RRS1 carries a WRKY domain that is not directly involved in the activation of immunity but is acetylated by PopP2 and is crucial for PopP2 perception. Therefore, the WRKY domain of RRS1 is interpreted as a decoy for the true WRKY targets of PopP2 that has lost its original function as a transcription factor but is still able to assist RSS1 in the perception of the biochemical activity of PopP2 (Figure 12) (Le Roux et al., 2015; Sarris et al., 2015). The RRS1 WRKY domain is also involved in the perception of a second effector, AvrRps4. However, the precise mechanisms of AvrRPS4 recognition and the true targets of this effector remain to be identified.

In rice, the NLR Pik-1 possesses an integrated HMA (heavy metal-associated) domain that is involved in the perception of the effector protein AVR-Pik from *M. oryzae*. Functional analysis showed that the HMA domain of Pik-1 interacts directly with surface exposed residues of the avirulence effector and that the affinity of this binding determines effector recognition (Maqbool et al. 2015). In addition, recent work in the group of Dr. Terauchi (Iwate, Japan)

showed that rice HMA proteins are virulence targets of AVR-Pik (personal communication, unpublished). Therefore, the HMA domain of Pik-1 seems to act indeed as a decoy for the host targets of AVR-Pik. An evolutionary advantage of this mode of recognition could be that it avoids the segregation between the decoy host target and the NLR receptor.

3.5 Frequently, NLRs pairs mediate disease resistance

In many cases single NLRs seem sufficient to mediate disease resistance. However, recently an increasing number of cases were reported where the resistance response to a single *Avr* gene product is mediated by **NLRs pairs** (Eitas and Dangl, 2010b). The first example of disease resistance conferred by NLR pairs were RPP2A/RPP2B from *A. thaliana* that are both required for immunity against certain *Hyaloperonospora arabidopsidis* isolates (Sinapidou et al., 2004). Similarly, the NLRs pairs, N/NRG1 and RPM1/TAO1 mediate recognition of the effector proteins p50 from Tobacco Mosaic Virus and AvrB from *P. syringae* respectively (Peart et al., 2005; Eitas et al., 2008). Other NLR pairs such us Lr10/RGA2, Pi5-1/Pi5-2 and Pikm1/Pikm2 mediate resistance to fungal pathogens (Loutre et al., 2009; Lee et al., 2009; Ashikawa et al., 2008). Furthermore, the pairs RGA5/RGA4 from rice and RRS1/RPS4 from *A. thaliana* recognize each at least two unrelated Avr-effectors (Cesari et al., 2013; Narusaka et al., 2009). These examples indicate that resistance mediated by NLR pairs is widespread and confers immunity to different and in some cases multiple pathogens (Figure 13).

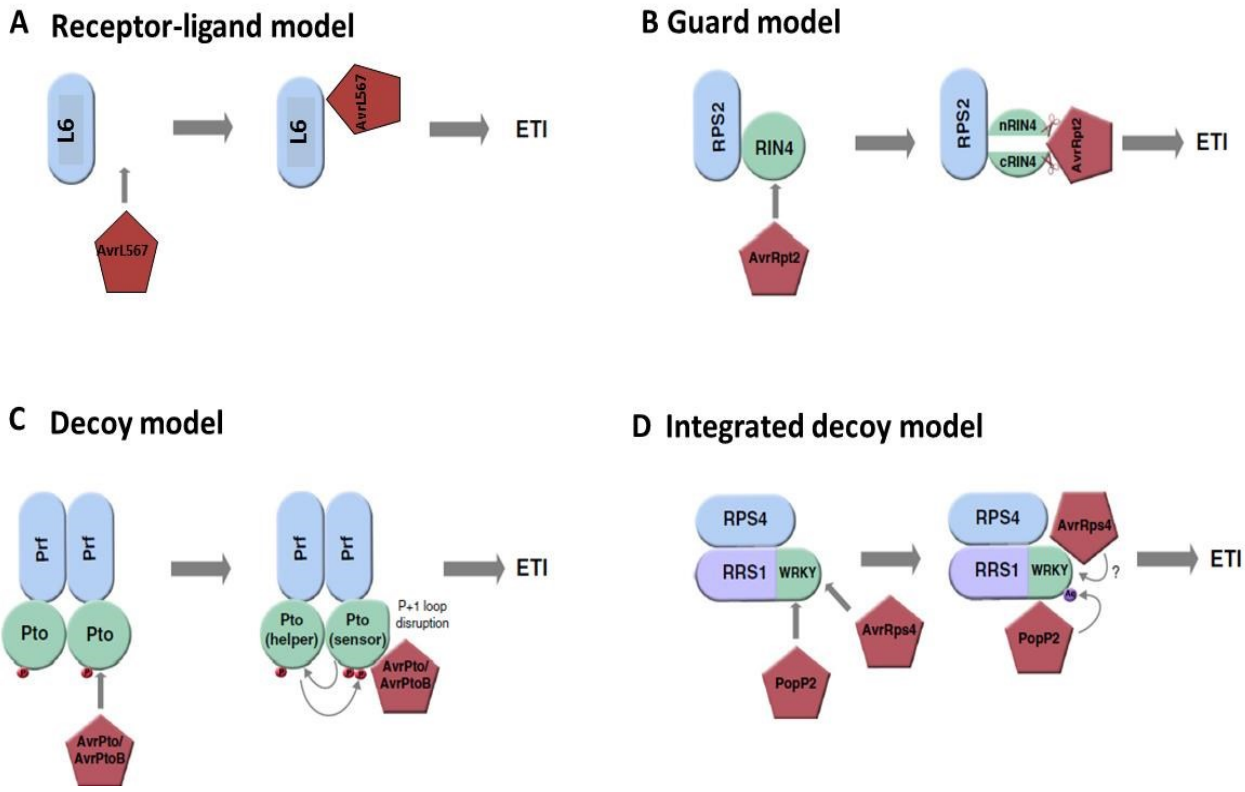
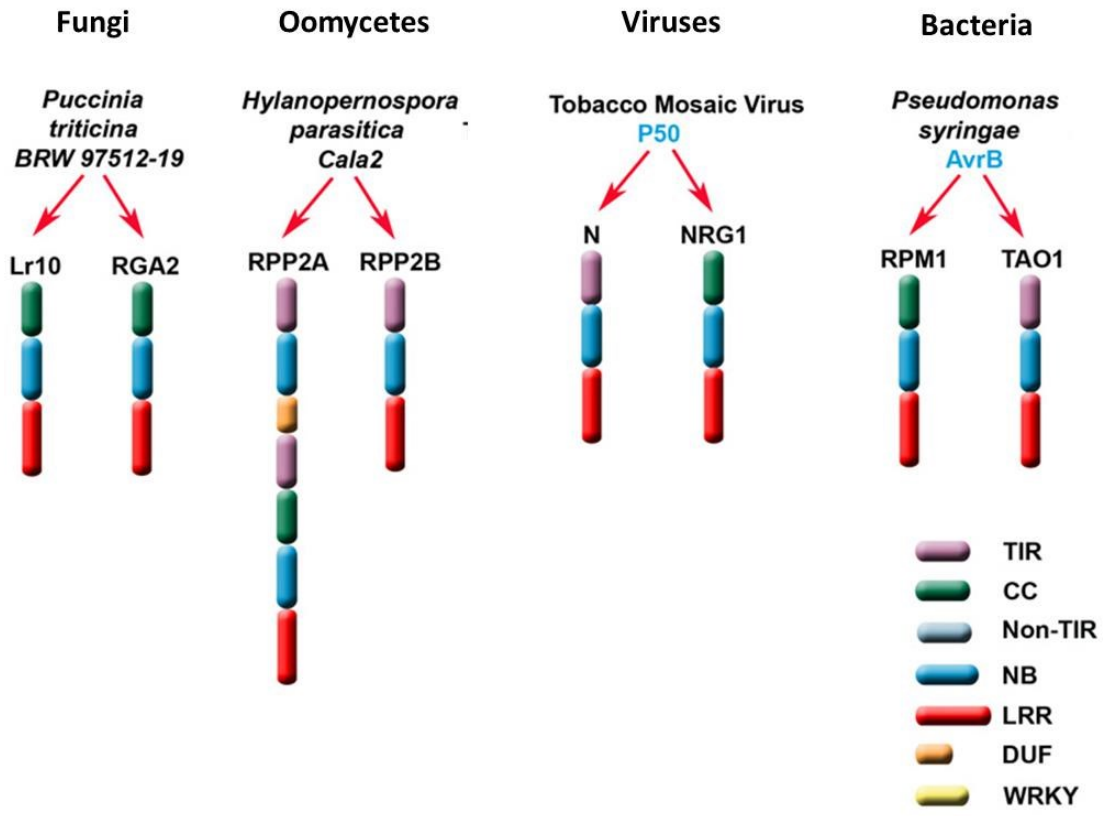


Figure 12. Examples of the different recognition model. (A) L6 directly interacts with AvrL567 to activate resistance. **(B)** RPS2 senses the cleavage of the guard protein RIN4 by AvrRpt2 and induces ETI. **(C)** Prf monitories the phosphorylation of the decoy protein Pto induce by AvrPto and AvrPtoB to induce resistance. **(D)** The interacted WRKY domain in RRS1 mediates recognition to AvrPps4 and PopP2 to trigger ETI. Adapted from Khan et al 2016



Adapted from Eitas & Dangl 2010

Figure 13. NLR pairs mediate resistance to very different groups of pathogens.

The genes coding for NLR pairs are genetically tightly linked and show generally a very typical clustered genomic organization characterized by an inverted tandem arrangement with a short shared 5' intergenic region (Figure 14). The only exception is the *RPP2A*/*RPP2B* pair, where both genes are adjacent in the genome but not arranged in an inverted orientation (Figure 14). This characteristic genomic organization is thought to contribute to a concerted regulation of the transcription of the two genes of one pair and to avoid their separation by segregation and recombination which would cause resistance loss and may result in autoimmune phenotypes (Eitas & Dangl 2010).

Recent studies on RPS4/RRS1 and RGA4/RGA5 gave first insight into the mechanism of effector recognition by paired NLRs. In both cases the NLR pairs function as receptor complexes in which one partner (RRS1 or RGA5) acts as a receptor or sensor of the effector while the other

(RGA4 or RPS4) acts as a cell death executor (Bernoux et al., 2014; Williams et al., 2014). In animals, a comparable mechanism of NLR functioning in pairs has been revealed in the case of NAIP/NLRC4 in mice. NAIP act as primary immune receptor and directly bind non-self ligands. This recognition event induces hetero-complex formation with NLRC4, oligomerization of NLRC4 and activation of defense signaling (Kofoed and Vance, 2011; Tenthorey et al., 2014).

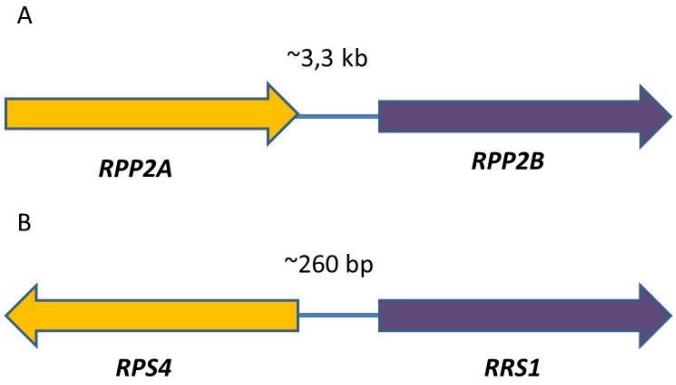


Figure 14. Examples of NLRs genetically linked. (A) Arabidopsis TNL pair in tail-to-head orientation. (B) Arabidopsis TNL pair in head-to-head orientation. Adapted from Griebel et al. 2014.

3.6 Activation of NLR resistance proteins

Since NLR proteins participate in the recognition of effector proteins and resistance induction, a fine-tuned regulation of both tasks is needed to avoid autoimmune responses due to unnecessary activation in the absence of a pathogen whereas a rapid response is expected upon pathogen attack (Takken and Tameling, 2009). In the absence of pathogens, NLRs predominately occur in a closed inactive conformation that is determined on the one hand by binding of an ADP molecule and on the other hand by intra and intermolecular interactions (Sukarta et al., 2016). How NLRs switch upon pathogen perception from this inactive state to the active conformation is still not well understood. In the cases where effectors interact directly with the NLR, the activation has been explained by mostly three different models, a simple “switch” model, the “bait and switch model” and the “equilibrium-based switch” model (Sukarta et al., 2016).

The **switch model** is based on the assumption that the NB-ARC domain acts as a molecular switch to regulate the activity of the NLR receptor (Takken et al., 2006; Faustin et al., 2007). Recognition of an effector by the inactive NLR which is in the ADP-bound/OFF state has consequences on the entire protein and results in a conformational change that triggers exchange of ADP for ATP and transition to the active/ON state and activate defense signaling (Figure 15). The conformational change is supposed to affect in particular the surfaces of the N-terminus of the LRR and the ARC2 that mediate inhibitory interactions and the alteration in the nucleotide binding status, from ADP to ATP is believed to lead to an open configuration of the NLR that exposes the N-terminal domain and facilities additional intermolecular interactions ultimately resulting in signal initiation.

Effector recognition may occur by direct binding like in the case of L6 and mainly involve the LRR domain or require co-factors that may be guardees, decoys or other host proteins (also defined as molecular baits of effectors) for effector recognition and activation. This latter situation has been further developed in the **bait and switch model**. This model state that co-factors or baits prime NB-LRR proteins to be functional and retain, in addition, the molecular switch in an inactive conformation until the effector-induced modification of the bait triggers release of autoinhibition. This bait modification can have multiple forms, such as post-translational modification, cleavage, degradation or simply complex formation with the bait protein. However, at least some baits, such as NRIP1, interact with their cognate NB-LRR proteins only in the presence of the effector (Figure 15). In these cases, priming of the NLR has to occur independently of the bait and it is the binding of the effector-bait complex to the NLR that triggers the switch. This binding may involve as in the case of N multiple interactions between the effector-bait complex and different domains of the NLR.

<u>"Equilibrium-based switch" model</u>

Recent investigation of the mode of activation of L6 led to the development of the alternative **equilibrium-based switch** model. It's baseline is that even in the resting state, in the absence of effector, a small but significant part of the NLRs are in the ATP-loaded "ON" state and that

effectors bind to these NLRs and not to the ADP-bound NLRs that are "OFF" as in the simple switch model (Williams, et al. 2011; Bernoux et al. 2016). In the resting state, most NLRs are OFF and this equilibrium is maintained by rapid hydrolysis of ATP and slow release of ADP. Upon binding of the effector to the ON state, ATP hydrolysis is slowed down preventing the receptor to switch towards an inactive configuration and stabilizing the ATP-loaded form. By this, the equilibrium is shifted efficiently to the side of the ON state in the presence of effector and defense responses are triggered (Figure 15).

This model is supported by two main findings. On the one, hand the effector AvrL567 binds preferentially to the ATP-loaded ON, state and not to the ADP-bound OFF/ state of the receptor protein L. On the other hand it was shown that auto-active mutants of the L6 and L7 proteins preferentially bind ATP molecules (Bernoux et al., 2016).

The role of the intramolecular interactions in the NLR activation

The activation of NLR resistance proteins not only relies in the direct or indirect interaction with their cognate effectors but also depends on the strength of the intramolecular, inhibitory interactions between the different NLR domains. Indeed, modifications of NLR sequences can result in expanded effector recognition or even effector-independent activation suggesting that the modification of the intramolecular interactions alters the control of the active and inactive states of NLRs. More recently it was demonstrated that the specific modification of residues involved in negative regulatory intramolecular interactions, e.g. in the NB-ARC or the TIR, can lead to NLR variants that are efficiently activated by Avr effector alleles that not or only weakly activate the non-modified NLRs (Giannakopoulou et al., 2015; Segretin et al., 2014; Harris et al., 2013). Such 'trigger-happy' NLR variants can thus have extended recognition spectra and illustrate that receptor activation results from an interplay between the sensitivity in sensing the effector and the ease with which inhibitory intramolecular interactions are released.

This is nicely illustrated by the role of the L-TIR domain in effector recognition specificity. Usually, the specificity of the recognition of different AvrL567 variants by different alleles of L is explained by the polymorphic residues concentrated in the LRR domain. However, the alleles L6 and L7 differently recognize AvrL567 variants even though they possess identical LRR domains and only differ in 10 polymorphic residues confined to the TIR domain (Bernoux et al. 2016). AvrL567 variants expressed in flax plants carrying L7 or L6 trigger a weak or strong cell death response respectively and AvrL567-A showed a weak interaction with the full length of L7 protein compared to L6. Swaps between L6 and L7 showed that two polymorphic regions in TIR and in NB domain are involved in the regulation of cell death and demonstrated that negative functional interaction between the TIR and NB domains favors the inactive ADP-bound state impairing the receptor to switch to an active ATP-bound state (Bernoux et al. 2016).

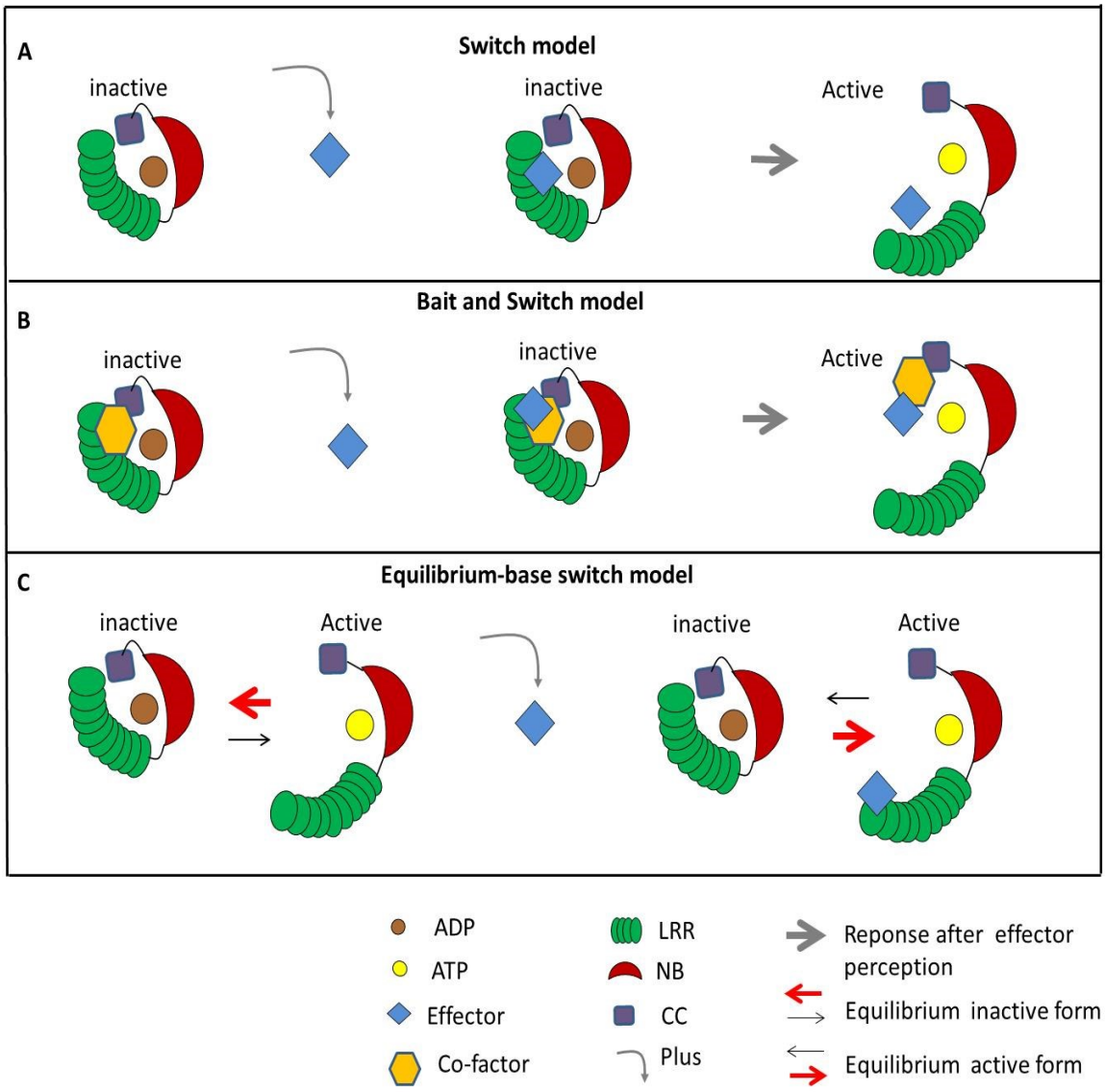


Figure 15. Models of resistance protein activation. The activation of resistance proteins after effector perception has been illustrated by three different models. **(A)** Switch model, the effector directly binds to an inactive protein inducting a conformational change orquestrated by the release of an ADP molecule and binding of ATP resulting in the activation of the protein. **(B)** Bait and switch model, the effector binds an inactive protein through a co-factor and as in (A) active the resistance protein. **(C)** The equilibrium-base switch model state that resistance proteins exist in equilibrium between an active and inactive configuration. Therefore, in this model the effector binds directly to the active protein, stabilizes its configuration and shifts the equilibrium towards activation. Adapted from Sukarta et al. 2016.

3.7 NLRs and their signaling components

NLRs are modular proteins mediating effector recognition and downstream signaling to induce defense responses. Transient expression of a truncated form of L10 encoding the N-terminal TIR homology region of the L10 protein led to cell death induction and suggested for the first time a role of the N-terminal part of the plant NLRs in signaling induction (Frost et al., 2004). Further analysis showed similar results e.g. transient expression of different truncations of RPP1A in tobacco revealed that its TIR-NB-ARC portion was sufficient to induce an elicitor-independent cell death (Weaver et al., 2006). Similarly, transient expression of the TIR_{1-45} and TIR_{1-80} fragments of RPS4 induced cell death in tobacco leaves. Furthermore, the cell death mediated by TIR_{1-80} required EDS1, SGT1 and HSP90 showing that this fragment of the protein seems have the same genetic requirements than the full length protein (Swiderski et al. 2009). Over-expression of the CC of MLA10 also resulted in cell death induction in barley (Maekawa et al. 2011). However, not always the transient expression of CC or TIR domains leads to cell death induction. For example, the transient expression of the TIR_{1-80} fragment of RPP2A or RPP2B did not induce cell death (Swiderski et al., 2009). Likewise, in the resistance proteins Rx, the over-expression of the NB domain and not the CC domain was sufficient to induce cell death. This lack of uniformity suggests that the NLR domains and downstream components that mediate signaling vary from one NLR to another. In addition, specific intra-and intermolecular interactions may be required to activate NLR signaling platforms (Eitas & Dangl 2010; Sukarta et al. 2016).

Role of complexes in NLR signaling

Homotypic or heterotypic interactions between the N-terminal domains of NLRs seem crucial for their function. This has in particular been demonstrated for MLA10 and L6 where the formation of homotypic complexes of their TIR or CC domains is crucial for downstream signaling. MLA10 and L6 mutants with single amino acid substitutions in the CC or TIR interaction interfaces are impaired in homodimer formation and fail to induce cell death (Bernoux et al. 2011; Maekawa et al. 2011).

The role of heterotypic TIR domain interactions has been investigated in the case of the RRS1/RPS4 pair. RRS1 and RPS4 TIR domains form homo-and heterodimers and structure -

function analysis revealed the interaction surfaces (Williams et al., 2014). The RPS4-TIR domain induces like the L6 TIR domain effector-independent cell death and for this homo-dimer formation is crucial since mutants that fail to dimerize are impaired in cell death induction. The RRS1 TIR domain, on the contrary, does not cause cell death but represses cell death caused by the RPS4 TIR domain. For this repression, formation of RRS1-RPS4 TIR domain hetero-complexes are required since RRS1 TIR mutants that fail in hetero-complex formation do not repress the cell death induced by the RPS4 TIR domain (Williams et al., 2014).

The tobacco NLR protein N oligomerizes in the presence of the effector protein p50. This oligomerization not only involves the TIR domain but also requires an intact P-loop suggesting that the NB-ARC domain may be involved in mediating NLR homo-complexes formation (Mestre and Baulcombe, 2006). In animals NLR activation leads to the creation of supramolecular structures that act as signaling platforms that enable the recruitment of additional adaptors and signaling components (Hu et al., 2015). Similar plant NLR super-structures mediating cell signaling have not yet been identified.

Role of accessory proteins in NLR signaling

Recruitment of executing signaling proteins by NLRs is thought to be a key step for the activation of the resistance responses. In plants, several signaling proteins have been shown to associate with NLRs and contribute to the activation of NLR-triggered defense. However, direct interaction between NLR receptors and signaling partners has been poorly demonstrated and little is known about the mechanism NLRs use to activate downstream signaling components. A recent study showed that the CC domains of several NLRs with homologies to the tomato NLR I2 induce cell death and interact with the chloroplastic protein Tylakoid Formation 1 (THF1) (Hamel et al., 2016). In addition, they have a negative effect on the accumulation of THF1. Further functional analysis with the NLR protein N indicated that THF1 functions as a negative regulator of the cell death and that the activation of N by P50 results in the destabilization of THF1. THF1 is crucial in chloroplast homeostasis and the cell death induced by N was light-dependent indicating a link between light, chloroplasts and the cell death

induced by I2-like NLRs (Hamel et al., 2016). Other signaling components in NLR-triggered immunity are NDR1, EDS1 and PAD4 (Aarts et al., 1998; Parker et al., 2000). NDR1 is a hub to mediate CNL defense responses whereas EDS1 and PAD4 are required for TNL-mediating signaling. However, the mechanistic link between NLRs and these signaling components and their molecular activity are still unknown.

Transcriptional regulators have also been shown to play a role in NLRs-mediate signaling. For example, in tobacco, association between N and the transcription factor SLP6 (squamosal promoter binding protein (SBP)-domain) was shown to be required for N-mediate signaling and defense responses (Padmanabhan et al., 2013). Similarly, after effector perception and activation, the NLR protein MLA10 translocates into the nucleus where it interacts with WRKY transcription factors that act as repressors of resistance and with the transcription factor MYB6 that acts as an activator (Shen et al., 2007; Chang et al., 2013). In rice, the CNL protein Pb1 was shown to interact with the transcription factor WRKY45 in the nucleus. Disruption of the Pb1-WRKY45 interaction mediated by the CC domain of Pb1 compromised Pb1-mediated resistance (Inoue et al., 2013a). Association of NLR domain with transcriptional repressors and activators may be a widespread strategy to rapidly modify gene expression for the activation of resistance responses.

Some NLRs, rely for their function on so-called helper NLRs (Eitas and Dangl, 2010b). Therefore, helper NLRs are thought to complement the function of NLRs that do not possess functional signaling domains and that probably act only as sensor proteins. Helper NLRs have been found to be not only involved in NLR signaling but also to be associated with RLK and RLP signaling. Thus, this group of proteins probably acts as a molecular bridge connecting these two different layers of plant immunity. Among the best characterized examples of helper NLRs are ADR1 that co-operates with several TNLs in *A. thaliana* and NRG1 which acts downstream of the TNL N (Bonardi et al., 2011; Collier et al., 2011). Unlike classical NLRs receptors, NGR1 and ADR1 family members are conserved among different plant species and do not require an entire P-loop to induce cell death (Bonardi et al., 2011).

3.8 Compartmentalization of NLR proteins

NLRs can act in different cellular compartments. The immune receptors RPS2, RPS5 and RPM1 possess N-terminal motifs that determine their localization at the **plasma membrane** (Figure 16). This membrane localization correlated to their activity as guards of proteins located in the plasma membrane where they are targeted by their corresponding effector proteins (Qi et al., 2012; Gao et al., 2011). The flax rust resistance proteins L6 and M carry N-terminal signal sequences that address them to different **endomembranes** (Figure 16). L6 to the Golgi and M to the tonoplast (Takemoto et al., 2012). Swapping of the N-terminal sequences changes the localization of the proteins but does not interfere with their function. However, complete deletion of the N-terminal signal sequence of L6 affects its stability and function (Takemoto et al., 2012).

Other plant NLRs move between the cytoplasm and other cellular structures (Figure 16). For instance, prior to pathogen attack, the potato NLR protein R3a is in the cytoplasm and after perception on the effector AVR3a it relocates to the endosomes (Engelhardt et al., 2012). The inhibition of endocytotic trafficking impairs R3a function suggesting that its re-localization is required for R3a-mediated defense activation.

Several plant NLRs re-localize to the nucleus upon pathogen attack and effector recognition (Figure 16). This is the case of the resistance protein MLA10 which in the absence of pathogen locates in the cytoplasm but after recognition of the effector Avr10, it is addressed to the nucleus where it interacts with different transcription factors (Shen et al., 2007; Chang et al., 2013). Other NLRs such as Rx1, N, RPS4 and RRS1 have also been found to move from the cytoplasm to the nucleus upon effector recognition (Figure 16). **Nucleo-cytoplasmic** partioning of several NLRs has been found to be an important mechanism to differentially activate downstream signaling. For example, the NLR protein RPS4 confers along with RRS1 the recognition of the effector proteins AvrRps4 and PopP2. RPS4 accumulates in the endomembranes and the nucleus (Wirthmueller et al., 2007). The forced accumulation of AvrRps4 in the nucleus leads to RPS4-mediated resistance but impairs RPS4-mediated HR. However, when AvrRps4 is retained in the cytosol RPS4 is able to mediate both resistance and

HR (Heidrich et al., 2011, 2012). The NLR Rx of potato recognizes the effector protein CP (coat protein) from Potato Virus X. The sequestration of Rx in the nucleus resulted in an impaired function of the resistance protein and the forced accumulation of Rx in the cytosol enhanced its function (Slootweg et al. 2010). However, when CP is forced to accumulate in the nucleus, Rx is impaired to activate resistance responses showing that recognition of effector protein and Rx-mediated signaling occurs in the cytoplasm (Tameling et al. 2010). The role of Rx accumulation in the nucleus is still an open question.

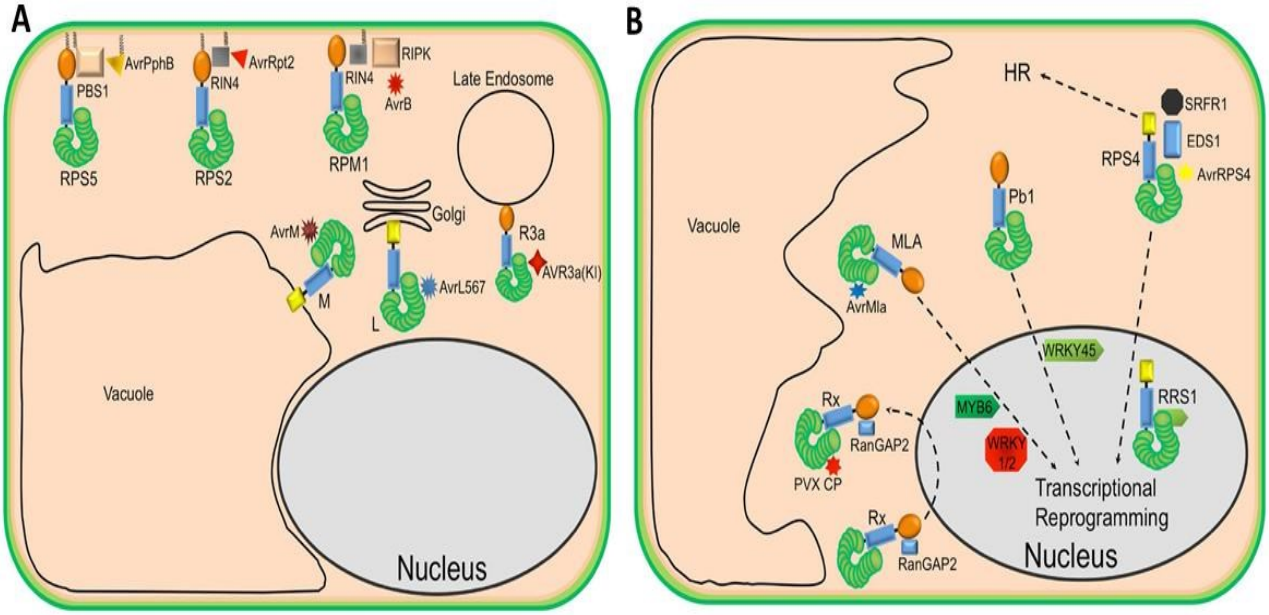


Figure 16. Localization of resistance proteins in plant cell compartments. NLR proteins can localize in the plasma membrane and endo-membranes **(A)** or accumulate in the cytoplasm and/or the nucleus **(B)** Adapted from Qi & Innes 2013

4. Rice-*M.oryzae* pathosystem: a model of host-fungal pathogen interactions

4.1 Rice: an important crop in food security

Rice (*Oryza sativa*) is a diploid, annual grass of the genus Poaceae (grass family) that self-pollinates resulting in a limited degree of outcrossing (< 0.5%). O*ryza sativa* possess two different subspecies *indica* and *japonica* both domesticated about 9.000-10.000 ago in India (Ganges plain) and China (middle and lower Yangtze valley) (Civá et al., 2015). Today, rice is cultivated in wet tropical, semi-tropical, and warm temperate areas around the world for the production of its cereal grain. There is high genetic diversity in rice germplasm and thousands of cultivars exist that have different grain color, size, and shape, as well as environmental tolerances and seasonality. In most of the cases rice is grown in fields that are flooded by either irrigation, rain-fed or floodplain systems. Moreover some varieties can be cultivated without flooding but they account for only 4% of rice cultivated worldwide (Garris et al., 2005).

Rice together with wheat is one of the most important cereal crops for human consumption in the world. Its production in 2012 was about 697.225 million tons in 158 million hectares that means 4.40 tons/hectare (http://ricestat.irri.org:8080/wrsv3/entrypoint.htm). To reach this tremendous production, the use of improved high yielding varieties, adequate irrigation, use of fertilizers and other complementary inputs are necessary. Despite the significant progress in enhancing rice productivity since the green revolution, population growth has outpaced increase in rice production (Hossain, 2007). Estimations state that crop yields must be increased by 150% before 2030 to satisfy the global food demand, nevertheless rice yield production has already stuck (Rao et al., 2014). Rice production is frequently threat for pathogen diseases. Among them blast disease caused by the fungus *M. oryzae* represents one of the most serious and widespread constraints.

4.2 Rice blast disease: a major threat for rice production

Rice blast, caused by *Magnaporthe oryzae,* is the most destructive disease of rice worldwide and represents a major economic concern and a serious threat for food security (Pennisi 2010;

Dean et al. 2012). Globally, annual yield losses provoked by this disease could feed 60 million of people (Pennisi, 2010).

M. oryzae attacks all aerial parts of the plants and causes characteristic necrotic lesions on leaves, stems and panicles (Figure 17) (Faivre-Rampant et al., 2008). Management of rice blast disease is based on a combination of blast resistant cultivars, fungicides and agricultural practice such as limitation of nitrogen fertilization (Wang & Valent, 2009). The selection pressure by continuous use of fungicides results in the emergence of fungicide-resistant populations, limiting its efficiency. In addition, fungicides used increase production costs and raise increasing concern due to their impact on the environment and human health. The use of disease resistant rice varieties represents the economically and ecologically most favorable solution for blast disease control. However, frequently, the resistance of new cultivars breaks down within a couple of years after release due to the rapid evolution of *M. oryzae* populations (Tharreau et al., 2009)

So far, the bases of durable disease resistance remain largely elusive and the elucidation of the mechanisms underlying fungal pathogenicity and plant resistance is believed to be important element to achieve of this goal. Because of the genetic and molecular tools available for rice and *M. oryzae* (molecular markers, genetic maps, whole genome sequence, and transformation facilities) as well as the agronomic relevance of rice for world population, the rice–*M. oryzae* pathosystem is one of the most suitable models to study plant–fungus interactions.

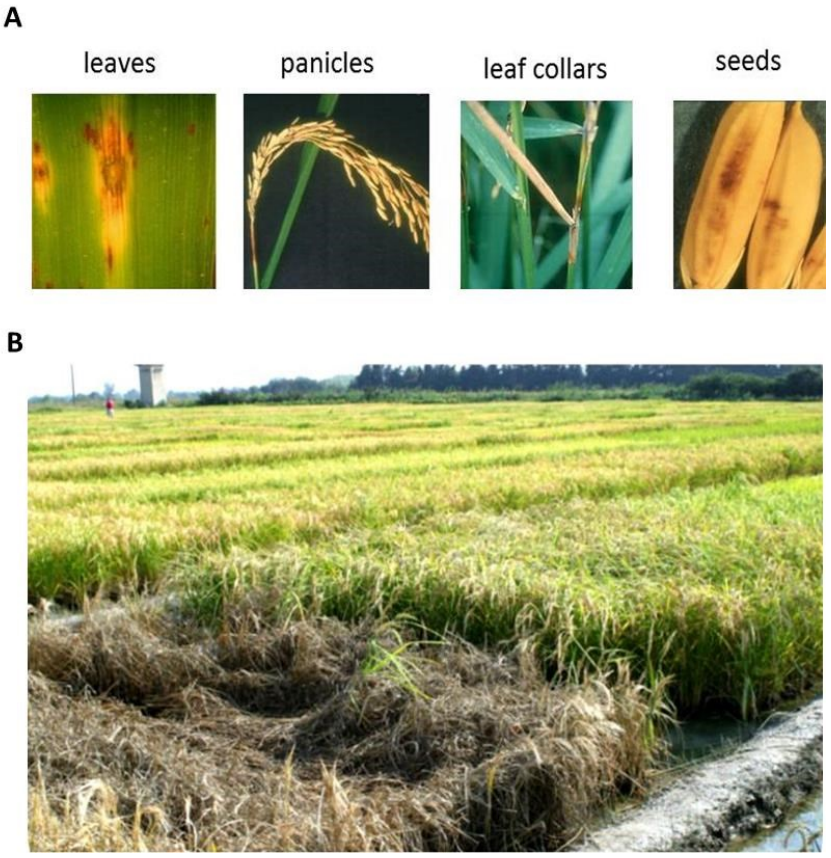


Figure 17. Symptoms of rice blast disease caused by *M. oryzae*. Symptoms on aerial parts of rice plants infected by *M. oryzae* **(A)**. Complete destruction by blast of a susceptible rice variety in an experimental field **(B)**. Photo from Didier Therrau

4.3 The fungus: *Magnaporthe oryzae*

Magnaporthe oryzae causes blast disease on a broad range of cereal and grass plants including food security crops such as rice, wheat, barley and millet (Couch et al., 2005). So, while the *M. oryzae* species complex can cause disease on almost hundred different plant species, individual isolates have very narrow host ranges and can infect generally only one single plant species. Rice isolates e.g. only infect rice whereas wheat isolates are only pathogenic on wheat. Isolates that share the same host plant are genetically more related than isolates with other host specificities (Couch et al., 2005; Chiapello et al., 2015).

Recently, the comparison of genome sequences and morphological structures of 6 isolates of *Magnaporthe* species from three different host plants, showed that the host specificity and adaptation of different pathotypes of *Magnaporthe* were not influenced by inherent morphological structures, but rather that host specificity and adaptation are evolutionary traits acquired by *Magnaporthe* species under strong selection pressures predetermined their host plants (Zhong et al., 2016). Another comparative genomic study of host-specific *M. oryzae* isolates showed that gain and loss of genes is a major mechanism in host specialization (Yoshida et al., 2016). Transposable elements seem to facilitate this process of gene evolution potentially by enhancing the rearrangement of chromosomes and other forms of genetic variation (Yoshida et al., 2016).

The rice blast fungus *M.oryzae* is currently found in the literature under two different names. *Pyricularia oryzae,* the anamorph (asexual) form of the fungus is the prevalent form in the field and is associated with the infection. However, the nomenclature of *M. oryzae,* the teleomorph (sexual) form, has been widely adopted by the rice blast scientific community during the last years. Hence, in this study we will refer to the recognizable name *M. oryzae*.

M. oryzae is an ascomycete haploid, heterotallic and hemibiotrophic fungus. After germination of conidia on leaf surfaces, germinating hyphae differentiate into a specialized cell called appressoria that produces tremendously high turgor pressure to mechanically break the leaf surface (cuticule and cell walls) to invade biotrophically one underlying epidermal cell. Thin primary invasive hyphae characterized by hyphal tip growth change rapidly their growth pattern and become more bulbous and multiply in a budding-like manner until the first invaded cell is completely filled with fungal hyphae. Subsequently, the fungus will invade adjacent cells in a similar biotrophic manner probably by exploiting pit fields for cell to cell passage. During this initial biotrophic phase, *M. oryzae* actively suppress the host immune system and cell death allowing the invasive hyphae to colonize living plant tissue from where nutrients are taken up. This is followed by a necrotrophic phase that is initiated by the

induction of host cell death by the fungus, accompanied by the growth of thin extracellular hyphae in the dying plant tissue and terminates in asexual spore production (Figure 18).

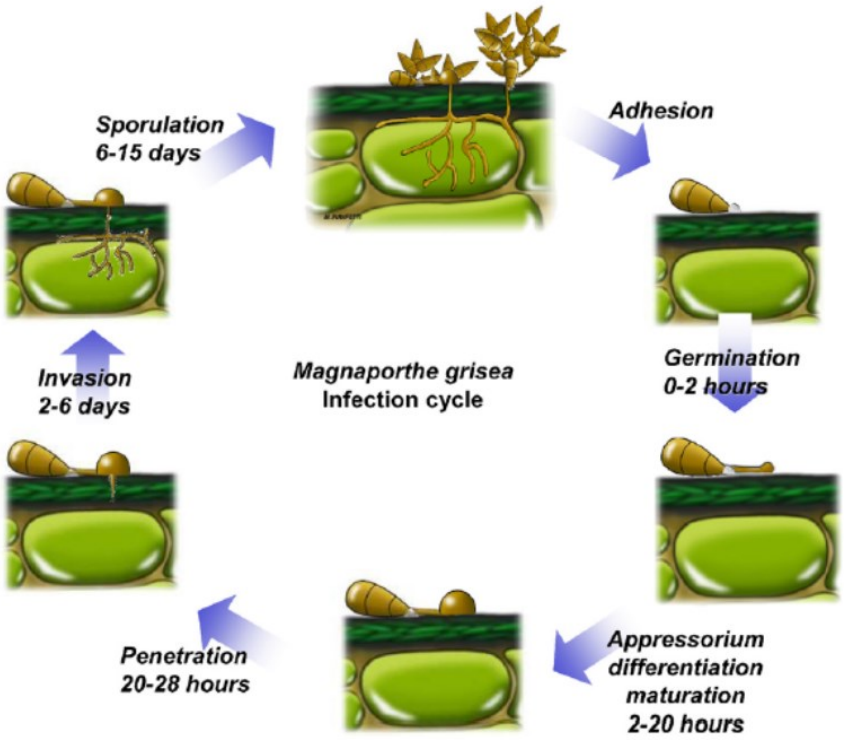


Figure 18. *Magnaporthe oryzae* infection cycle.
The *M. oryzae* infection cycle begins when a conidia from the fungus lands on rice leaves. Conidia attach firmly and germinate after few hours. Germ tube stops polar growth and forms a highly melanized dome-shape infection structure named appressorium. During the next hours, turgor pressure increases in the appressorium resulting in cell penetration via the penetration pore from which a narrow primary invasive hyphae emerges. After penetration, biotrophic growth starts and the primary hypha differentiate into a series of bulbous hypha those continuously divide in a budding type manner to completely fill the first invaded cell. Before expand the infection to other cells. After 4-5 days the fungus initiates necrotrophic growth that results in host tissue killing, development of disease symptoms and finally, clonal reproduction and sporulation. Adapted from Ribot et al. 2008

4.4 Effector proteins in *M. oryzae*; key determinants of virulence

Effector proteins are key elements in *M. oryzae* virulence and are particularly important during the biotrophic phase of fungus infection (Valent & Khang 2010; Giraldo & Valent 2013; Koeck et al. 2011). *M. oryzae* has a genome of approximately 38 Mb coding about 11.000 genes (Dean et al., 2005). Depending on the annotation pipeline, 750 to 1570 putative secreted proteins were identified in the genome of the rice blast fungus (Dean et al. 2005; Yoshida et al. 2009; Chiapello et al. 2015; Sperschneider et al. 2015). More than 750 of them have no functional annotation and approximately 40% are smaller than 250 amino acids resulting in huge sets of potential *M. oryzae* candidate effectors even if more refined analysis are performed. For instance, a recent genome-wide effector search relying on multiple parameters and a machine learning approach identified 485 effector candidates (Sperschneider et al., 2015). An important additional criteria to further refine the *M. oryzae* candidate effector lists is *in planta* expression. For example, by comparing the expression of *M. oryzae* genes in early-infected rice leaf sheaths with their expression in *in vitro* grown mycelium, 58 candidate effectors specifically expressed during infection and named BAS (biotrophy-associated secreted) proteins were identified. However, the functional analyses of hundreds of candidate effectors remains a tremendous challenge and so far, not more than 35 *M. oryzae* effectors and effector candidates, including 10 Avrs, have been further characterized.

The crucial role of small secreted proteins in the *M. oryzae* effector repertoire is underscored by the finding that all cloned *M. oryzae Avr* genes (Table 1) with the exception of ACE1, a cytosolic polyketide synthase/peptide synthetase and AVR-Pita a secreted zinc protease, code for small secreted proteins. Analysis of their localization during infection revealed that all these AVR effectors are translocated into host cells indicating that they act as cytoplasmic effectors (Giraldo and Valent, 2013). For six of them as well as for AVR-Pita a corresponding intracellular NLR rice immune receptor has been identified (Table 1).

Characterization of putative effectors in *M. oryzae*

Four BAS (Biotrophy-Associated Secreted) proteins, BAS1-4, that are specifically and highly expressed during early biotrophic infection were confirmed to be secreted in compatible but

not incompatible interactions (Mosquera et al., 2009). Furthermore, **BAS1** and **BAS2** were shown to be translocated into host cells and to accumulate in BICs (biotrophic interfacial complexes) a structure characteristic for the biotrophic phase (more details on the BIC further in this chapter). **BAS3** and BAS4 are extracellular effector and while BAS3 specifically co-localizes with cell wall crossing points, **BAS4** uniformly outlines invasive hyphae (Mosquera et al., 2009). However, disruption of these four *BAS* genes had no measurable impact on the virulence of *M. oryzae*.

851 putative effectors were identified in transcriptome profiles of rice blast infected leaves by using an integrative genome expression profiling approach that combines robust-long serial analysis of gene expression (RL-SAGE) with massively parallel signature sequencing (MPSS) and sequencing by synthesis (SBS) (Chen et al., 2013). 42 of these predicted effector proteins were selected and transiently expressed in rice protoplast to evaluate their ability to induce plant cell death. **Five** of the 42 effectors called **MoCDIP1-5** (*M. oryzae* cell death–inducing proteins) induced cell death in rice protoplast when they contained the signal peptide for secretion to the extracellular space (Chen et al., 2013). In addition, four of them also induced cell death in tobacco plants, suggesting that these effectors from *M. oryzae* may be key elements for facilitating colonization of the fungus during infection.

Disruption of 78 *M. oryzae* genes coding for putative secreted proteins was carried out to identify effector proteins mediating fungal pathogenicity during the early stages of infection (Saitoh et al., 2012). With the exception of **one** putative effector protein named **MC69**, the knock-out of these putative effectors did not affect growth, conidiation or pathogenicity of *M. oryzae* probably because they have redundant functions or no functions in these conditions. The *mc69* mutants, on the contrary, showed significant reduction of blast disease symptoms on rice suggesting that it has a crucial role in *M. oryzae* pathogenicity. Further, *M. oryzae* infection analyses showed that MC69 is dispensable for appressorium formation but is required for invasive hyphae development in rice leaf sheath. This suggests that MC69 is important in the interaction of the blast fungus with living rice cells during early infection. Live-

cell imaging indicates that MC69 is not translocated into the rice cytoplasm. Targeted mutagenesis of two cysteine residues (Cys36 and Cys46) in MC69 did not affect secretion but blocked its function, suggesting that disulfide bonds may be required for proper folding, stability or the molecular function of the MC69 in pathogenicity. Knock-out of an orthologue of *MC69* in *Colletotrichum orbiculare* responsible of cucumber anthracnose disease, also reduced fungal pathogenicity showing that MC69 is a secreted protein required for infection in at least two different plant pathogens (Saitoh et al., 2012).

A set of 247 candidate *M. oryzae* effectors was generated by a combination of expression studies, *in silico* identification of secreted proteins, a size limit of 250 amino acids and an elevated cysteine content (Sharpee et al. 2016). From these 247 putative effectors, 73 were successfully cloned and transiently expressed in tobacco plants to identify suppressors of BAX1 or Nep1-induced cell death. 11 of them named **SPDs** (suppressors of plant cell death) blocked cell death induced by the Nep1 and with exception also cell death triggered by BAX. Five of the eleven SPDs have been previously characterized, and are either essential for disease development, secreted during biotrophic infection, or cell death suppressors. Sequence analysis of the 11 SPD genes from 43 re-sequenced *M. oryzae* genomes showed that SPD2, SPD4, and SPD7 are particularly polymorphic when compared with other *M. oryzae* effector proteins suggesting that these effectors are under diversifying selection pressure to escape recognition from the host and have important roles in blast infection.

Mode of action of *M. oryzae* effector proteins

The molecular function of one apoplastic effector and two cytoplasmic avirulence effector proteins from *M. oryzae* has been recently characterized in more detail.

The effector protein **Slp1** (Secreted LysM Protein1) is a secreted LysM protein crucial for *M. oryzae* pathogenicity. Live-cell imaging of infected rice tissue showed that Slp1 accumulates at the plant-fungal interface during the early stages of infection where it specifically binds chitin oligomers with high affinity. Competition assays showed that Slp1 actually out-compete the

rice immune receptor CEBiP for chitin binding and suppresses by this chitin-induced plant immune responses (Mentlak et al. 2012). Deletion of *Slp1* considerably reduced the ability of *M. oryzae* to proliferate in rice tissue and cause disease suggesting that the suppression of chitin-triggered immunity by this apoplastic effector plays crucial roles in *M. oryzae* virulence.

The avirulence effector **AvrPiz-t** is a cytoplasmic effector recognized by the intracellular NLR immune receptor Piz-t (Li et al., 2009). In the absence of Piz-t, AvrPiz-t contributes to *M. oryzae* virulence since transgenic rice plants expressing AvrPiz-t showed enhanced susceptibility to the blast fungus (Park et al., 2012). This was correlated with suppressed immune responses in response to flg22 and chitin. Furthermore, AvrPiz-t interacts with the rice RING E3 ubiquitin ligase APIP6 and induces APIP6 degradation. APIP6 ubiquitinates AvrPiz-t and promotes its degradation indicating that both proteins stimulate the degradation of each other when co-expressed together (Park et al., 2012). *APIP6*-silencing in rice plants reduced PTI response and basal resistance to *M. oryzae* suggesting that APIP6 positively regulates rice innate immunity. However, APIP6 is not required for Piz-t-mediated resistance (Park et al., 2012).

In a separate study, AvrPiz-t was shown to target APIP10, another rice E3 ligase different from APIP6. AvrPiz-t suppressed the ubiquitin ligase activity of APIP10 which, in turn, ubiquitinates AvrPiz-t *in vitro* (Park et al. 2016). In the absence of *Piz-t*, silencing of *APIP10* compromised basal defense against *M. oryzae.* However, in a *Piz-t* background, silencing of APIP10 triggered Piz-t accumulation, cell death and enhanced resistance to *M. oryzae*. This seems due to promotion of Piz-t degradation via the 26S proteasome by APIP10. Therefore, it seems that APIP10 acts as a negative regulator of Piz-t and functionally connects AvrPiz-t and its cognate NLR Piz-t. Collectively, these 2 studies indicate that AVR-Piz-t targets the host ubiquitin proteasome system to suppress PTI (Park et al., 2012, 2016).

The effector **AVR-Pii** from *M. oryzae* is recognized by the paired NLRs Pii-1 and Pii-2 coded by the resistance locus *Pii.* AVR-Pii has been shown to form complexes with two rice Exo70

proteins, OsExo70-F2 and OsExo70-F3 (Fujisaki et al., 2015). Exo70 proteins are components of the exocyst complex mediating tethering and fusion of vesicles and plasma membrane at sites of polarized exocytosis. However, in the absence of *Pii*, the over-expression of *AVR-Pii* or the knockdown of both *Exo70* genes did not affect virulence of *M. oryzae* (Fujisaki et al., 2015). Therefore, an important role of AVR-Pii in *M. oryzae* virulence could not be demonstrated. However, knock-down of, *OsExo70-F3* but not *OsExo70-2,* abolished defense responses and resistance triggered by AVR-Pii and dependent on Pii. This suggests that the interaction of AVR-Pii with OsExo70-F3 plays an important role in Pii-mediated immunity to blast and OsExo70-F3 is expected to act as a decoy or a helper in Pii-mediated AVR-Pii recognition.

4.4 Delivery of *M. oryzae* effector proteins

Live-cell imaging of chimeric effector proteins expressed by transgenic *M. oryzae* isolates and labelled with fluorescent proteins such as green fluorescent protein (GFP), monomeric red fluorescent protein (mRFP) or monomeric Cherry (mCherry) has been useful to monitor effector localization during infection (Giraldo et al., 2013). This assays showed that apoplastic effectors, such as Bas4, Bas113 and Slp1 are retained within the extra-invasive hyphal membrane (EIHM) compartment located between the fungus and the plant plasma membrane and outline the entire IH in a uniform manner. On the contrary, translocated cytoplasmic effectors such as Pwl2, AVR-Pita, Bas1 and Bas107 accumulate in a punctate structure in the biotrophic interface named the biotrophic interfacial complex (BIC). The BIC is a plant membrane-derived interfacial structure initially localized on the tip of primary invasive hyphae and later positioned subapically, beside the first bulbous IH cell (Giraldo et al., 2013). Translocated effectors are supposed to be delivered from the BIC inside host cells by unknown mechanisms. Some of these intracellular effectors subsequently move up to three or four cells away from the invading hyphae. The movement from cell to cell observed for some cytoplasmic effectors possibly occurs through plasmodesmata and likely contributes to preparing host cells for pathogen colonization and infection (Figure 19) (Khang et al., 2010; Giraldo et al., 2013).

Fungal effectors are usually secreted into the plant-fungus interface via the classical ER-Golgi secretory pathway (Mentlak et al. 2012). This involves guidance of nascent secreted proteins to the rough, ribosome-rich ER by their N-terminal hydrophobic signal peptide and subsequent co-translational translocation into the lumen of the ER. From there, secreted proteins move by vesicle-mediated transport through different compartments of the Golgi-apparatus to finally reach the plasma membrane where transport vesicles fuse with the plasma membrane to deliver the secreted proteins to the extracellular space. Fusion of transport vesicles with the plasma membrane is a highly controlled process and involves two key steps: tethering of the vesicle to the plasma membrane by the exocyst complex and membrane fusion mediated by SNARE (soluble N-ethylemaleimide sensitive factor attachment protein receptors) protein. The exocyst complex is an octameric protein complex (in yeast Sec3, Sec5, Sec6, Sec8, Sec10, Sec15, Exo70, and Exo84) composed of subunits initially associated to the vesicle (e.g. Sec15), the plasma membrane (e.g. Sec3 and EXO70) or cytoplasmic. It assembles when the vesicle reaches the plasma membrane. Subsequent membrane fusion is mediated by v-and t-SNAREs located respectively on the vesicle and in plasma membrane (Mentlak et al. 2012).

While apoplastic effectors secreted from invasive hyphae into the extracellular space are believed to follow this conventional secretory pathway, cytoplasmic effectors were recently described to follow a novel pathway of secretion that is independent of the Golgi. This pathway, that also relies on exocyst components and the t-SNARE Sso1 is mainly defined by insensitivity to Brefeldin A an inhibitor of vesicle flow from the ER to Golgi vesicles. Indeed, Brefeldin A inhibits secretion of apoplastic effectors but not the secretion of cytoplasmic effectors that therefore seem to follow another secretion pathway that does not involve the Golgi (Giraldo et al., 2013)

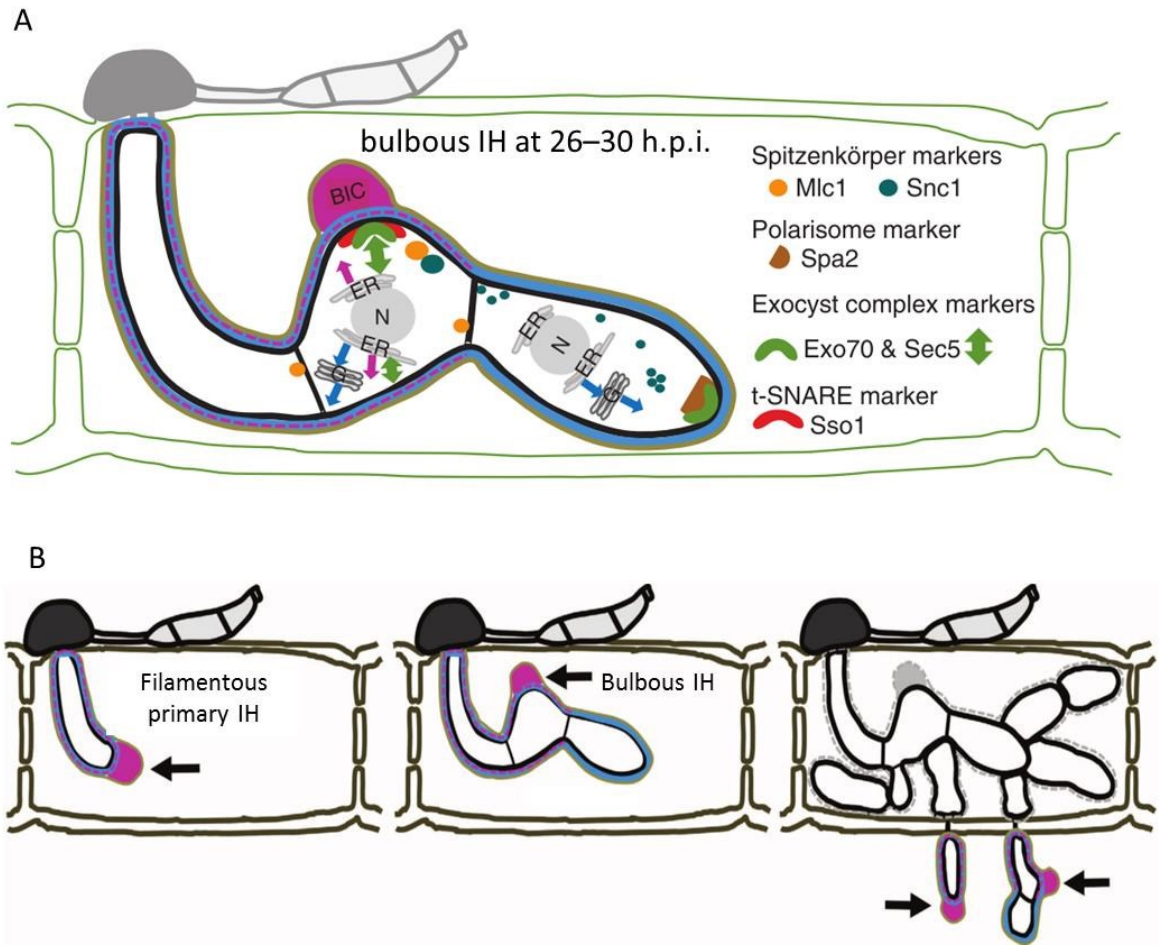


Figure 19. Secretion and accumulation of *M. oryzae* effector proteins. A) Apoplastic effectors (blue), accumulate in the EIHM compartment surrounding the IH, resulting in uniform outlining of the IH. These apoplastic effectors follow the conventional, BFA-sensitive, Golgi-dependent secretion pathway. In contrast, cytoplasmic effectors (magenta), accumulate in the BIC beside the first-differentiated bulbous IH cell. Cytoplasmic effectors follow a nonconventional, BFA-insensitive secretion pathway **B)** Schematic representation of the effector accumulation during the differentiation of a filamentous primary invasive hypha into a pseudohyphal-like bulbous invasive hypha in the first-invaded rice cells. Cytoplasmic effectors show preferential accumulation in the BIC (black arrows), which is first located in front of the growing primary hyphal tips, and then remains behind beside the first-differentiated bulbous IH cell. Typical accumulation patterns for cytoplasmic (magenta) and apoplastic (blue) effectors are shown within the EIHM (tan) compartment enclosing the IH. The EIHM appears to lose integrity when the fungus has moved into neighbor cells (dotted line). Adapted from Giraldo et al. 2013

4.5 Rice Immune receptors

Rice plants possess an arsenal of resistance genes that counteract *M. oryzae* attack. More than 100 blast resistance genes have been genetically identified and 25 genes from 11 loci have been cloned (Table 1). All rice blast immune receptors cloned so far code for NLRs with the exception of Pi-d2 that encodes a receptor-like kinase (Chen et al., 2006), suggesting that blast resistance mainly relies on the recognition of cytoplasmic effectors by NLR immune receptors.

Rice PRRs

Rice RLPs and RLKs contribute to basal resistance against the blast fungus, however with the exception of immune receptor Pi-d2 they appear not to confer complete resistance. Pi-d2 has a unique structure composed of a predicted extracellular bulb-type mannose-specific binding lectin (B-lectin) and an intra-cellular serine-threonine kinase domain. Pi-d2 has been shown to confer resistance to the Chinese rice blast strain ZB15 in transgenic rice plants (Chen et al., 2006, 2010). However, the ligands and the precise function of Pi-d2 remain elusive.

Among the best understood examples of rice PRR function is the perception of chitin, one of the main structural components of fungal cell walls, by the rice immune receptor complex CEBiP/CERK1 (Chitin oligosaccharide elicitor-binding protein/Chitin Elicitor Receptor Kinase). CEBiP is an RLP protein containing a transmembrane domain and three LysM motifs that bind chitin (more details about PRR receptors above in this chapter). Rice CEBiP mutants display an increased susceptibility to attack by the rice blast fungus as well as a reduced chitin-triggered immune response. Upon chitin-binding, CEBiP forms heterodimers with CERK1, a RLK protein with LysM motifs in the extracellular domain and an intracellular Ser/Thr kinase domain. Chitin-bound CEBiP/CERK1 heterodimers represent a signaling-active receptor complex that activates chitin-triggered immunity in rice (Shimizu et al., 2010).

Additional rice PRRs such as LYP4 and LYP6 binds both chitin and peptidoglycan. Knockdown of *CEBiP*, *OsCERK1*, *LYP4* or *LYP6,* all containing at least one LysM domain, resulted in a reduced chitin-triggered immune responses and compromised the resistance against *M. oryzae*

suggesting that LysM domain has an crucial role in chitin oligosaccharides perception in rice plants (Couto and Zipfel, 2016).

Other rice RLPs or RLKs have also been shown to be involved in blast resistance. For instance, the knockout of several rice wall-associated kinases (OsWAK), a sub-family of receptor-like kinases, compromised resistance responses to the rice blast fungus (Hu et al. 2014; Delteil et al. 2016; Harkenrider et al. 2016). Their mode of action and their ligands are however not known.

Rice NLRs

Genome-wide investigation of *NLRs* in two sequenced rice varieties, Nipponbare (*japonica*) and 93-11 (*indica*) identified 623 and 725 *NLRs* respectively (Luo et al., 2012). Among them, 347 (55.7%) in Nipponbare and 345 (47.6%) in 93-11 are pseudogenes caused either by large deletions, nonsense point mutation or small frame-shift insertion/deletions (indels). 75% of the *NLRs* are organized in clusters in the rice genome with two or more *NLRs* separated by no more than eight *non-NLRs*. In addition, *NLRs* are also found as single-copy genes. Thus, from 279 *NLR* loci identified in Nipponbare, 160 were single-copy loci and 119 were multiple-copy loci (Luo et al., 2012). Many *NLRs* conferring rice blast resistance are localized in *NLR* clusters. An important cluster for rice blast resistance is e.g. located on chromosome 6 an harbours *Pi2*, *Pi9, Pi50* and *Piz-t* (Jiang et al., 2012).

Rice *NLRs* can be classified in two main types based on their nucleotide diversity. **Type I *NLRs*** are rapidly evolving genes with extensive chimeric structures and high diversity among rice accessions (Luo et al., 2012). Frequently, orthologous relationships cannot be established between Type I *NLRs* from different varieties since there is copy number variation between different varieties and intense recombination between more or less closely related Type I *NLRs.* **Type II *NLRs*** are slowly evolving genes without sequence exchanges between different *NLRs* and highly conserved in different accessions (Luo et al., 2012). Therefore, for Type II *NLRs*, orthologous can be identified in different rice varieties. Type II *NLRs* are the

predominant class of NLRs in the genome of Nipponbare and almost all the *NLR R* genes cloned so far belong to this category. Although type I *NLRs* only account for a small proportion of all *NLRs,* e.g. 5.8% in Nipponbare, these *NLRs* significantly contribute to the diversity of *NLRs* in the rice genome. Furthermore, conserved rice *NLR*s are mostly single-copy and/or localize as a singleton whereas *NLRs* contributing to rice genetic variation are usually part of complex clusters (Yang et al., 2006). Presence and absence (P/A) polymorphisms are also frequent in rice *NLR* loci. For instance, 84 of the 323 *NLR* loci in 93-11 are absent from the Nipponbare genome. Similarly, about 40 *NLR* loci of Nipponbare are absent from 93-11. In total, at least 124 *NLR* loci exhibit P/A polymorphism between the two rice genomes suggesting that gain and loss of *NLRs* is an important mechanism in the evolution of rice immunity (Luo et al., 2012).

Bioinformatic searches of highly polymorphic *NLRs* associated with rapid molecular evolution enabled recently the massive identification of novel, functional blast resistance *NLRs* (Yang et al. 2013). 60 *NLRs* from three rapidly evolving *NLR* families were randomly chosen in maize, sorghum and brachypodium and 20 from 5 different rice varieties. After cloning and transformation into 2 different blast susceptible rice varieties, transgenic rice plants were inoculated with 12 different isolates of *M. oryzae* to test for sensitivity or resistance to the fungus (Yang et al., 2013b). By this, in total, 28 functional blast resistance *NLRs* were obtained, 15 from non-rice grasses and 13 from rice. This suggests that rapidly evolving *NLR* gene family are important for the resistance to fast evolving fungal pathogens such as *M. oryzae* and present an outstanding resource for resistance breeding (Yang et al. 2013).

Table 1. Cloned blast rice resistance genes and *M. oryzae* avirulence genes

	Resistance genes		Avirulence genes	
	R gene	**Encoded protein**	**Avr gene**	**Encoded protein**
***R-Avrs* pairs**	*Pib*	NB-LRR	*AvrPib*	75 aa secreted protein
	Pi-ta[a]	NB-LRR	*AvrPi-ta*	224 aa secreted protein
	Pi9	NB-LRR	*AvrPi9*	91 aa secreted protein
	Piz-t	NB-LRR	*AvrPiz-t*	108 aa secreted protein
	Pik	NB-LRR	*Avr-Pik/km/kp*	113 aa secreted protein
	Pikm	NB-LRR	*Avr-Pik/km/kp*	113 aa secreted protein
	Pikp[ab]	NB-LRR	*Avr-Pik/km/kp*	113 aa secreted protein
	Pia[abc]	NB-LRR	*AVR-Pia*	85 aa secreted protein
	Pi-CO39[abc]	NB-LRR	*AVR1-CO39*	89 aa secreted protein
***R* genes only**	Pi36	NB-LRR	--	--
	Pi37	NB-LRR	--	--
	Pi50	NB-LRR	--	--
	Pi64	NB-LRR	--	--
	Pit	NB-LRR	--	--
	Pi5[b]	NB-LRR	--	--
	Pid3	NB-LRR	--	--
	Pid3-A4	NB-LRR	--	--
	Pi54	NB-LRR	--	--
	Pi25	NB-LRR	--	--
	Pi1	NB-LRR	--	--
	Pb1	NB-LRR	--	--
	Pish	NB-LRR	--	--

	Pi2	NB-LRR	--	--
	pi21	Proline-containing protein	--	--
	Pi-d2	B lectin RKL		
***AVRs* only**			*ACE1*	Polyketide synthase
			AvrPii	70 aa secreted protein
			PWL2	145 aa secreted protein

Adapted from Liu & Wang 2016

a= These genes have been reported to have integrated domains

b= These genes requires two NLRs members to function

c= These genes recognized two unrelated effector proteins

4.6 Recognition of effector proteins by NLRs

Rice blast resistance can be mediated by either single *NLR*s such as *Piz-t*, *Pi9* and *Pib* or by functional *NLRs* pairs such as *Pik*-1/ *Pik*-2, RGA4/RGA5 and *Pi5*-1/ *Pi5*-2 (Liu, 2016). Effector recognition has been particularly well studied in the case of Pik-1/Pik-2 and RGA4/RGA5. *Pik* and the corresponding AVR effector gene, *AVR-Pik* are highly polymorphic. *Pik* has six different alleles (*Pik*, *Pikm*, *Pikp*, *Piks*, *Pikh* and *Pi1*) while AVR-Pik has five different alleles (*AVR-Pik A, B, C, D* and *E*) (Kanzaki et al., 2012). The interaction is allele-specific and individual NLR alleles recognize only particular subsets of the AVR-Pik alleles. AVR-PikD is e.g. recognized by Pikp, Pik and Pikm while AVR-PikA is only recognized by Pikm (Kanzaki et al., 2012). Detailed molecular and structural analysis demonstrated that AVR-Pik binds directly to the unconventional HMA domain of Pik-1 and that this physical interaction is required for recognition. The binding surface in AVR-Pik has been structurally characterized and it was shown that polymorphisms in this surface as well as in the Pik-1 HMA domain explain differences in AVR-Pik-HMA binding and immune activation (Maqbool et al. 2015). This work also showed that the HMA domain has been integrated into Pik-1 to mediate effector recognition.

The other well-studied example of NLR proteins functioning in pairs to recognize the blast fungus are RGA5 and RGA4. The corresponding genes show the characteristic tandem inverted

clustered organization of paired NLRs. RGA5/RGA4 mediate recognition to two different avirulence effectors AVR-Pia and AVR1-CO39 (Cesari et al., 2013). Functional analyses showed that RGA5 works as a sensor NLR protein to recognize effector proteins and as a repressor of RGA4, that acts as a cell death executor (Césari et al. 2014). Furthermore, RGA5 contains like Pik-1 an HMA/RATX1 domain which is also involved in effector binding (Cesari et al., 2013), confirming that integrated domains such as HMA/RATX1 play important roles in blast effector recognition.

The interaction of the blast resistance NLR Pi-ta and the *M. oryzae* effector AVR-Pita was the first report of direct interaction of fungal Avrs and NLRs (Jia et al 2000). This model was supported by the physical interaction between AVR-Pita with the Pi-ta-LRR which is disrupted by a single amino acid substitution in the Pi-ta-LRR domain. However, a second NLR named Ptr seems to be required for full Pi-ta function (Jia and Martin, 2008). Furthermore, an splice variant of *Pi-ta* encoding a TRX (thioredoxin) integrated domain in its C-terminal has been shown to present the highest level of expression while the originally described *Pi-ta* splice variant is of minor importance (Costanzo and Jia, 2009). Whether the TRX domain found in the Pi-ta variant is involved in AVR-Pita binding or whether Pi-ta cooperates with Ptr to mediate effector recognition remains to be determined.

The investigation of RGA4/RGA5 and Pik-1/Pik-2 indicates that the recognition of effector proteins by NLR pairs containing integrated domains and following the, “integrated decoy model” (Césari, et al. 2014), is an important mode to sense *M. oryzae* attack. The investigation of the Piz-t/AvrPiz-t and Pi-ta/AVR-Pita systems suggests also a role for indirect recognition following the guard or decoy model as well as direct ligand-receptor type interactions. Further investigation of these latter systems and additional NLR/blast effector couples such as Pib/AVR-Pib, Pi9/AvrPi9 or Pii/AVR-Pii are required to get a better understanding of the role and mode of action of NLRs in rice blast resistance.

4.7 Downstream responses in blast resistance

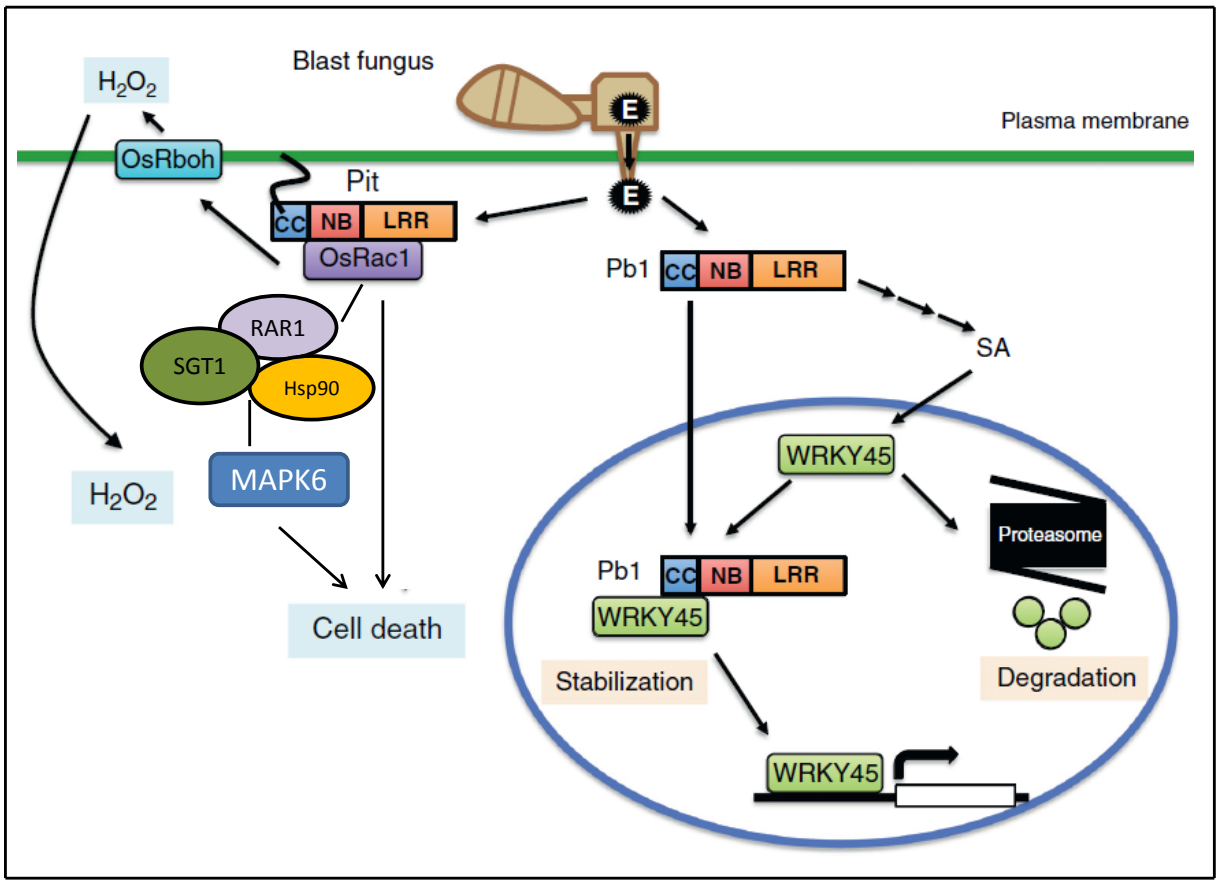
Relatively few rice proteins required for NLR-triggered immune signaling have been characterized so far. One example is OsRac1 that belongs to a family of small GTPases that function as molecular switches in different signaling cascades (Kawano et al., 2010). OsRac1 participates to PTI signaling and is also required for Pit, a blast NLR, to mediated resistance to *M. oryzae*. OsRac1 therefore seems to be a key regulator in rice immunity (Figure 20) (Kawano et al., 2010).

Most other proteins identified to signal downstream of NLRs are either protein kinases or transcription factors. Among them, MAPKs are particularly well established as important immune signaling components. For example, OsMAPK6 kinase activity is induced by a sphingolipid elicitor derived from *M. oryzae* (Lieberherr, 2005). Furthermore, OsMAPK6 interacts with OsRac1 to mediate resistance responses and is also required in signaling transduction triggered by Pit and the OsRac1, RAR1, HSP90 and SGT1 protein complex (Figure 20) (Kawano et al. 2010).

Many transcription factors of the WRKY family have been shown to be involved in rice blast resistance responses. Under *M. oryzae* attack the expression of 15 *OsWRKY* is increased. Overexpression of *OsWRKYs* such as *OsWRKY30* and *OsWRKY45* enhances resistance to the blast fungus. OsWRKY45 is regulated by the ubiquitin proteasome system (UPS). Interaction of OsWRK45 with Pb1, a blast resistance NLR, avoids OsWRKY45 degradation and is required for Pb1 to mediate resistance to *M. oryzae* (Figure 20) (Inoue et al., 2013b).

Different rice hormones such as salicylic acid (SA), jasmonic acid (JA), gibberellic acids (GAs), abscisic acid (ABA) cytokinine (CK) and brassinolide (BL) are important components in rice immunity. Rice plants treated with BL or JA show e.g. increased resistance to *M. oryzae* and induction of defense gene expression while ABA treatment enhances susceptibility (Yang et al. 2013). Recently, CKs were reported to be required for oxidative stress tolerance to the rice

blast fungus (Chanclud et al., 2016). However, the understanding of hormone function in rice immunity remains limited and requires further studies.



Adapted from Kawano & Shimamoto 2013

Figure 20. Signaling pathways in rice blast resistance. The rice GTPase OsRac1 forms a complex with the rice NB-LRR protein Pit. OsRac1 is required for Pit-mediated disease resistance to the rice blast fungus. The active form of Pit induces the activation of OsRac1 at the plasma membrane.OsRac1 functions in rice innate immunity through the RAR1-SGT1-HSP90 cytosolic complex. OsMAPK6 transduces the signaling mediated by Pit, OsRac1 and RAR1-SGT1-HSP90 complex. Pb1 interacts with a transcription factor WRKY45 through its CC domain and this interaction protects the degradation of WRKY45 by the ubiquitin-proteasome system. Pb1-mediated blast resistance is dependent on SA. Adapted from kawano et al. 2013

THESIS OUTLINE

General objective

The main objective of this thesis is to contribute to a better understanding of the molecular determinants of plant immunity. More specifically, we were interested in the elucidation of the molecular bases of the recognition of fungal effector proteins by plant NLR immune receptors and the link between effector recognition and immune activation. For this, we structurally characterized the *M. oryzae* effector proteins AVR1-CO39 and AVR-Pia and studied the diversity and evolution of structure-related effector proteins in ascomycete fungi. Furthermore, we investigate the molecular determinants mediating the recognition of AVR-Pia by the immune receptors RGA4/RGA5. Finally we analyzed the composition and formation of the RGA4/RGA5 receptor complex.

Since the *M. oryzae*-rice pathosystem is a model system for fungal-plant interactions, the knowledge generated in this work is highly relevant for a better comprehension of plant diseases caused by fungi.

Research questions and approaches

The following specific questions were addressed to reach the general objective of the PhD project

1. **Is the three-dimensional structure of effector proteins a key determinant of their function, diversity and evolution?**

To answer this question we first determined the three-dimensional structure of the *M. oryzae* effector proteins AVR1-CO39 and AVR-Pia using Nuclear magnetic resonance (NMR). Since both effectors had similar structures that they shared, in addition, with two other fungal effector proteins, AvrPiz-t and ToxB, from *M. oryzae* and *Pyrenophora tritici-repentis* respectively, we could define the novel MAX-effector family (***M****. oryzae* **A**vr and To**x**B effectors) by structure-informed sequence patterns. Bioinformatic searches exploiting this

pattern showed that MAX effectors are phylogenetically widely distributed, expanded and diversified specifically in the order of the Magnaporthales **(Chapter I)**.

2. How does AVR-Pia bind the RGA5$_{RATX1}$ domain?

Previously, our research group demonstrated that the *M.oryzae* effector AVR-Pia directly interacts with the C-terminal RATX1 domain of RGA5. In addition, it was shown that *AVR-Pia-H3,* a natural *AVR-Pia* allele, lost the interaction with the RATX1 domain and does not trigger blast resistance in rice plants carrying RGA4/RGA5.

In this study, we deepened the characterization of the AVR-Pia-RGA5$_{RATX1}$ interaction. We compared the NMR structures of AVR-Pia and AVR-Pia-H3 and identified a candidate interaction surface by NMR-titration. We characterized *in vitro* AVR-Pia-RGA5$_{RATX1}$ complex formation by isothermal calorimetry (ITC) and validated the candidate interaction surface as well as two key residues by yeast two hybrid and co-immunoprecipitation experiments **(Chapter II)**.

3. **What is the role of AVR-Pia - RGA5$_{RATX1}$ binding in AVR-Pia recognition?**

To determine the importance of AVR-Pia-RGA5$_{RATX1}$ binding for recognition and avirulence activity of AVR-Pia, structure-informed AVR-Pia mutants were analyzed in transgenic *M. oryzae* isolates and transient *N. benthamiana* cell death induction assays. These experiments showed that AVR-Pia-RGA5$_{RATX1}$ binding is required for resistance but that even weakly binding AVR-Pia mutants are still efficiently recognized **(Chapter II)**.

4. How are other RGA5 domains involved in AVR-Pia recognition?

To better understand RGA5-mediated recognition of AVR-Pia, interactions of AVR-Pia with individual RGA5 domains (CC, NB-ARC, LRR, RATX1) were characterized by yeast two hybrid and co-immunoprecipitation assays. This revealed interactions of AVR-Pia with the RGA5 NB-ARC and LRR domains in addition to RATX1 **(Chapter III)**.

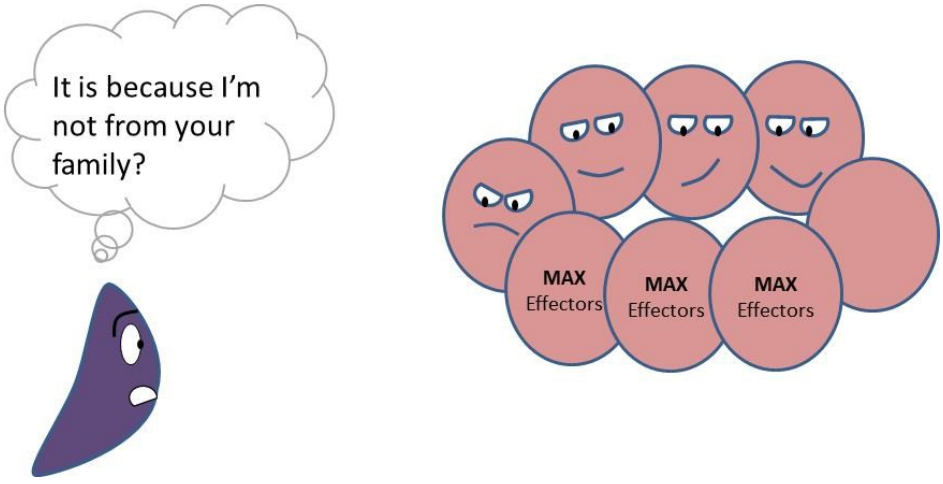
5. How does the RGA5/RGA4 receptor complex form and how does it mediate AVR-Pia recognition?

RGA4 and RGA5 functionally and physically interact to recognize AVR-Pia. CC domain homo and hetero-complex formation has been reported (Césari et al., 2014c). In addition, it was observed that RGA4 mediates cell death activation, while RGA5 acts as RGA4 repressor and AVR-Pia sensor.

In this study, to better characterize RGA5 and RGA4 interactions, we evaluated the physical interaction between these two NLRs in presence of AVR-Pia$_{WT}$. Next, we tested the hetero-complex formation of RGA4$_{\Delta CC}$ and RGA5$_{\Delta CC}$ and the association of individual domains (NB, CC, NB-LRR) of RGA4 and RGA5 by co-IP experiments **(Chapter IV).**

CHAPTER I

Structural analyses of *Magnaporthe oryzae* effector proteins



INTRODUCTION

Effectors are among the most powerful arms of plant pathogens to cause disease. Their study provides deep insight into the mechanisms that pathogens employ to provoke disease and into the co-evolution with their hosts. In addition, effectors are key elements for disease resistance breeding since their recognition by immune receptors is exploited for the development of resistant varieties. Recently, effector-informed breeding has been developed as a powerful tool to accelerate the identification, functional characterization and deployment of resistance genes in commercial crops. Effector-assisted breeding has been particularly successfully applied in late blight control (Whisson et al. 2011). Effector proteins from the oomycete *P. infestans* have been used to identify new resistance genes in potato germplasm, to assist resistance selection using them as "molecular markers" and finally to deploy resistance genes in potato commercial crops by monitoring the spacio-temporal distribution of effector proteins in pathogen populations (Vleeshouwers et al. 2008; Win et al. 2009; Boyd et al. 2013)

In recent years, structural biology has proven to be a particularly powerful approaches for the understanding of effectors (Wirthmueller et al. 2013). For example, the three-dimensional structure of an NLP from *P. aphanidermatum* was a key element to uncover the mechanism of NLP phytotoxicity (Ottmann et al., 2009; Küfner et al., 2009). The NLP structure revealed a fold formed by a central β-sandwich surrounded by α-helices and loops that exhibits structural similarities to soluble proteins produced by marine organisms named actinoporins (Ottmann et al., 2009). Actinoporins are cytosolic toxins that form trans-membrane pores via their N-terminal regions (Črnigoj Kristan et al., 2009). In NLPs the N-terminal region is also required to induce necrosis and plant defense activation (Fellbrich et al., 2002) suggesting that NLPs and actinoporins share a cytolytic, membrane-disintegrating mode of action.

Comparison of the structures of the four oomycete RXLR effector proteins AVR3a4, AVR3a11, PexRD2 and ATR1 that share less than 20% sequence identity; revealed a common three-α-helix fold, termed the 'WY domain' that appear to provide a flexible, stable scaffold that supports surface diversification of RXLR effectors (Win et al., 2012). Similarly, analyses of the structure of two alleles of the flux rust effector AvrL567 revealed a β-sandwich fold that has slight structural similarity to the host-selective toxin ToxA from the fungal pathogen *Pyrenophora tritici-repentis.* In addition, two patches of positive charge on the AvrL567 surface were identified as potential DNA-binding sites (Wang et al. 2007; Wirthmueller et al. 2013). The crystal structures of ATR1, AVR-Pik, AvrL567-A and AvrL567-D were also highly instrumental to map polymorphic residues involved in the interaction with their cognate NB-LRRs and, therefore, in mediating recognition (Wirthmueller et al. 2013).

Altogether, these examples illustrate how structural studies can lead to the elucidation of effector functions and mechanisms of effector action as well as the identification of interaction interfaces in effectors and hidden similarities between effector by sequences.

The *M. oryzae* effectors AVR-Pia and AVR1-CO39 are both recognized by the rice NLR pair RGA5/RGA4 through direct binding with the C-terminal domain of RGA5 containing the RATX1 domain. Interestingly, despite recognition by the same NLRs, AVR-Pia and AVR1-CO39 have very different amino acid sequences, suggesting molecular factors other than amino acid sequence are important determinants in effector recognition in this case.

In this chapter, we describe the structure as well as important biochemical surface properties of the *M. oryzae* AVR-Pia and AVR1-CO39. We also present the discovery of a family of sequence-unrelated effector proteins with a structurally conserved fold we that seems particularly abundant and important in *M. oryzae*.

Article I

PLOS | PATHOGENS

RESEARCH ARTICLE

Structure Analysis Uncovers a Highly Diverse but Structurally Conserved Effector Family in Phytopathogenic Fungi

Karine de Guillen[1,2☯], Diana Ortiz-Vallejo[3,4☯], Jérome Gracy[1,2], Elisabeth Fournier[3,4], Thomas Kroj[3,4‡*], André Padilla[1,2‡*]

1 INSERM U1054, Centre de Biochimie Structurale, Montpellier, France, **2** CNRS UMR5048, Montpellier University, Montpellier, France, **3** INRA, BGPI, Biology and Genetics of Plant-Pathogen Interactions, Campus International de Baillarguet, Montpellier, France, **4** CIRAD, BGPI, Biology and Genetics of Plant-Pathogen Interactions, Campus International de Baillarguet, Montpellier, France

☯ These authors contributed equally to this work.
‡ TK and AP also contributed equally to this work.
* thomas.kroj@supagro.inra.fr (TK); andre.padilla@cbs.cnrs.fr (AP)

Citation: de Guillen K, Ortiz-Vallejo D, Gracy J, Fournier E, Kroj T, Padilla A (2015) Structure Analysis Uncovers a Highly Diverse but Structurally Conserved Effector Family in Phytopathogenic Fungi. PLoS Pathog 11(10): e1005228. doi:10.1371/journal.ppat.1005228

Editor: Jin-Rong Xu, Purdue University, UNITED STATES

Received: June 18, 2015

Accepted: September 24, 2015

Published: October 27, 2015

Data Availability Statement: All relevant data except sequences from resequenced Magnaporthe isolates are within the paper and its Supporting Information files. Sequences from resequenced Magnaporthe isolates are available at http://genome.jouy.inra.fr/gemo

Funding: This work was supported by a PhD grant for DO from the Colombian National Agency for the Science and Technology COLCIENCIAS, COLFUTURO (call 528) and by grants from the "Agence nationale de la recherche" of France ANR-10-INSB-05-0 (French Infrastructure for Integrated

Abstract

Phytopathogenic ascomycete fungi possess huge effector repertoires that are dominated by hundreds of sequence-unrelated small secreted proteins. The molecular function of these effectors and the evolutionary mechanisms that generate this tremendous number of singleton genes are largely unknown. To get a deeper understanding of fungal effectors, we determined by NMR spectroscopy the 3-dimensional structures of the *Magnaporthe oryzae* effectors AVR1-CO39 and AVR-Pia. Despite a lack of sequence similarity, both proteins have very similar 6 β-sandwich structures that are stabilized in both cases by a disulfide bridge between 2 conserved cysteins located in similar positions of the proteins. Structural similarity searches revealed that AvrPiz-t, another effector from *M. oryzae*, and ToxB, an effector of the wheat tan spot pathogen *Pyrenophora tritici-repentis* have the same structures suggesting the existence of a family of sequence-unrelated but structurally conserved fungal effectors that we named MAX-effectors (***M****agnaporthe* **A**vrs and To**x**B like). Structure-informed pattern searches strengthened this hypothesis by identifying MAX-effector candidates in a broad range of ascomycete phytopathogens. Strong expansion of the MAX-effector family was detected in *M. oryzae* and *M. grisea* where they seem to be particularly important since they account for 5–10% of the effector repertoire and 50% of the cloned avirulence effectors. Expression analysis indicated that the majority of *M. oryzae* MAX-effectors are expressed specifically during early infection suggesting important functions during biotrophic host colonization. We hypothesize that the scenario observed for MAX-effectors can serve as a paradigm for ascomycete effector diversity and that the enormous number of sequence-unrelated ascomycete effectors may in fact belong to a restricted set of structurally conserved effector families.

Structural Biology—FRISBI) to AP, KG and JG, ANR-09-GENM-029 (Analysis and comparison of genomes of the fungal pathogen Magnaporthe oryzae—GEMO) to EF and TK and ANR-07-GPLA-0007 (Molecular bases of disease and resistance in the Interaction of Rice and Magnaporthe grisea – IRMA) to TK. This work benefited from interactions promoted by COST Action FA 1208. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist

Author Summary

Fungal plant pathogens are of outstanding economic and ecological importance and cause destructive diseases on many cultivated and wild plants. Effector proteins that are secreted during infection to manipulate the host and to promote disease are a key element in fungal virulence. Phytopathogenic fungi possess huge effector repertoires that are dominated by hundreds of sequence-unrelated small secreted proteins. The molecular functions of this most important class of fungal effectors and the evolutionary mechanisms that generate this tremendous numbers of apparently unrelated proteins are largely unknown. By investigating the 3-dimensional structures of effectors from the rice blast fungus *M. oryzae*, we discovered an effector family comprising structurally conserved but sequence-unrelated effectors from *M. oryzae* and the phylogenetically distant wheat pathogen *Pyrenophora tritici-repentis* that we named MAX-effectors (*M. oryzae* Avrs and ToxB). Structure-informed searches of whole genome sequence databases suggest that MAX-effectors are present at low frequencies and with a patchy phylogenetic distribution in many ascomycete phytopathogens. They underwent strong lineage-specific expansion in fungi of the *Pyriculariae* family that contains *M. oryzae* where they seem particularly important during biotrophic plant colonization and account for 50% of the cloned Avr effectors and 5–10% of the effector repertoire. Based on our results on the MAX-effectors and the widely accepted concept that fungal effectors evolve according to a birth-and-death model we propose the hypothesis that the majority of the immense numbers of different ascomycete effectors could in fact belong to a limited set of structurally defined families whose members are phylogenetically related.

Introduction

Pathogenic microorganisms have to cope with the immune system of their host and therefore deploy measures to hide their presence, disturb host immunity or inactivate defense responses. In all these strategies, proteins secreted by the pathogen during infection and acting on host proteins and cellular processes play a key role [1–3]. These proteinaceous virulence factors named effectors act either extra-cellularly or inside host cells and can possess, depending on the microorganism, very different molecular features.

In fungal pathogens, the main class of effectors are small secreted proteins of less than 200 amino acids expressed specifically during infection and often rich in cysteins [4–6]. Genome sequencing and expression analysis identified hundreds of such effector candidates in individual plant pathogenic fungal species. Few of them, mainly those acting extra-cellularly, are widely distributed among phytopathogenic fungi and contain known motifs or domains, such as NLPs (necrosis and ethylene-inducing peptide 1 (Nep1)-like proteins), LysM domain-containing proteins or protease inhibitors [5,6]. The vast majority of the fungal effectors do not share sequence similarities with other proteins and do not contain conserved motifs. This is very different from the situation in other phytopathogens and in particular oomyctes, an important class of plant pathogens that have similar lifestyles and infection strategies and whose virulence relies also on large effector repertoires. In oomycete pathogens, large families of cytoplasmic effectors with hundreds of members in individual species are defined by the presence of the RXLR or the LFLAK host cell translocation motifs [7–9]. The effector domains of these RXLR and Crinkler (CRN) effectors that mediate virulence functions are highly diversified but contain, in the majority of cases, conserved motifs or domains that are shared between effectors from the same or other species allowing their classification in distinct

families. On the contrary, most fungal effectors are species-specific while few are lineage specific and occur in closely related species. In most phytopathogenic fungi, no large effector gene families were identified [5,6]. The majority of their effectors are singletons and a small proportion belongs to small paralogous groups of rarely more than 3 members. Effector repertoires dominated by gene families of large size counting more than 5 members were only detected in particular cases such as powdery mildew and rust fungi lineages [10–13]. Due to their high diversity and the lack of similarity with other proteins, the mode of action and the role in infection of fungal effectors have to be elucidated case by case and remain still largely unknown [5,6]. In addition, this tremendous diversity raises the question of the evolutionary trajectories of fungal effectors that do not show traces of common origins.

Rice blast disease caused by the ascomycete fungus *M. oryzae* is present in all rice growing areas and causes important harvest losses. Since rice is the main source of calories for half of the human population and since disease control strategies are frequently overcome by the pathogen due to its high genetic plasticity, blast is considered one of the most dangerous plant diseases threatening global food security and hampering attempts to increase rice yield in many parts of the world [14–16]. Due to its economic importance, the status of the host plant rice as a model plant and the ease of cultivation and genetic manipulation of *M. oryzae*, blast disease has become a model for the molecular and genetic investigation of fungal plant diseases [14]. In particular, molecular mechanisms of fungal disease development were studied intensively in *M. oryzae* uncovering important features of fungal virulence [17,18]. Key steps in infection by *M. oryzae* are (i) penetration into epidermal cells by the breakage of the leaf cuticle and epidermal cell walls by an appressorium, a specialized unicellular structure, (ii) biotrophic growth inside the first invaded host cells, followed by (iii) necrotrophic growth associated with active killing of host tissue and the development of disease symptoms and finally, (iv) clonal reproduction and sporulation.

Effectors and in particular cytoplasmic effectors are key elements in *M. oryzae* virulence and particularly important during the biotrophic phase of infection [6,19,20]. However, the function of individual effectors in the infection process has only been established for the LysM effector SLP1 that sequesters chitin fragments and thereby interferes with their recognition by the rice chitin receptor CEBiP, and AvrPiz-t that interferes with host immunity by inhibiting the E3 ubiquitin ligase APIP6 [21,22]. Mutant analysis aiming to demonstrate that individual effectors are important for virulence have often been unsuccessful, probably due to functional redundancy among effectors [23,24]. Approximately 700 of the 1300–1500 secreted proteins encoded in the *M. oryzae* genome are considered effector candidates according to their size of less than 200 amino acids and their lack of homology to proteins of known function [25,26]. Hundreds of them were found to be expressed during appressoria formation or infection [23,26–28].

Some effectors are recognized in certain plant accessions by immune receptors localized either at the plasma membrane or in the cytosol leading to the induction of strong defense responses and resistance to pathogen isolates possessing this effector [29]. The recognized effector is, in these cases, named an avirulence (Avr) protein. In *M. oryzae*, 8 different effectors acting as Avr proteins named PWL2, AVR-Pia, AVR1-CO39, AVR-Pii, AVR-Pik, AvrPiz-t, AVR-Pita and Avr-Pi9 have been cloned molecularly [26,30–35]. They are all translocated into host cells and do not show similarities to proteins of known function with the exception of AVR-Pita that shows homology to neutral zinc proteases [6]. For 7 of them, the matching rice immune receptors that are in all cases cytoplasmic nucleotide-binding and leucine-rich repeat domain proteins (NLRs) have been identified [36–41].

In the present study, the 3-dimensional structures of the *M. oryzae* effectors AVR-Pia and AVR1-CO39 were investigated to deepen our understanding of fungal effector function and diversity. NMR analysis revealed that the structures of both proteins consist of two anti-parallel β-sheets, each having three strands, and linked by one disulfide bond Structural similarity

searches revealed that the *M. oryzae* effector AvrPiz-t and the effector ToxB from the wheat pathogen *Pyrenophora tritici-repens* have similar 6 β-sandwich structures with the same topology [42,43]. Comparisons of the structures of the four effectors that we named MAX-effectors revealed that they share a common architecture but no sequence consensus. Structure-informed and pattern-based searches identified large numbers of weakly homologous MAX-effector candidates in *M. oryzae* and *M. grisea*, and limited numbers or no homologs in other phytopathogenic ascomycete fungi. Expression profiling indicated that the majority of the *M. oryzae* MAX-effector candidates are expressed during early infection. MAX-effectors therefore seem to have undergone a lineage-specific expansion in the *Pyricularia* genus that may be driven by duplications and rapid adaptation to new functions involving important changes of surface properties but conservation of protein architecture. This evolutionary process has the potential to generate large families of structurally related proteins without sequence similarity and may serve as a paradigm for effector evolution and diversification in phytopathogenic ascomycete fungi.

Results

Protein expression

AVR1-CO39 and AVR-Pia proteins, deleted for their endogenous secretion signal, were expressed in *E. coli* with an N-terminally fused signal peptide for secretion in the bacterial periplasm that is cleaved upon secretion, an N-terminal His_6-tag for purification and a TEV1 cleavage site. Recombinant proteins were soluble and were purified to homogeneity from periplasmic protein extracts by Ni-agarose affinity and gel exclusion chromatography (S1 Fig). Both recombinant Avr proteins eluted as monomers from gel exclusion chromatography.

NMR analysis

Recombinant, ^{15}N and ^{13}C-labelled AVR1-CO39 and AVR-Pia proteins produced in ^{15}N and ^{13}C-labelled minimal medium were used for structure determination by two- and three-dimensional NMR experiments. Three-dimensional (3D) HNCO, HNCA, HN(CO)CACB, HN(CA)CO, HNCACB, 2D ^{13}C-detected CON, CACO and 2D-COSY-DQF(D_2O) and TOCSY(D_2O) experiments were used for the backbone and aliphatic side chain resonance assignments. 3D ^{15}N-edited NOESY-HSQC and 2D-NOESY(D_2O) spectra were collected to confirm the chemical shift assignments and generate distance restraints for structure calculations. (Fig 1 and S1 Table). The assigned ^{1}H,^{15}N-HSQC spectra were well dispersed. Residues from the N-terminal tags are still resolved. All amino acids of AVR-Pia and almost all of AVR1-CO39 have {^{1}H-^{15}N} NOE values above 0.8 indicating highly defined structures with low flexibility (S2 Fig). Only N-terminal tags, below residue number 22–23, and C-terminal sequences of AVR1-C039 (amino acids 80–89) show increased flexibility. The strong $d\alpha_N(i, i+1)$ NOEs and weak $d_{NN}(i, i+1)$ NOEs are indicative of a β-structure and consistent with the six β-strands observed in AVR-Pia and AVR1-CO39 (S3 Fig). NHs in slow exchange were consistent with hydrogen bonding networks and were used to derive constraints for the structure calculations.

The ratios of R_2 to R_1 relaxation rates of AVR-Pia and AVR1-CO39 were consistent with a monomeric molecular size (AVR-Pia $\tau_c = 6.2 \pm 0.3$ ns and AVR1-CO39 $\tau_c = 5.7 \pm 0.4$ ns) and thus confirm that both Avrs form monomers in solution (S1 Table) [44].

AVR-Pia and AVR1-CO39 have similar β-sandwich structures

The solution structures of AVR-Pia and AVR1-CO39 were determined based on 1541 and 1286 NOE-derived distance restraints, 90 and 72 dihedral angle restraints and 20 and 15

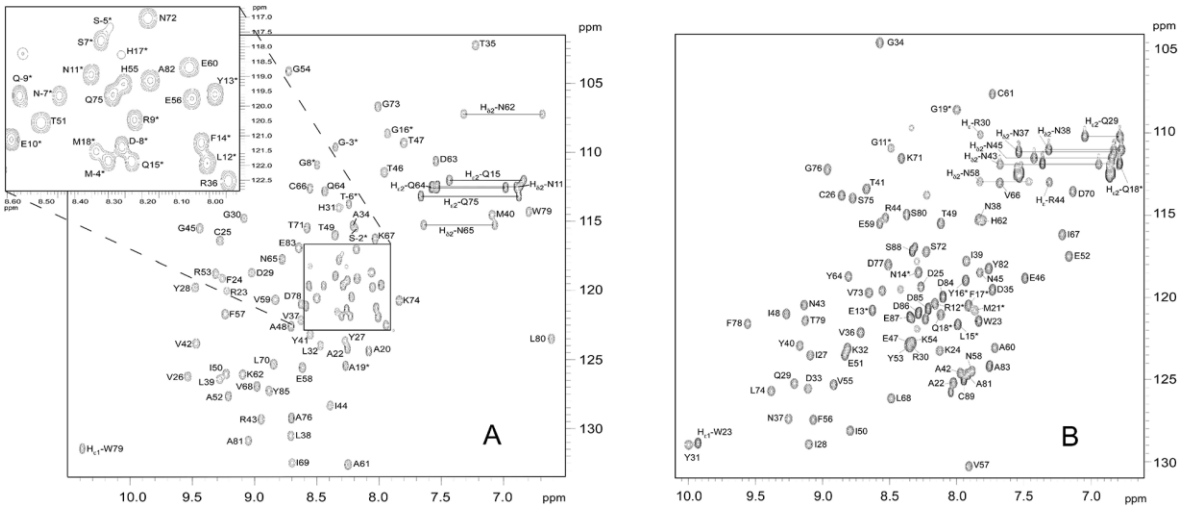


Fig 1. ^{15}N-HSQC spectra of (A) AVR-Pia and (B) AVR1-CO39. In the 15N-HSQC spectra of (A) AVR-Pia and (B) AVR1-CO39, each peak comes from N-H chemical connectivity and has ^{15}N and ^{1}H chemical shift coordinates. There is one backbone N-H group per residue, leading to one HSQC peak per residue. A side-chain NH_2 group gives two HSQC peaks with one common N coordinate. Other side-chains NH groups may also be observed, as Trp Nε1-H and Arg Nε-H. (*) indicates residues in the N-terminal tail. The mature proteins start at residue Ala20 and Trp23 for AVR-Pia and AVR1-CO39, respectively [26,32]. The NH_2 side chains resonances were assigned. The resonances of the tryptophan indole groups are specifically labelled Hε1.

doi:10.1371/journal.ppat.1005228.g001

hydrogen bond restraints, respectively (Fig 2 and Table 1 and S4 Fig). A disulfide bridge between Cys25-Cys66 for AVR-Pia and between Cys26-Cys61 for AVR1-CO39 was added based on cysteine ^{13}Cβ chemical shifts and DTNB quantification of free thiols. The Pro65 in AVR1-CO39 has been determined to be in a cis-conformation according to the ^{13}Cβ chemical shift at 34.4 ppm and strong sequential Hα-Hα NOE. The best conformers with the lowest energies, which exhibited no obvious NOE violations and no dihedral violations > 2° were selected for final analysis.

Surprisingly, both AVR-Pia and AVR1-CO39 proved to possess the same secondary structure elements arranged with the same topology in similar three-dimensional structures (Fig 2). Both proteins are composed of 6 β-strands that form two antiparallel β-sheets packed face-to-face and connected by loops (Fig 2). The first sheet is formed by the three β-strands β1, β2 and β6 while the second sheet contains β3, β4 and β5. In both cases, the two β-sheets pack together by an internal core of hydrophobic residues and one disulfide bridge and the structures belong to the β-sandwich classification. In both Avrs, the β-strands overlay and are similarly oriented (vide infra) but loops differ in length and structure.

AVR-Pia and AVR1-CO39 possess a hydrophobic surface patch

The surface properties of AVR-Pia and AVR1-CO39 are different with the exception of a hydrophobic patch located in both proteins on the side of the β-sandwich that is formed by the first β-sheet (β1-β2-β6) (Fig 2C and 2D). In AVR-Pia, this solvent exposed hydrophobic surface is constituted by the residues F24, V26 and Y28 in β1, V37, L38 and Y41 in β2, and Y85 in β6, and has an area of 372 Å^2. In AVR1-CO39, the solvent exposed hydrophobic surface of the first β-sheet is formed by the residues I27 and Y31 in β1, V36 and I39 in β2 and V73 in β6, as well as W23 from the N-terminus and Y82 from the C-terminus, and has a surface area of 280 Å^2.

ToxB and AvrPiz-t are structural homologs of AVR1-CO39 and AVR-Pia

To identify structural homologs of AVR-Pia and AVR1-CO39, structural similarity searches were performed using the Dali server and the Protein Data Bank [45]. Both queries, with AVR1-CO39 and AVR-Pia, identified the secreted effector protein ToxB from the wheat tan spot pathogen *Pyrenophora tritici-repentis* as well as its natural allele Toxb as the closest structural homologs with the highest Z-scores (S2 Table and Fig 3) [43]. Like, AVR-Pia and AVR1-CO39, ToxB is secreted during infection and is an important determinant of virulence for the tan spot fungus [46]. In addition, the search with AVR1-CO39 identified AvrPiz-t, another avirulence effector of *M. oryzae* that is sequence-unrelated to AVR-Pia and

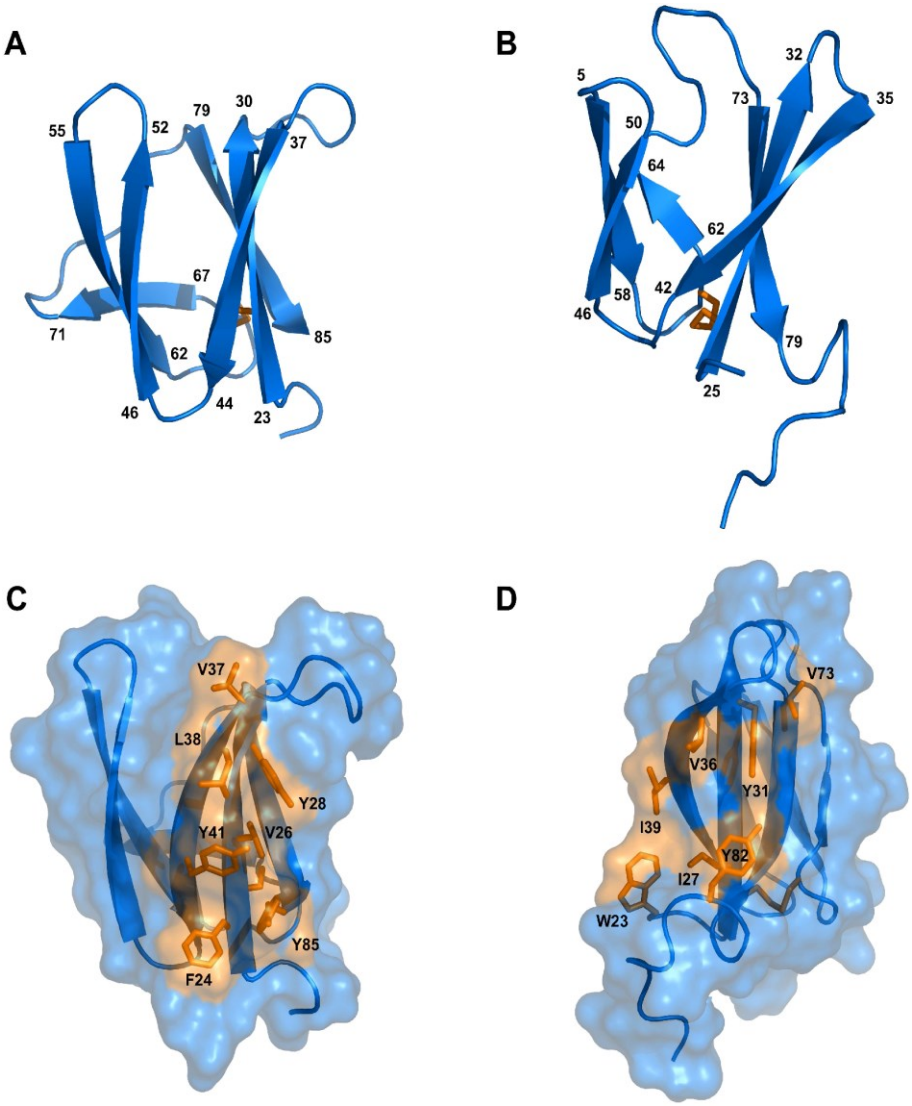


Fig 2. Solution structures of mature AVR-Pia and AVR1-CO39. Cartoon representations of AVR-Pia (A) and (B) AVR1-CO39 highlight the similar β-sandwich structure of both proteins. Yellow sticks represent disulfide bonds. Numbers indicate the residues at β-strands borders. A surface view reveals extended hydrophobic patches on one of the surfaces of AVR-Pia (C) and AVR1-CO39 (D) that are composed of exposed hydrophobic residues labelled in yellow. The Figs were generated using PyMOL (http://www.pymol.org).

doi:10.1371/journal.ppat.1005228.g002

AVR1-CO39 but structurally similar [42]. A pairwise similarity matrix using root-mean-square deviation (rmsd, measured in Å) and DALI Z-scores [45] was established revealing that all proteins are structurally related and that ToxB is closer to all other three structures than the others among them (S2 Table). ToxB and AvrPiz-t are like AVR-Pia and AVR1-CO39, composed of two three-stranded antiparallel β-sheets, β1-β2-β6 and β3-β4-β5, forming a six β-sandwich (Fig 3A–3D). Structure-based sequence alignments provided by DALI revealed, at a first glance, no obvious conservation, but also no clear consensus except buried hydrophobic residues alternating with exposed polar amino acids in the β-strands (Fig 3E). The β-strands β1 and β2 are very similar in length and position in all four proteins, while β3, β4 and β6 display more variation. β5 is the shortest and the most irregular strand. As expected for β-strands, buried and exposed residues alternate, with the exception of β1 where residues have a tendency to be more buried. This is due to the packing of β1 in between the β2 and β6 strands. The loops connecting the β-strands have variable length, and are the sites where most of the residue insertions occur. The disulfide bond between C2 and C43 (ToxB numbering) is well conserved but shifted "in phase" by two residues in AVR-Pia (Fig 3E).

Psi-Blast searches identify in *M. oryzae* and *M. grisea* multiple effector candidates with similarities to *Magnaporthe* Avrs and ToxB

The unexpected finding, that all three *M. oryzae* effectors that have been characterized for their structure so far and one effector from an only very distantly related fungal group are

Table 1. Statistics for 20 NMR structures of AVR-Pia and AVR1-CO39.

	AVR-Pia	AVR1-CO39
NOE restraints	1541	1286
Short range (|i-j|≤1)	1022	745
Medium range (1<|i-j|<5)	128	157
Long range (|i-j|≥5)	391	384
Dihedral restraints (a)	90	72
Number of NOE violations		
> 0.1 Å	9.30 ± 1.29	22.45 ± 2.76
> 0.2 Å	0.35 ± 0.47	0.60 ± 0.74
> 0.3 Å	0.05 ± 0.01	0
> 0.4 Å	0	0
Dihedral violations		
> 2°	0.15 ± 0.26	0.05 ± 0.10
> 4°	0	0
Ramachandran plot statistics (b)		
most favorable regions (%)	86.2	83.2
additionally allowed regions (%)	13.1	15.4
generously allowed regions (%)	0	1.3
disallowed regions (%)	0.7	0.3
Pairwise RMSD (Å) (c)		
Backbone	0.72 ± 0.14	0.49 ± 0.11
Heavy atoms	1.34 ± 0.18	1.02 ± 0.14

Structures were calculated using CYANA, refined using CNS, and analyzed using PROCHECK.

(a) Residues in regular secondary structures were derived from the chemical shifts using TALOS+ software.

(b) PROCHECK was used over the residues 24–85 for AVR-Pia and over the residues 23–83 for AVR1-CO39.

(c) Main chain atoms (N, Cα, C) over the residues 24–85 for AVR-Pia and over the residues 23–83 for AVR1-CO39.

doi:10.1371/journal.ppat.1005228.t001

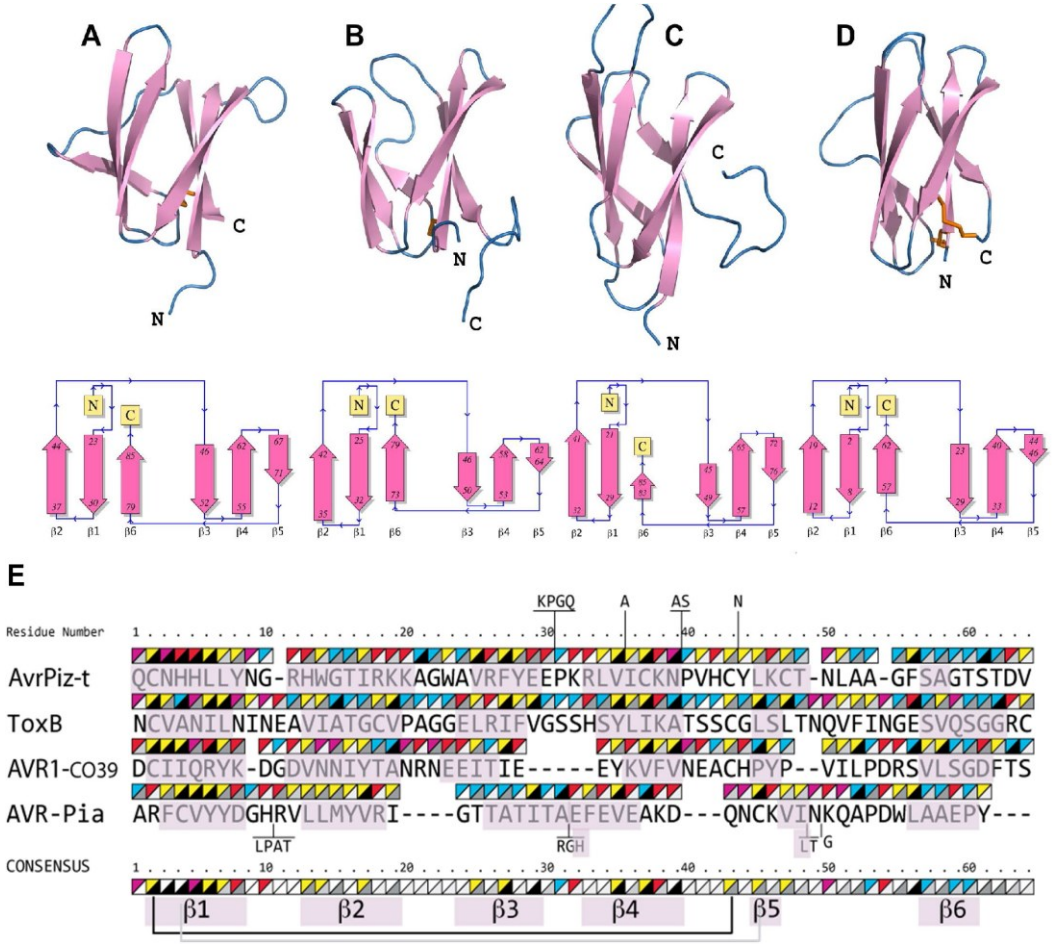


Fig 3. AVR-Pia, AVR1-CO39, AvrPiz-t and ToxB have similar 6 β-sandwich structures. Topology diagrams (lower row) show that AVR-Pia (A), AVR1-CO39 (B), AvrPiz-t (C) and ToxB (D) possess the same fold. Ribbon diagrams (upper row, generated with PyMOL (http://www.pymol.org)) highlight similarities of their structures. Disulfide bonds are shown in the ribbon diagrams by orange sticks. All four structures were superimposed and a structural alignment was derived using DALI with the ToxB sequence as the reference for numbering (E). Residues not aligned to ToxB are connected by vertical lines and correspond to insertions in loops of AvrPiz-t and AVR-Pia. Triangles over the residues indicate chemical properties (upper-left triangle: yellow for hydrophobic, red for charged, pink for Asn and Gln and blue for other residues) and solvent accessibility (lower-right triangle: from black for buried to white for solvent-exposed). The consensus is defined by at least three similar residues per position. Residues forming β-strands are pink. Disulfide bridges in AVR1-CO39 and ToxB are shown below the consensus by a black line and for AVR-Pia by a grey line. For AvrPiz-t, no disulfide bridge was reported despite presence of the two conserved cysteins [42].

doi:10.1371/journal.ppat.1005228.g003

structurally related raised the possibility that these four effectors are members of a widely distributed and abundant fungal effector family characterized by a common β-sandwich structure and high sequence divergence. Simple Blast searches are not suited to identify such distantly related proteins and when performed with the protein sequence of effectors from ascomycete fungi, generally identify no or only very few conserved homologs in the same species. Therefore, more sensitive Psi-Blast searches that use position-specific scoring matrices were performed with AVR-Pia, AVR1-CO39, AvrPiz-t and ToxB. The searches were performed on a protein sequence database combining the protein sequences of the *M. oryzae* reference isolate 70–15, of 5 other rice-infecting *M. oryzae* isolates (TH16, TH12, PH14, FR13 and Guy11), three *M. oryzae* isolates with other host specificities (BR32, US71 and CD156 specific for

wheat, *Setaria italica* and *Eleusine coracana*) and one isolate of the sister species *M. grisea* (BR29). These additional *M. oryzae* and *M. grisea* protein sequences were obtained by whole genome re-sequencing and de novo annotation of proteins and are accessible at http://genome.jouy.inra.fr/gemo [47]). After 4 Psi-Blast iterations and filtering of the results for sequences having an alignment length of at least 40 residues, an overall protein size of less than 180 amino acids and the presence of a predicted signal peptide, 3, 8 and 4 homologs of AVR-Pia, 16, 25 and 16 homologs of AVR1-CO39 and 5, 9 and 6 homologs of ToxB were detected in respectively 70–15, TH16 and BR29 (S3 Table, orthologous sequences present in 70–15 and TH16 were only counted for 70–15). For the other *M. oryzae* isolates similar numbers of homologs as in TH16 were found. The elevated number of homologs present in these isolates but not in 70–15 are due to the fact that the pipeline used for protein annotation in the re-sequenced genomes identified many additional small secreted proteins that are not annotated in 70–15 although the corresponding coding sequences are present in its genome [47]). The similarities were weak (frequently less than 25% identity) but they were consistent with the structural alignment (Fig 3) and included the two cysteine residues. For AvrPiz-t, no homologs that were not already identified by standard Blast were identified in the Psi-Blast search. When 25 additional fungal genomes, including the closely related fungi *M. poae* and *Gaeumannomyces graminis* were added to the database for the Psi-Blast searches, only very limited numbers of homologs (0, 1 or 2) with frequently low e-value scores were identified in other fungi. This suggested that effectors with similaritiy to ***M****agnaporthe* **A**vrs and To**x**B named in the following MAX-effectors that potentially also have 6 β-sandwich structure are present with low frequency in other fungal pathogens but were strongly amplified and diversified in *M. oryzae* and *M. grisea* that both belong to the genus *Pyricularia* in the *Pyriculariae* family [48].

HMM searches identify a huge MAX-effector family in *M. oryzae* and *M. grisea*

To exclude that the Psi-Blast search missed MAX-effectors in the additional fungal genomes due to biases in the search matrix or too low sensitivity and to deepen the search for this class of effectors in *M. oryzae* and *M. grisea* genomes, a hidden Markov model (HMM)-based profile search was performed. This type of profile search is among the most powerful procedure for detecting with high accuracy remote homologies between proteins.

As a first step, a high stringency Blast search with the three *M. oryzae* effectors and a Psi-Blast search with ToxB was performed and the resulting set of closely related sequences was aligned in a multiple sequence alignment constrained by the structural alignment of AVR-Pia, AVR1-CO39, AvrPiz-t and ToxB (S5A Fig). For the *M. oryzae* effectors, the Blast search identified orthologs of the effectors with few polymorphisms in different *M. oryzae* isolates. In addition, for each *M. oryzae* effector, one paralog was identified in *M. oryzae* and one or two paralogs were identified in the *M. grisea* isolate BR29 (S5B Fig). For the *M. oryzae* paralogs, generally several different alleles were identified. For ToxB, in addition to highly homologous sequences from *P. tritici-repentis* and *P. bromi*, 1 homolog was identified in *M. oryzae*, *Bipolaris oryzae* and *Colletotrichum higginsianum*, 2 in C. *fioriniae*, 3 in *C. orbiculare* and 4 in *C. gloeosporioides*. (S5B Fig).

As a second step, an HMM profile was built, starting from the structure-guided multiple sequence alignment from step1 (S5A Fig) and by iteratively searching for homologs in a database containing the small secreted proteins (<170 amino acids) of 25 pathogenic and non-pathogenic ascomycete fungi and of the 9 re-sequenced *M. oryzae* and *M. grisea* isolates from which completely redundant sequences had been removed. At each iteration, the recovered sequences were filtered for alignment of the two cysteins with a spacing of 34 to 49 amino acids

and used to generate a new profile used in the next iteration. The interval of 34 to 49 amino acids was fixed, based on the frequencies of cystein spacings in HMM searches run without this constraint.

This search recovered 161 new, more distantly related sequences of which 154 were from *M. oryzae* or *M. grisea*, 5 from 3 different *Colletotrichum* species, 1 from *Lepthosphaeria maculans* and 1 from *Mycosphaerella graminicola* (recently renamed *Zymoseptoria tritici*) (S6A Fig). This suggests that MAX-effectors have been massively and specifically expanded in *M. oryzae* and *M. grisea*. However, it also indicates their presence in other fungal species, i. e. in *Colletrichum* spp. where they seem to occur at elevated frequencies. The alignment and clustering of the set of 200 sequences combining the 39 sequences used for the initial profile and the 161 new sequences revealed clusters of orthologous sequences originating from the different *M. oryzae* isolates with weak sequence polymorphism between orthologs (S6A and S6B Fig). Frequently, orthologs of *M. oryzae* can be identified in *M. grisea* but never in other fungi. Sequences from different orthologous clusters have high sequence diversity. Only in 3 cases, statistically significant clusters, supported by bootstrap values bigger than 50% can be identified that contain 2 distantly related MAX-effectors or MAX-effector clusters of *M. oryzae*.

A sequence logo derived from the multiple alignment shows the invariant cysteine residues (position 2 and 43 in mature ToxB) that constitute the alignment framework, as well as additional positions that are specifically enriched (Fig 4A). There is a propensity for hydrophobic residues in positions 4 and 6, corresponding to hydrophobic positions in strand β1, in position 27, corresponding to a hydrophobic residue in β3 and in positions 35, 37 and 39 corresponding to β4. Positions 10, 23, 40 and 49 are in loop regions between the pairs of strands β1-β2, β2-β3, β4-β5 and β5-β6 respectively, and are enriched in glycine, polar or charged residues.

The resulting HMM profile was used to search with a relaxed cut-off two different databases: (i) the UniRef90 database that contains non-redundant sequences from a wide range of different organisms and that was used to determine in which type of organisms proteins with the MAX-effector motif occur and to evaluate by this the specificity of the motif and (ii) the previously described fungal genomes and *M. oryzae* and *M. grisea* database to get a precise view of the occurrence of MAX-effectors in a broad range of ascomycete fungi.

The search of the UniRef90 database recovered 70 sequences. All but 3 were from phytopathogenic ascomycete fungi (S7A Fig). The exceptions were from a bacteria, *Pseudomonas* sp. *StFLB209*, living in association with plants, from tomato (*Solanum lycopersicum*) and from a nematode-parasitic fungus (*Arthrobotrys oligospora*) and had low e-values. Among the fungal sequences, 49 were from *M. oryzae* and included AVR1-CO39 and AVR-Pia. The remaining 18 corresponded to previously identified effectors from *Colletotrichum* species (5 *C. orbiculare*, 2 *C. higgensianum*, 3 *C. gloeosporioides*, 2 *C. fioriniae*) that belong as *M. oryzae* to the class of Sordariomycetes and *Z. tritici*, *L. maculans* and *B. oryzae* as well as ToxB from *P. tritici-repentis* and *P. bromi* that are all from the class of Dothideomycete fungi. Clustering of the sequences revealed high sequence diversity and, apart from the Tox-B cluster, no or extremely limited relatedness could be identified (S7A and S7B Fig). Interestingly, with slightly different settings, this search also recovered the well characterized AVR-Pik effector from *M. oryzae* [26]. AVR-Pik clearly fits the MAX-effector pattern but was discarded in the other searches since its secretion signal is not recognized by the SignalP4.1 program used for filtering of the results.

The search of the previously described *Magnaporthe* and other fungal genomes database not filtered for redundancy recovered only limited numbers of MAX-effectors in non-*Magnaporthe* fungal genomes that had, with the exception of one effector from *Fusarium fujicuroi*, already been retrieved in the other searches (Fig 4B and S8A Fig). In *M. oryzae*, between 67 and 38 MAX-effectors per isolate were identified while in *M. grisea*, 37 MAX-effectors were identified (Fig 4B). 46 of the 55 MAX-effectors identified by Psi-Blast in *M. oryzae* 70–15 and

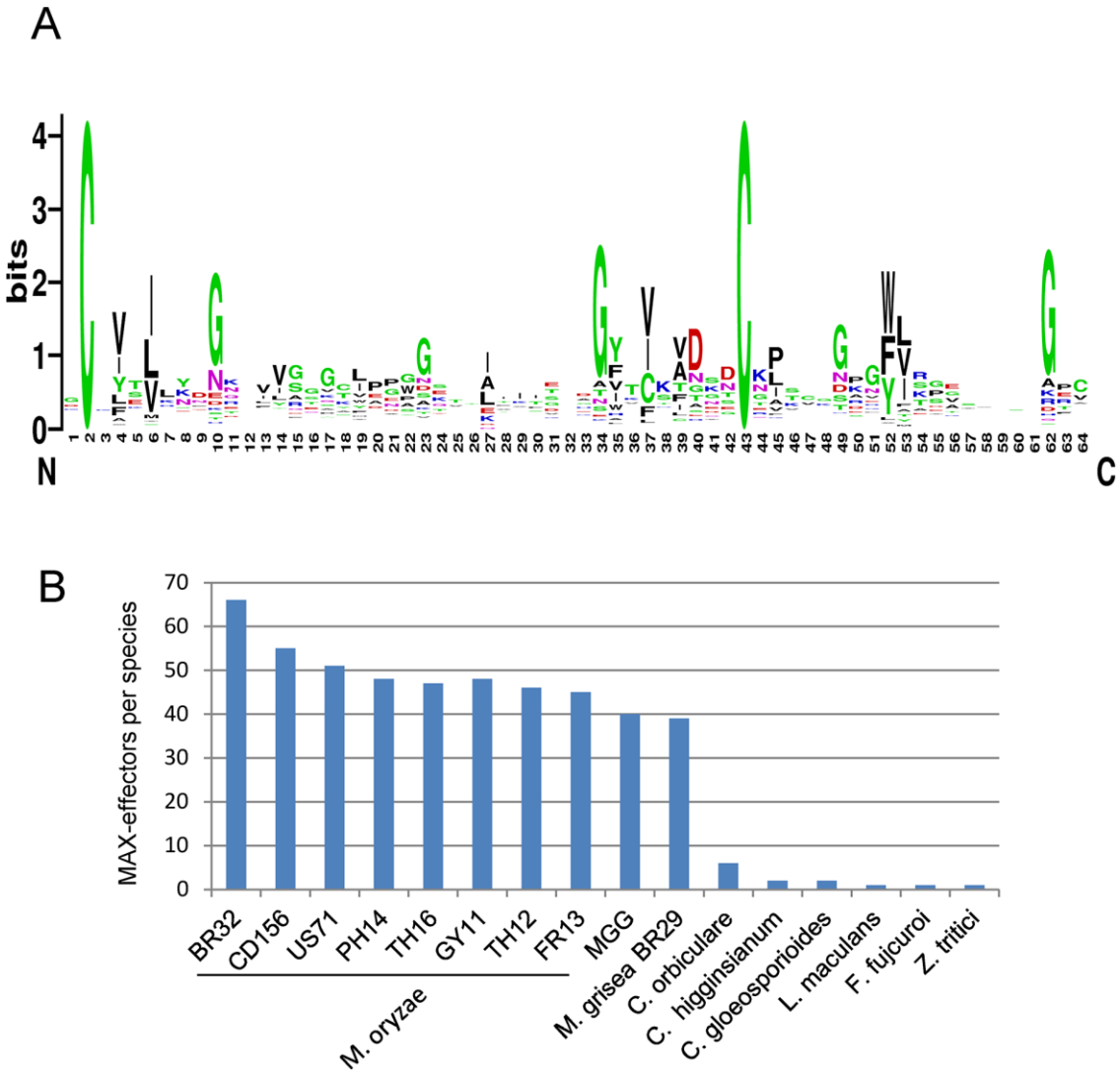


Fig 4. Large numbers of MAX-effectors sharing a characteristic sequence pattern are present in *M. oryzae* and *M. grisea*. A) Sequence pattern of MAX-effectors. The sequence logo was generated using the alignment of MAX-effector candidates identified by a high stringency HMM search (S6 Fig). (B) Numbers of MAX-effector candidates detected by a low stringency HMM sequence pattern search. A database combining 25 pathogenic and non-pathogenic ascomycete fungi and 9 *M. oryzae* and *M. grisea* isolates was searched with an HMM pattern based on a structural alignment of AVR-Pia, AVR1-CO39, AVR-Pia and AvrPiz-t.

doi:10.1371/journal.ppat.1005228.g004

TH16 and in *M. grisea* BR29 (S3 Table) were also found by this HMM search. Alignment and clustering shows that the *M. oryzae* MAX-effectors are generally present in the majority of *M. oryzae* isolates and are grouped in clusters of orthologs (S8A and S8B Fig). Many of these orthologous clusters also contain an ortholog from the *M. grisea* isolate BR29 that shows however higher sequence divergence. Only six statistically significant clusters (bootstrap > 50%) that contain more distantly related *M. oryzae* effectors from different orthologous groups are identified. Otherwise, the sequence diversity between proteins from different *M. oryzae* ortholog clusters is so strong that classical tree building methods do not detect statistically significant sequence relatedness. The non-*Magnaporthe* MAX-effectors do not cluster significantly with

Magnaporthe MAX-effectors and 8 of the 10 *Colletotrichum* effectors are comprised in three different *Colletotrichum*-specific clusters.

Taken together, the different HMM searches reveal that the MAX-effector motif is specific for effectors from phytopathogenic ascomycete fungi. MAX-effectors are identified with low frequencies in phytopthogenic ascomycete fungi from the class of Dothideomycetes and seem to have expanded moderately in different *Colletotrichum* species (i.e. *Colletotrichum orbiculare*). Only in *M. oryzae* and *M. grisea*, MAX-effectors expanded and diversified massively to become a dominating family of virulence effectors in these pathogens.

Expression profiling shows that a majority of MAX-effectors is expressed specifically during biotrophic infection

To test if the *M. oryzae* MAX-effectors identified by the HMM profile search could be involved in plant infection, the expression of 50 different candidate MAX-effector-coding genes was analyzed by qRT-PCR in infected rice leafs and in *in vitro* grown mycelium (S4 Table). 30 genes showed early infection-specific expression with a majority of profiles (25) that strongly resemble the biotrophy effector marker gene *BAS3* (Fig 5 and S9A and S9D Fig) [23]. The expression pattern of all these genes and of 3 genes coding for MAX-effectors identified only by Psi-Blast searches was clearly different from the markers of very early or late infection (Orf3 and MGG01147, respectively). For 18 genes, no significant expression was detected and only 2 genes were expressed constitutively with significant expression in the mycelium (Fig 5 and S9C and S9D Fig). Therefore, the majority of the MAX-effector candidates seems specifically expressed during biotrophic infection and can therefore be considered as potential virulence effectors.

Discussion

In this study, we have determined by NMR spectroscopy the 3-dimensional structures of two different effectors of *M. oryzae*, AVR1-CO39 and AVR-Pia. Although the two proteins have no evident sequence similarity they possess similar 6 β-sandwich structures formed in both cases by two β-sheets each formed of three β-strands oriented in an antiparallel manner. Interestingly, similar β-sandwich structures have previously been found for AvrPiz-t, the only other structurally characterized *M. oryzae* effector and for ToxB, an effector from an only very distantly related plant pathogenic ascomycete fungus, *P. tritici-repentis*. Overlay of the structures and structural alignments revealed that the nature and number of secondary structural elements and the topology of their fold are the same in all four effectors. In addition, all four proteins are stabilized by buried hydrophobic residues coming for their majority from the β-strands and by a disulfide bond between conserved cysteins located in the beginning of β1 and in the beginning or just before β5. However, the orientation and the length of certain β-strands, i.e. β-5, vary considerably and the sequences and the length of loops are highly variable resulting in proteins with very different shapes and surface properties. Due to the high sequence diversity, similarity among the MAX-effectors is therefore only detected when their structure is taken into consideration.

Hydrophic surface patches in AVR-Pia and AVR1-CO39 are potential sites of protein-protein interaction

The only similarity of the surfaces of AVR1-CO39 and AVR-Pia is an extended hydrophobic area on the surface formed by β1, β2 and β6. Such extended and exposed hydrophobic areas are uncommon since protein surfaces are generally in contact with solvent water molecules

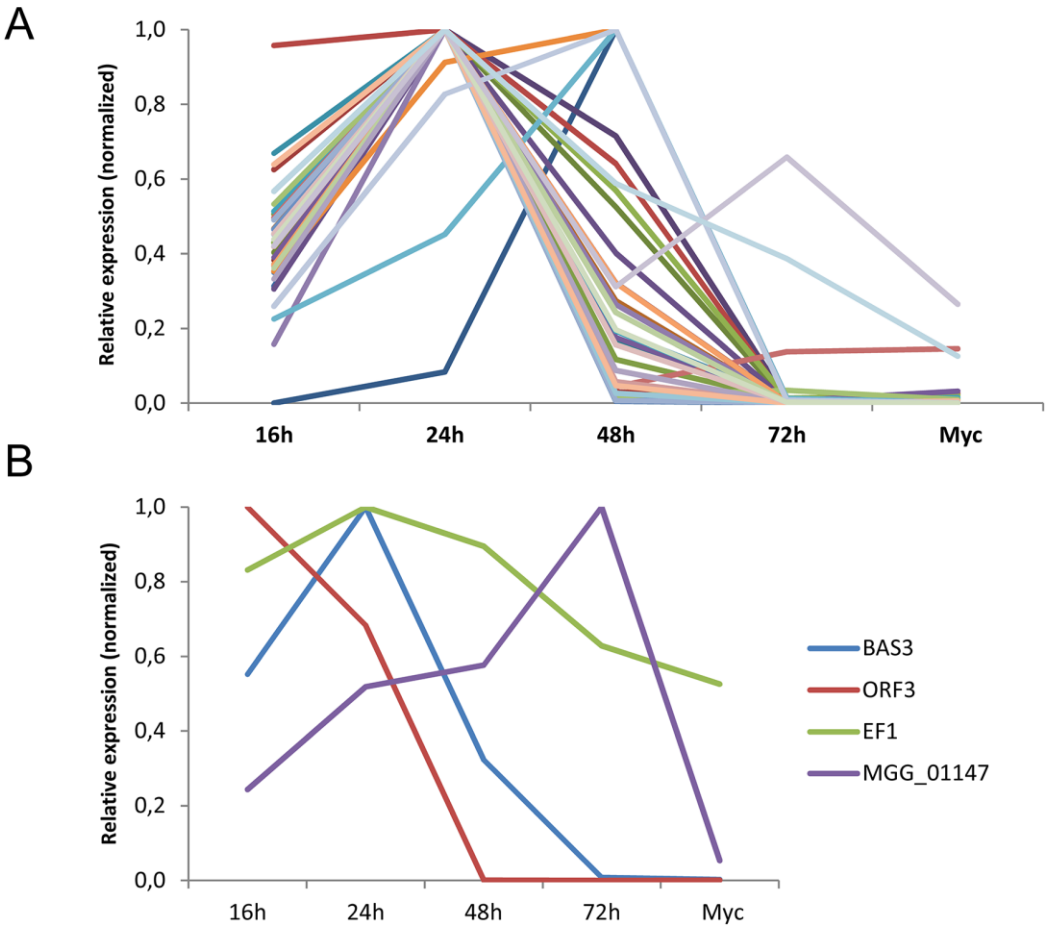


Fig 5. The majority of *M. oryzae* MAX-effectors is expressed specifically during biotrophic infection. mRNA levels of *M. oryzae* genes coding for 32 different MAX-effectors (A) and marker genes (B) for appressorium formation and very early infection (*ORF3* of the *ACE1* cluster, *MGG_08381*), biotrophic infection (*BAS3*, *MGG_11610*), late infection *(MGG_01147*) and constitutive expression (EF1α, *MGG_03641*) were determined by q-RT-PCR in rice leaf samples harvested 16, 24, 48 and 72 h after inoculation and mycelium grown *in vitro*. Relative expression levels were calculated by using expression of a constitutively expressed *Actin* gene (*MGG_03982*) as a reference and normalized with respect to the highest expression value. Values are means calculated from the relative expression values of three independent biological samples. Individual expression profiles are in S9 Fig.

doi:10.1371/journal.ppat.1005228.g005

and they are frequently involved in protein-protein interactions. Previous studies on the recognition of AVR-Pia by the rice NLR immune receptor RGA5 support that the hydrophobic surface of AVR-Pia could indeed be involved in protein binding [37]. AVR-Pia binds physically to a C-terminal domain of RGA5 homologous to heavy metal-associated (HMA) domain proteins related to the copper chaperone ATX1 from *Saccharomyces cerevisiae* (RATX1 domain). This binding is required to derepress a second NLR RGA4 that activates resistance signaling [49]. A natural allele of AVR-Pia (AVR-Pi-H3) where the surface exposed phenylalanine 24 and threonine 46 situated respectively in and at the border of the hydrophobic patch are replaced by serine and asparagine loses binding to $RGA5_{RATX1}$ and does not trigger resistance [37]. Structural information will now guide further functional studies to elucidate if other amino acids situated in or at the border of the hydrophobic patch are also involved in $RGA5_{RATX1}$-binding and to validate by this the role of the hydrophobic patch as a protein-protein interaction surface.

MAX-effectors have different molecular properties and activities

Common features of the *M. oryzae* MAX-effectors are that they act intracellular in host cells [21,24,32] and are recognized by NLR immune receptors in resistant rice genotypes: AVR1-CO39 and AVR-Pia by the same NLR pair RGA4/RGA5 and AvrPiz-t by the NLR immune receptor Piz-t [37,39,41]. While the molecular bases of the recognition of AVR1-CO39 and AVR-Pia by RGA4/RGA5 are beginning to be elucidated, details of AvrPiz-t recognition are not known. Also, whether the three *M. oryzae* MAX-effectors target similar host processes and host proteins is not known. AvrPiz-t was described to target the host ubiquitin proteasome system by binding and inactivating the RING E3 ubiquitin ligase APIP6 [21] but virulence targets of AVR-Pia and AVR1-CO39 have not been described. However, it has been hypothesized that both proteins target RATX1 proteins homologous to the $RGA5_{RATX1}$ domain that was suggested to act as a mimic for AVR-Pia and AVR1-CO39 targets [50]. Therefore, we assume that AvrPiz-t on the one hand and AVR-Pia and AVR1-CO39 on the other have different molecular activities and target different host proteins. This would be in accordance with the high divergence of their shapes and their surface properties. That AVR-Pia and AVR1-CO39 interact with the same immune receptor by binding to the same sensor domain and potentially interact with the same host targets is striking because apart from the extended hydrophobic patch on the β1β2β6 surface they share no apparent similarities with respect to their shapes and surfaces. It will therefore be important to elucidate in the future which amino acids of AVR-Pia and AVR1-CO39 bind to $RGA5_{RATX1}$ and which surfaces of $RGA5_{RATX1}$ are involved in binding to each of the two effectors to better understand specificity in effector recognition. In addition, identification of AVR1-CO39 and AVR-Pia targets as well as ToxB targets for which molecular details of activity are also lacking will be important to understand how MAX-effectors promote virulence and to understand the link between MAX-effector structure and function.

MAX-effectors are a highly diversified effector family specific to phytopathogenic ascomycete fungi and underwent expansion in *M. oryzae* and *M. grisea*

Structure-informed pattern searches identified huge numbers of MAX-effector candidates that possess as the structurally characterized MAX-effectors very high sequence diversity and probably also possess a 6 β-sandwich structure stabilized by buried hydrophobic residues from β-strands and a disulfide bond between conserved cysteins connecting β1 and β5. Systematic prediction of the secondary structure of the MAX effector candidates using SSPRO 5 software identified with high frequency two β-strands, β1 located after the first cysteine and β4 located before the second cysteine (S10 Fig). The other regions of the sequences had more variable secondary structure predictions which is also reflected by a less defined pattern in these regions (Fig 4A). High sequence diversity among MAX-effector candidates could as in the case of the structurally characterized MAX-effector be the consequence of interchangeability of buried hydrophobic core residues, variation in the lengths of some β-strands (i.e. β5), exchange of surface exposed residues and deletion or insertion of residues in exposed loops.

MAX-effectors were specifically detected in phytopathogenic ascomycetes from the classes of Sordariomycetes and Dothideomycetes. One MAX-effector per species was detected in phytopathogenic fungi of the class of Dothideomycetes (*L. maculans*, *P. tritici-repentis*, *Z. tritici* and *B. oryzae*) and higher numbers (2–6) occur in fungi from the genus *Colletotrichum*. Only in *M. oryzae* and *M. grisea* that are both from the genus *Pyricularia* huge numbers of MAX-effector candidates were detected and expression profiling confirmed that most of them are likely *bona fide* effectors expressed specifically during biotrophic early infection. With 40–60

effectors which represents 5–10% of the candidate effectors of individual *M. oryzae* or *M. grisea* isolates, MAX-effectors can be considered a dominant class of effectors in these fungi [24,47]. This is further supported by the finding that 5 of the 51 biotrophy-associated proteins identified by transcriptome analysis are MAX effectors (MG02546, MG08414, MG08482, MG09425 and MG09675) [23]. Also, the *M. oryzae* effector AVR-Pik fits the MAX-effector pattern further highlighting the outstanding importance of this effector family that comprises 4 out of 8 cloned Avr effectors in the blast fungus [6]. It is striking that the only other group of fungi with elevated numbers of MAX effectors are *Colletotrichum* species. *Colletotrichum* fungi are phylogenetically only distantly related to *M. oryzae* and *M grisea* but employ a similar hemibiotrophic infection strategy characterized by appressorium-mediated penetration into the host and growth inside invaded plant cells during biotrophic infection. It will be interesting to determine in the future whether MAX effectors play similar roles in these early infection processes in both groups of fungi.

In *Gaeumannomyces graminis* and *M. poae* that belong to the closely related *Magnaporthaceae* family no MAX-effectors were detected [48]. The expansion of MAX-effectors therefore occurred probably in a common ancestor of *M. oryzae* and *M. grisea* since clear orthologous relations can be established between many MAX-effectors from *M. oryzae* and *M. grisea* but after the split of the *Magnaporthaceae*. Expansion and diversification of the MAX-effectors is clearly continuing since frequently orthologs in *M. oryzae* or *M. grisea* cannot be identified and duplication, loss and diversification of MAX-effectors in host specific lineages of *M. oryzae* is observed (S8B Fig). Genome sequencing of additional species from *Pyricularia* and other genera in the *Pericullariae* will allow to further strengthen the hypothesis of lineage-specific expansion of MAX-effectors.

Lineage specific expansion of effector families has been observed in other fungi such as mildew and rust fungi whose effector repertoires are dominated by effector families that contain frequently numerous members and are for their majority restricted to individual species or precise clades [10,51]. However, in these cases, sequence divergence is not as strong as in MAX-effectors since sequence-based comparisons allow the establishment to these effector families.

On the contrary, the effector repertoires of ascomycete phytopathogens outside the mildew lineage contain hundreds of sequence-unrelated effectors and the evolutionary origin of these huge amounts of species or clade specific genes is an open question. Duplication and diversification eventually driven by localization of the genes in transposon rich regions, genome reshuffling or transfer of accessory chromosomes were convincingly proposed as potential mechanisms to create effector diversity but the apparent lack of relatedness of ascomycete effectors remains unexplained [52–55]. Establishment of a huge effector family in *M. oryzae* and *M. grisea* that is also present at much lower frequency in other ascomycete pathogens sheds new light on the origin and relatedness of ascomycete effectors.

Diversifying rather than convergent evolution leads to highly diversified effector families

Theoretically, convergent evolution as well as diversifying evolution can explain the situation observed for the MAX-effectors characterized by a broad and patchy distribution, high diversification and limited sequence homology as well as a shared sequence pattern and probably the same structure. Convergent evolution would apply if these proteins with similar functions and a similar fold appeared repeatedly in phytopathogenic ascomycetes and eventually evolved independently in different clades. Under diversifying evolution, a protein or protein family present in a common ancestor has been strongly diversified in different lineages of ascomycete fungi and frequently lost during evolution in certain lineages and species. The scenario of

convergent evolution of MAX-effectors cannot be excluded but is clearly less parsimonious. It raises the question why MAX-effectors do not occur in organisms with similar lifestyles outside the Sordariomycete and Dothideomycete pathogens such as phytopathogenic basidiomycetes or oomycetes. In addition, there are no well-documented examples of convergent evolution towards similar folds or sequence patterns for pathogenic effectors or secreted fungal proteins involved in adaption to the environment while comparative genomics studies in fungi and oomycetes are beginning to identify such widely distributed gene families that are shaped by strong diversifying selection and that can only be properly reconstructed when pattern-based searches and structure information are taken into consideration. The best documented example is certainly the WY-domain family among the RXLR effectors that is specific to the *Peronosporales* clade in oomycetes and evolves by diversifying evolution [8,9,56–58]. Careful sequence analysis involving pattern searches identified the W, Y and L sequence motifs in the effector domains of a majority of the *Phytophtora* RXLR effectors that are frequently completely sequence unrelated [9]. Functional analysis confirmed the importance of these motifs for effector function [59] and structure analysis of the effector domain of different RXLR effectors with limited sequence homology revealed that conserved sequence motifs reflected a conserved, highly similar 3-dimensional structure named the WY-domain fold [56,60–62]. PexRD2 and AVR3a11 show e.g. only 14% amino acid identity in a structure-based alignment but overlay of their structures has an RMSD score of 0.73 Å. As in the case of the β-sandwich fold of the MAX-effectors, the WY-domain fold tolerates insertion or deletion of amino acids in the loops, exchange of surface exposed amino acids and is stabilized by hydrophobic core residues that can be exchanged as long as hydrophobicity is maintained [56]. This flexible structure allows to generate effectors with highly variable shapes and surface properties and studied WY-domain effectors showed very diverse molecular activities, target different host proteins and are recognized by different NLR immune receptors [7,56].

An example of rapidly evolving proteins from fungi that are structurally but not sequence-conserved are hydrophobins that are low molecular mass secreted proteins important for the impermeabilization of fungal cell walls, adhesion to hydrophobic surfaces and pathogenicity [63]. Hydrophobins were shown to evolve rapidly according to a birth-and-death mechanism [64], are widely distributed in a broad range of basidio- and ascomycete fungi and are characterized by sequence patterns but no sequence homology [63,65]. Structure analysis demonstrated that distantly related hydrophobins are structurally related supporting a common evolutionary origin [66].

Another example of a fungal gene family that is rapidly evolving according to a birth-and-death model are the Hce2 proteins (**h**omologs of ***C****ladosporium fulvum* **E**CP2) that are present in a wide range of basidio and ascomycete fungi and seem to act as effectors in pathogenic fungi and potentially in stress responses in non-pathogenic fungi [67]. Much like MAX-effectors they show patchy distribution, lineage-specific expansions and high sequence diversification.

MAX-effectors may serve as a paradigm for the evolution and diversification of effectors in phytopathogenic ascomycetes

Based on our discovery of the MAX-effector family and the widely accepted concept that fungal effectors evolve according to a birth-and-death model we propose the hypothesis that the majority of the immense number of different ascomycete effectors could in fact belong to a restricted set of structurally defined families whose members are phylogenetically related. These families of structurally conserved effectors are expected to be, as the MAX-effectors widely distributed with frequent losses on the one hand and lineage specific expansions on the

other leading to effector families that are particularly important in certain fungal clades but not in others. The evolution of individual effectors is so rapid and their adaptation to new functions so profound that sequence homology and resulting phylogenetic signals are rapidly lost although the basic protein architecture may frequently be conserved because it represents a good solution to many general constraints effectors have to face such as stability in the fungus-host interface or translocation into host cytosol. Sequence homology can therefore only be detected in orthologs from closely related species but in paralogs from the same species or homologs from more distantly related species no similarity is detected on the sequence level. Only structure-informed and pattern-based searches reveal the hidden relatedness of ascomycete effectors. This hypothesis is also supported by the recent identification of an effector super family in the powdery mildew fungus *Blumeria graminis* fsp *hordei* by structural modelling [51]. 72 effectors from different families established by sequence homology or with no homology to other proteins had 3D structure models with similarity to ribonucleases suggesting a common origin and a conserved structure in this superfamily of sequence diverse effectors.

Knowledge on the structures of fungal effector proteins is extremely limited and outside of the MAX-effectors the structures of only three cytoplasmic fungal effectors have been determined. AvrL567 from the rust fungus *Melampsora lini* and ToxA from *P. tritici-repentis* have distantly related β-sandwich structures whose topologies are completely different from the MAX-effectors and AvrM has a helical structure [68–70]. Therefore, the elucidation of the 3-dimensional structures of additional fungal effectors is a priority for a better understanding of their diversity and will teach us to what extent structurally conserved but sequence-diversified effector families dominate the huge and extremely diverse effector repertoires of phytopathogenic fungi.

Methods

Protein expression and purification

The sequence for the mature protein (residues 20–85 for AVR-Pia, and residues 23–89 for AVR1-CO39) was inserted into the pET-SP vector by ligation of PCR using NdeI-BamHI sites. PCR products were generated using the forward and reverse oligos tatcatatggctGCGCCAGCT AGATTTTGCGTCTAT and tatggatccCTAGTAAGGCTCGGCAGCAAG or tatcatatGCTTG GAAAGATTGCATCATCCA and tatggatccGATCAACAAGACTCATCGTCGTCA for respectively AVR-Pia or AVR1-CO39. The pET-SP vector was constructed from pET-15b (Merck-Millipore, Darmstadt Germany) by inserting a periplasmic secretion sequence, a hexahistidine tag and a TEV cleavage site at the N-terminus of the protein adding an extra 31 amino acid sequence at the N-terminus of the recombinant proteins (sequence MKKTAIAIA VALAGFATVAQA_APQDNTSMGSSHHHHHHSSGRENLYFQGHMA). The plasmids pET-SP-AVR-Pia and pET-SP-AVR1-CO39 were used to transform *E. coli* BL21 (DE3).

Transformed cells were grown in an autoinducing minimal media C-750501 [71] at 37°C for 24h. To generate isotopically-labeled samples for NMR spectroscopy, we used $^{15}NH_4Cl$, $^{13}C_3$-glycerol and $^{13}C_6$-glucose as the primary nitrogen and carbon sources. Cells were harvested by centrifugation and the pellet was resuspended in lysis buffer (200 mM TrisHCl pH8, 200mM Sucrose, 0.05mM EDTA, 50μM lysozyme). After 30 minutes incubation, cell debris were removed by centrifugation at 12 000 *g* for 15 min at 4°C. The resulting crude protein extracts were loaded on an AKTA basic system into a HisTrap 5ml HP columm (GE Healthcare), equilibrated in buffer A (50 mM TrisHCl, pH 8.0, 300 mM NaCl, 1 mM DTT, 0.1 mM Benzamidine). The His-tagged protein was eluted from the affinity column with buffer B (buffer A supplemented with 500 mM imidazole). Fractions containing the protein were

identified by SDS-PAGE and pooled. The protein was further purified by gel filtration using a Superdex S75 26/60 (GE Healthcare) column in buffer A and pure fractions were pooled.

The elution profiles indicated that AVR-Pia and AVR1-CO39 eluted as single monomeric species (Fig 1). Ellman's reagent, 5, 5'-dithio-bis-(2-nitrobenzoic acid), DTNB, was used for quantitating free sulfhydryl groups [72]. Briefly, aliquots of standard (cysteine, Sigma, 12.5 μM to 75 μM) or sample (50 μM) were reacted with 0.1 mM DTNB reagent in 100 mM sodium phosphate pH 8.0, 1mM EDTA buffer. Free sulfhydryl groups were also measured in denaturating conditions using the same buffer supplemented with 6M Guanidinium Chloride. Absorbance was read at 412 nm on a NanoDrop 2000, and the concentration of free thiols was determined from the standard curves.

NMR samples

The NMR samples were prepared with 1mM of purified protein at 10% D_2O and 0.5 mM DSS as a reference. For AVR-Pia the purification buffer was exchanged with phosphate buffer (20 mM potassium-sodium phosphate, pH 5.4 and 150 mM NaCl), by filtrating with Centricon. The purified AVR1-CO39 proteins were dialyzed in 20 mM sodium phosphate, pH 6.8, 150 mM NaCl and 1 mM DTT. For the D_2O experiments, a non-labeled sample was lyophilized and dissolved in D_2O.

Nuclear magnetic resonance spectroscopy

Spectra were acquired on 500 and 700 MHz Avance Bruker spectrometers equipped with triple-resonance (^1H, ^{15}N, ^{13}C) z-gradient cryo-probe at 305 K. Experiments were recorded using the TOPSPIN pulse sequence library (v. 2.1) (S2 Fig). 2D-NOESY experiments with excitation sculpting water suppression were acquired at 305K, with mixing times from 100 to 150 msec. All spectra are referenced to the internal reference DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) for the ^1H dimension and indirectly referenced for the ^{15}N and ^{13}C dimensions [73].

NMR data was processed using Topspin (v. 3.2) and were analyzed using strip-plots with Cindy in house software and CCPN [74] [analysis v 2.3]. Side chain assignments were carried out using 2D-NOESY, 2D-TOCSY and COSY-DQF experiments with D_2O samples, combined with ^{15}N-NOESY-HSQC and ^{15}N-TOCSY-HSQC 3D spectra. For AVR-Pia, the two N-terminal residues Ala-Pro and the His-tag, Ser-His_6-Ser were not assigned. For AVR1-CO39, the tag-residues Asp(-7)-Asn(-8) and the stretch Ser_2-His_6-Ser_2 were not assigned. The ^{15}N and ^{13}C assignments were derived from the 3D spectra at 500 MHz.

^{15}N backbone amide NMR relaxation data

Relaxation data were acquired at 305K on a Bruker Avance 500 MHz spectrometer using R_1, R_2 and $^{15}N\{^1H\}$ heteronuclear NOE pulse sequences (TOPSPIN library, v 2.1). NMR samples of 500 μL at 0.85 mM and 0.3 mM were used for AVR-Pia and AVR1-CO39, respectively. R_1 experiments were performed with nine relaxation delays (18, 54, 102, 198, 294, 390, 582, 774 and 966 ms). R_2 experiments were carried out employing a Carr–Purcell–Meiboom–Gill (CPMG) pulse train [75,76] with eight relaxation delays (16, 32, 48, 64, 96, 128, 192 and 256 ms). A recycle delay of 2.5 s was employed in R_1 and R_2, experiments, and ^{15}N decoupling during acquisition was performed using a GARP-4 sequence. In heteronuclear $^{15}N\{^1H\}$NOEs, proton saturation was achieved during the relaxation time by application of high-power 120° pulse spaced at 20 ms intervals for 3 s prior to the first pulse on ^{15}N [77]. A relaxation delay equal to 6 s between each scan was used. Relaxation parameters, R_1, R_2 and NOEs were determined from the analysis module of CCPN [74].

Structure calculation

The programs CYANA [78] and CNS [79] were used for automatic NOE assignments and structure calculations. The NH, Hα, ^{15}N, ^{13}Cα and ^{13}Cβ chemical shifts were converted into Φ/Ψ dihedral angle constraints using TALOS+ (v. 1.2) [80]. The CANDID procedure of CYANA (v 2.1) was used to assign the 3D-peaks list from the ^{15}N-NOESY-HSQC spectra. NOE assignments were inspected and used in a new CANDID assignment run including peaks from the 2D-NOESY spectra (with 100 and 150 msec mixing times for AVR-Pia and 100 and 200 msec for AVR1-CO39). A disulfide bridge Cys25-Cys66 for AVR-Pia and between Cys26-Cys61 for AVR1-CO39 was added based on cysteine Cβ chemical shifts and DTNB quantification of free thiols. NOE constraints were inspected and classified from very strong, strong, medium weak and very weak, corresponding to 2.4, 2.8, 3.6, 4.4 and 4.8 Å upper bound constraints, respectively. Final structure calculations were performed with CYANA (v. 2.1) using 1541 and 1286 distance restraints, for AVR-Pia and AVR1-CO39, with 90 and 72 Φ/Ψ dihedral angle constraints, respectively. The 30 conformers with lowest target function starting from 200 initial structures, were refined by CNS (v. 1.2) using the refinement in water of RECOORD [81]. The final 20 conformers were selected with the lowest NOE and dihedral angle violations. These are the structures discussed herein and deposited (PDBs, 2MYV and 2MYW). The final 20 structures contained no NOE violations greater than 0.3 Å and no dihedral angle constraint violations greater than 2°. Structures were validated using PROCHECK [82].

Sequence analysis

Two sequence databases were used, the UniRef90 release 2015_03 [83] and a database build from the genomes of the ascomycete fungi *Magnaporthe oryzae* (reference isolate 70–15), *Colletotrichum graminicola*, *Colletotrichum higginsianum*, *Fusarium graminearum*, *Fusarium oxysporum*, *Gaeumannomyces graminis*, *Magnaporthe poae*, *Neurospora crassa*, *Pyrenophora tritici-repentis*, *Verticillium dahliae*, *Aspergillus fumigatus*, *Aspergillus nidulans*, *Blumeria graminis*, *Botrytis cinerea*, *Colletotrichum gloeosporioides*, *Colletotrichum orbiculare*, *Dothistroma septosporum*, *Fusarium fujikuroi*, *Fusarium pseudograminearum*, *Fusarium verticillioides*, *Leptosphaeria maculans*, *Phaeosphaeria nodorum*, *Pyrenophora teres*, *Trichoderma virens*, *Tuber melanosporum* and *Zymoseptoria tritici* (all from the Ensembl Fungi database http://fungi.ensembl.org) as well as the genomes of eight *M. oryzae* isolates specific for *Eleusine coracana* (CD156), *Triticum aestivum* (BR32), *Setaria italica* (US71) and *Oryza sativa* (TH16, GY11, FR13, TH12, PH14) and one *M. grisea* isolate (BR29) pathogenic to *Digitaria* ssp (genome sequences at http://genome.jouy.inra.fr/gemo) [47]. Sequences without signal peptide (according to SIGNALP 4.1 [84]) bigger than 170 amino acids or with less than 2 cysteine residues were removed. For the initial HMM search, identical sequences were reduced to only one occurrence in the databases.

The start of the search was a structural alignment with TM-align [85] and the structures of AVR-Pia, AVR1-CO39, AvrPiz-t and ToxB complemented with sequence homologues found by single queries using BLAST (v 2.2.27+) with a stringent cut-off E-value = 1e-6. For the ToxB query, two iterations of NCBI PSI-BLAST were used on the NR database with a cut-off E-value = 1e-4 (S5A Fig).

This initial alignment was used as input to look for homologs in the filtered and non-redundant fungi database using HMMERsearch program from the HMMER package v 3.0 [86] with a 1e-6 E-value cut-off. For each run, only sequences where the two cysteine residues were aligned were kept, and the output alignment was used as input query for a new HMMERsearch run. This run was repeated until reaching convergence. New iterations were then done with

increased E-value cut-off at 1e-5 and 1e-4. From the last alignment, a histogram indicated that the two aligned cysteine residues were separated by at least 34 and at most 49 amino acids.

The full homolog search was re-started, as described above, but this time using also the aligned cysteine separations as an additional constraint for filtering homologs after each HMMERsearch run. The HMMERsearch runs were repeated until convergence for raised threshold E-values 1e-6, 1e-5, 1e-4 and finally 1e-2. The homolog ensembles obtained for the three E-values cut-off, 1e-6, 1e-4 and 1e-2 were aligned with Muscle v3.8 [87] (S6A Fig for E-value 1e-4). The derived logo was built from the HMMER search with E-value of 1e-4 using Weblogo3 [88]. The multiple sequence alignment (MSA) derived from the HMMER search with E-value 1e-4 was used as input to look for homologs in the redundant fungi database and the UniRef90 database, using HMMERsearch with an E-value threshold of 1e-1. Diversity trees were built from alignments generated with Muscle v3.8 using the Neighbor-Joining method with the MEGA6 package [89].

Fungal growth and infection assays

For analysis of gene expression *in vitro* grown mycelium, *M. oryzae* isolate Guy11 was grown in liquid medium (glucose 10g/L, KNO_3 3g/L, KH_2PO_4 2g/L and yeast extract 2g/L) at 120 rpm on a rotary shaker at 25°C for five days. Mycelium was harvested over a piece of cheese-cloth (Merck-Millipore, Darmstadt Germany).

For production of spores for infection assays, *M. oryzae* isolate Guy11 was grown on rice flour agar for spore production [90]. A suspension of fungal conidiospores was prepared at a density of $2x10^5$ spores/ml and spotted on detached leaves of the japonica rice variety Saraceltik grown for 3 weeks as described [91]. Infected leaf samples were harvested 16, 24, 48 and 72 hours post inoculation (hpi).

RNA extraction and qRT–PCR analysis

RNA extraction and reverse transcription was performed as described [92]. Quantitative PCR were performed with a LightCycler 480 instrument (Roche, Basel, Switzerland) using LC 480 SYBR Green I Master Mix (Roche) and the primers listed in the S4 Table. Amplification was performed as follows: 95°C for 10 min; 40 cycles of 95°C for 15 s, 60°C for 20s and 72°C for 30 s; then 95°C for 5 min and 40°C for 30 s. Data were analyzed using the delta-delta Ct method and applying the formula 2-ΔCT, where ΔCT is the difference in threshold cycle (CT) between the gene of interest and the housekeeping gene *Actin* (*MGG_03982*) used as a constitutively expressed reference gene. For each condition, three biological replicates were analyzed.

Supporting Information

S1 Table. NMR experiments acquired for structure calculations and chemical shift assignments.
(PDF)

S2 Table. DALI statistics for structural alignment of AVR-Pia, AVR1-C039, AvrPiz-t and ToxB.
(PDF)

S3 Table. MAX-effector candidates identified by psi-Blast in the genomes of the *M. oyzae* isolates 70–15 and TH16 and the *M. grisea* isolate BR29.
(PDF)

S4 Table. Primers used for qRT-PCR.
(PDF)

S1 Fig. Gel filtration profile and SDS-PAGE analysis of purified AVR-Pia (A) and AVR1-CO39 (B) proteins.
(PDF)

S2 Fig. 15N Relaxation data at 500 MHz and 305K for AVR-Pia (panels A, B and C) and AVR1-CO39 (panels D, E and F).
(PDF)

S3 Fig. Backbone sequential and medium range NOEs observed for (A) AVR-Pia and (B) AVR1-CO39. The line width is proportional to the NOE intensity. The dots (•) indicate slow exchange NH observed in 2D-NOESY in D2O. Grey arrows indicate the ß-strands determined from the structure analysis (vide infra).
(PDF)

S4 Fig. Solution structures of (A) AVR-Pia and (B) AVR1-CO39. Superposition of the backbone atoms of the 20 lowest energy conformers used to calculate the final structures. Only mature chains are shown, from residues Ala20 and Trp23 for AVR-Pia and AVR1-CO39, respectively.
(PDF)

S5 Fig. Structure-guided alignment and diversity of MAX-effector homologs identified by Blast. A) Homologs of AVR1-CO39, AvrPiz-t and AVR-Pia identified by Blast in *M. oryzae* and *M. grisea* genomes and ToxB homologs identified by Psi-Blast in the GeneBank database were aligned to the structural alignment of mature ToxB, AVR1-CO39, AvrPiz-t and AVR-Pia. (B) A diversity tree was constructed by the neighbor-joining method using the alignment in (A). It highlights the high diversity of MAX-effector homologs. Branch supports are based on 1000 bootstraps and horizontal branch length reflects sequence divergence. Accession numbers of non-Magnaporthe sequences were completed by a 2 letter identifier for the species: BO for *Bipolaris oryzae*, CF is for *Colletotrichum fioriniae*, CH for *C. higgensianum*, CG for *C. gloeosporioides*, CO for *C. orbiculare*, LM for *Lepthosphaeria maculans*, PT for *Pyrenophora tritici-repentis* and PB for *Pyrenophora bromi*.
(PDF)

S6 Fig. MAX-effector homologs identified by a high stringency HMM search. (A) Histogram showing the numbers of MAX-effectors identified by an HMM pattern search in a non-redundant database comprising the small secreted proteins of 25 ascomycete fungi and of 8 additional *M. oryzae* and one *M. gisea* isolate. (B) MAX-effectors were aligned to the structural alignment of mature ToxB, AVR1-CO39, AvrPiz-t and AVR-Pia and gaps were removed. (C) A diversity tree was constructed by the neighbor-joining method using the alignment in (B). Branch supports are based on 1000 bootstraps and horizontal branch length reflects sequence divergence. Accession numbers of non-Magnaporthe sequences were completed by a 2 letter identifier for the species: BO for *Bipolaris oryzae*, CF for *Colletotrichum fioriniae*, CH for *C. higgensianum*, CG for *C. gloeosporioides*, CO for *C. orbiculare*, LM for *Lepthosphaeria maculans*, PT for *Pyrenophora tritici-repentis* PB for *Pyrenophora bromi* and ZT for *Zymoseptoria tritici*.
(PDF)

S7 Fig. MAX-effector homologs identified in the UniRef90 database by a low stringency HMM search. (A) Histogram showing the numbers of MAX-effectors identified by an HMM

pattern search in a non-redundant UniRef90 database. (B) MAX-effectors identified by HMM pattern search were aligned to the structural alignment of mature ToxB, AVR1-CO39, AvrPiz-t and AVR-Pia. (C) A diversity tree was constructed by the neighbor-joining method using the alignment in (B). This highlights the high diversity of MAX-effector homologs. Branch supports are based on 1000 bootstraps and horizontal branch length reflects sequence divergence. Accession numbers contain the following information on the species: MAGGR, MAGO7, MAGOP and MAGOR are from *M. oryzae*, COLGC and COLGN from *C. gloeosporioides*, COLHI from *Colletotrichum higginsianum*, 9PEZI from *C. fioriniae* and COLOR *from Colletotrichum orbiculare*, 9PLEO from*P. tritici-repentis* or *P. bromi*, ARTOA from *Arthrobotrys oligospora*, COCMI from *Bipolaris oryzae*, LEPMJ from *Leptosphaeria maculans*, MYCGM from *Zymoseptoria tritici*, 9PSED from *Pseudomonas sp. StFLB209* and SOLLC from *Solanum lycopersicum*.
(PDF)

S8 Fig. MAX-effector homologs identified by a low stringency HMM search in the fungal genomes database. (A) MAX-effectors identified by an HMM pattern search in a redundant database comprising the small secreted proteins of 25 ascomycete fungi, 8 additional *M. oryzae* and one *M. gisea* isolate were aligned to the structural alignment of mature ToxB, AVR1-CO39, AvrPiz-t and AVR-Pia and gaps were removed. (B) A diversity tree was constructed by the neighbor-joining method using the alignment in (A). Branch supports are based on 1000 bootstraps and horizontal branch length reflects sequence divergence. Accession numbers of non-Magnaporthe sequences were completed by a 2 letter identifier for the species: BO for *Bipolaris oryzae*, CF for *Colletotrichum fioriniae*, CH for *C. higgensianum*, CG for *C. gloeosporioides*, CO for *C. orbiculare*, GF for *Fusarium fujcuroi*, LM for *Lepthosphaeria maculans*, PT for *Pyrenophora tritici-repentis*, PB for *Pyrenophora bromi* and ZT for *Zymoseptoria tritici*.
(PDF)

S9 Fig. Expression of *M. oryzae* MAX-effector candidates and marker genes during rice infection and in *in vitro* grown mycelium. mRNA levels of *M. oryzae* genes coding for MAX-effectors (A, B and C) and marker genes (D) was determined by q-RT-PCR in rice leaf samples harvested 16, 24, 48 or 72 h after inoculation and mycelium grown liquid medium for 72 hours. (A) Infection specific MAX-effectors identified in the HMM search, (B) infection specific MAX-effectors identified in the Psi Blast search but nor in the HMM search, (C) constitutively expressed MAX-effectors identified in the HMM search and (D) marker genes for appressorium and very early infection (*ORF3* of the ACE1 cluster, *MGG_08381*), biotrophic infection (*BAS3*, *MGG_11610*), late infection (*MGG_01147*), constitutive expression (*EF1α*, *MGG_03641*). Relative expression levels were calculated by using expression of a constitutively expressed *Actin* (*MGG_03982*) as a reference. Mean values and standard deviation were calculated from three independent biological samples.The analyzed genes, were in most cases not or extremely weakly expressed in the mycelium. For genes with significant expression in the mycelium (ratio gene versus actine $>$ 0,01) a T-test was performed to determine if in planta expression was significantly different from expression in the mycelium. In these cases (MGG_11967, MGG_14793, MGG_15207, MGG_17266, MGG_18019, M_TH16_00000541, M_TH16_00040131, M_TH16_00079081, M_TH16_00104561, M_TH16_00120731, M_TH16_00124981), a star or two stars (* or **) mark conditions where the expression was different from expression in the mycelium at respectively $p<0,05$ or $p<0,005$.
(PDF)

S10 Fig. Prediction of the secondary structure of M. oryzae MAX effectors. The secondary structures of the MAX-effectors from the 70–15, TH16 and BR29 genomes was predicted with SSPRO5 [93].The predictions are shown at the bottom of the figure and are aligned onto the corresponding primary sequence alignment shown at the top of the figure. Sequence identifiers for the secondary structure predictions are suffixed with ".2d.SSPRO5". Blue"H", red "E" and yellow "C" correspond respectively to helix, extended sheet and coil predictions. The sequences of the 4 MAX effectors with experimentally determined structures are displayed at the top of the multiple sequence alignment and, for clarity, the alignment positions corresponding to shared gaps in the known structures were removed.
(TIF)

Acknowledgments

Véronique Chalvon is acknowledged for technical assistance. Martin Cohen-Gonsaud, Christian Roumestand, Jean-Benoit Morel and Stella Césari are acknowledged for fruitful discussions.

Author Contributions

Conceived and designed the experiments: TK AP JG. Performed the experiments: DOV KdG JG AP. Analyzed the data: JG DOV KdG EF TK AP. Contributed reagents/materials/analysis tools: EF. Wrote the paper: TK AP JG.

References

1. Hogenhout S a, Van der Hoorn R a L, Terauchi R, Kamoun S. Emerging concepts in effector biology of plant-associated organisms. Mol Plant Microbe Interact. 2009; 22: 115–122. doi: 10.1094/MPMI-22-2-0115 PMID: 19132864
2. Jones JDG, Dangl JL. The plant immune system. Nature. 2006; 444: 323–9. PMID: 17108957
3. Doehlemann G, Requena N, Schaefer P, Brunner F, Connell RO, Parker JE. Reprogramming of plant cells by filamentous plant-colonizing microbes. 2014; 803–814.
4. Stergiopoulos I, de Wit PJGM. Fungal effector proteins. Annu Rev Phytopathol. 2009; 47: 233–63. doi: 10.1146/annurev.phyto.112408.132637 PMID: 19400631
5. Presti L Lo, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, et al. Fungal Effectors and Plant Susceptibility. Annu Rev Plant Biol. 2015; 66: 513–545. doi: 10.1146/annurev-arplant-043014-114623 PMID: 25923844
6. Giraldo MC, Valent B. Filamentous plant pathogen effectors in action. Nat Rev Microbiol. Nature Publishing Group; 2013; 11: 800–14. doi: 10.1038/nrmicro3119 PMID: 24129511
7. Bozkurt TO, Schornack S, Banfield MJ, Kamoun S. Oomycetes, effectors, and all that jazz. Curr Opin Plant Biol. Elsevier Ltd; 2012; 1–10.
8. Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, et al. Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. Nature. 2009; 461: 393–8. doi: 10.1038/nature08358 PMID: 19741609
9. Jiang RHY, Tripathy S, Govers F, Tyler BM. RXLR effector reservoir in two Phytophthora species is dominated by a single rapidly evolving superfamily with more than 700 members. Proc Natl Acad Sci U S A. 2008; 105: 4874–4879. doi: 10.1073/pnas.0709303105 PMID: 18344324
10. Duplessis S, Cuomo C a, Lin Y-C, Aerts A, Tisserant E, Veneault-Fourrey C, et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc Natl Acad Sci U S A. 2011; 108: 9166–71. doi: 10.1073/pnas.1019315108 PMID: 21536894
11. Pedersen C, Themaat EVL van, McGuffin LJ, Abbott JC, Burgis TA, Barton G, et al. Structure and evolution of barley powdery mildew effector candidates. BMC Genomics. 2012; 13: 694. doi: 10.1186/1471-2164-13-694 PMID: 23231440
12. Spanu PD, Abbott JC, Amselem J, Burgis T a, Soanes DM, Stüber K, et al. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science. 2010; 330: 1543–6. doi: 10.1126/science.1194573 PMID: 21148392

13. Hacquard S, Joly DL, Lin Y-C, Tisserant E, Feau N, Delaruelle C, et al. A comprehensive analysis of genes encoding small secreted proteins identifies candidate effectors in Melampsora larici-populina (poplar leaf rust). Mol Plant Microbe Interact. 2012; 25: 279–93. doi: 10.1094/MPMI-09-11-0238 PMID: 22046958

14. Dean R, Kan JANALVAN, Pretorius ZA, Hammond-kosack KIME, Pietro ADI, Spanu PD, et al. The Top 10 fungal pathogens in molecular plant pathology. 2012;13: 414–430.

15. Gurr S, Samalova M, Fisher M. The rise and rise of emerging infectious fungi challenges food security and ecosystem health. Fungal Biol Rev. 2011; 25: 181–188.

16. Skamnioti P, Gurr SJ. Against the grain: safeguarding rice from rice blast disease. Trends Biotechnol. 2009; 27: 141–50. doi: 10.1016/j.tibtech.2008.12.002 PMID: 19187990

17. Galhano R, Talbot NJ. The biology of blast: Understanding how Magnaporthe oryzae invades rice plants. Fungal Biol Rev. Elsevier Ltd; 2011; 25: 61–67.

18. Liu J, Wang X, Mitchell T, Hu Y, Liu X, Dai L, et al. Recent progress and understanding of the molecular mechanisms of the rice-Magnaporthe oryzae interaction. Mol Plant Pathol. 2010; 11: 419–427. doi: 10.1111/j.1364-3703.2009.00607.x PMID: 20447289

19. Mentlak TA, Talbot NJ, Kroj T. Effector Translocation and Delivery by the Rice Blast Fungus Magnaporthe oryzae. In: Francisrtin, Kamoun S, editors. Effectors in Plant–Microbe Interactions. Wiley-Blackwell; 2011. pp. 219–241.

20. Valent B, Khang CH. Recent advances in rice blast effector research. Curr Opin Plant Biol. Elsevier Ltd; 2010; 13: 434–41. doi: 10.1016/j.pbi.2010.04.012 PMID: 20627803

21. Park C-H, Chen S, Shirsekar G, Zhou B, Khang CH, Songkumarn P, et al. The Magnaporthe oryzae Effector AvrPiz-t Targets the RING E3 Ubiquitin Ligase APIP6 to Suppress Pathogen-Associated Molecular Pattern-Triggered Immunity in Rice. Plant Cell. 2012.

22. Mentlak TA, Kombrink A, Shinya T, Ryder LS, Otomo I, Saitoh H, et al. Effector-mediated suppression of chitin-triggered immunity by magnaporthe oryzae is necessary for rice blast disease. Plant Cell. 2012; 24: 322–35. doi: 10.1105/tpc.111.092957 PMID: 22267486

23. Mosquera G, Giraldo MC, Khang CH, Coughlan S, Valent B. Interaction transcriptome analysis identifies Magnaporthe oryzae BAS1-4 as Biotrophy-associated secreted proteins in rice blast disease. Plant Cell. 2009; 21: 1273–90. doi: 10.1105/tpc.107.055228 PMID: 19357089

24. Saitoh H, Fujisawa S, Mitsuoka C, Ito A, Hirabuchi A, Ikeda K, et al. Large-scale gene disruption in Magnaporthe oryzae identifies MC69, a secreted protein required for infection by monocot and dicot fungal pathogens. PLoS Pathog. 2012; 8: e1002711. doi: 10.1371/journal.ppat.1002711 PMID: 22589729

25. Soanes DM, Alam I, Cornell M, Wong HM, Hedeler C, Paton NW, et al. Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis. PLoS One. 2008; 3: e2300. doi: 10.1371/journal.pone.0002300 PMID: 18523684

26. Yoshida KK, Saitoh H, Fujisawa S, Kanzaki H, Matsumura H, Tosa Y, et al. Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen Magnaporthe oryzae. Plant Cell. 2009; 21: 1573–91. doi: 10.1105/tpc.109.066324 PMID: 19454732

27. Chen X, Coram T, Huang X, Wang M, Dolezal A. Understanding Molecular Mechanisms of Durable and Non-durable Resistance to Stripe Rust in Wheat Using a Transcriptomics Approach. Curr Genomics. 2013; 14: 111–126. doi: 10.2174/1389202911314020004 PMID: 24082821

28. Kim S, Hu J, Oh Y, Park J, Choi J, Lee Y-H, et al. Combining ChIP-chip and expression profiling to model the MoCRZ1 mediated circuit for Ca/calcineurin signaling in the rice blast fungus. PLoS Pathog. 2010; 6: e1000909. doi: 10.1371/journal.ppat.1000909 PMID: 20502632

29. Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet. Nature Publishing Group; 2010; 11: 539–48. doi: 10.1038/nrg2812 PMID: 20585331

30. Miki H, Matsui K, Kito H, Otsuka K, Ashizawa T, Yasuda N, et al. Molecular cloning and characterization of the AVR-Pia locus from a Japanese field isolate of Magnaporthe oryzae. 2009;10: 361–374.

31. Orbach MJ, Farrall L, Sweigard J a, Chumley FG, Valent B. A telomeric avirulence gene determines efficacy for the rice blast resistance gene Pi-ta. Plant Cell. 2000; 12: 2019–32. PMID: 11090206

32. Ribot C, Césari S, Abidi I, Chalvon V, Bournaud C, Vallet J, et al. The Magnaporthe oryzae effector AVR1-CO39 is translocated into rice cells independently of a fungal-derived machinery. Plant J. 2012; 74: 1–12.

33. Sweigard J a, Carroll a M, Kang S, Farrall L, Chumley FG, Valent B. Identification, cloning, and characterization of PWL2, a gene for host species specificity in the rice blast fungus. Plant Cell. 1995; 7: 1221–33. PMID: 7549480

34. Wu J, Kou Y, Bao J, Li Y, Tang M, Zhu X, et al. Comparative genomics identifies the Magnaporthe oryzae avirulence effector AvrPi9 that triggers Pi9 -mediated blast resistance in rice. 2015;

35. Li W, Wang B, Wu J, Lu G, Hu Y, Zhang X, et al. The Magnaporthe oryzae avirulence gene AvrPiz-t encodes a predicted secreted protein that triggers the immunity in rice mediated by the blast resistance gene Piz-t. Mol Plant Microbe Interact. 2009; 22: 411–20. doi: 10.1094/MPMI-22-4-0411 PMID: 19271956

36. Ashikawa I, Hayashi N, Yamane H, Kanamori H, Wu J, Matsumoto T, et al. Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer Pikm-specific rice blast resistance. Genetics. 2008; 180: 2267–76. doi: 10.1534/genetics.108.095034 PMID: 18940787

37. Cesari S, Thilliez G, Ribot C, Chalvon V, Michel C, Jauneau A, et al. The rice resistance protein pair RGA4/RGA5 recognizes the Magnaporthe oryzae effectors AVR-Pia and AVR1-CO39 by direct binding. Plant Cell. 2013; 25: 1463–81. doi: 10.1105/tpc.112.107201 PMID: 23548743

38. Qu S, Liu G, Zhou B, Bellizzi M, Zeng L, Dai L, et al. The broad-spectrum blast resistance gene Pi9 encodes a nucleotide-binding site-leucine-rich repeat protein and is a member of a multigene family in rice. Genetics. 2006; 172: 1901–14. PMID: 16387888

39. Okuyama Y, Kanzaki H, Abe A, Yoshida K, Tamiru M, Saitoh H, et al. A multifaceted genomics approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS-LRR protein genes. Plant J. 2011; 66: 467–79. doi: 10.1111/j.1365-313X.2011.04502.x PMID: 21251109

40. Bryan GT, Wu KS, Farrall L, Jia Y, Hershey HP, McAdams S a, et al. tA single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene Pi-ta. Plant Cell. 2000; 12: 2033–46. PMID: 11090207

41. Zhou B, Qu S, Liu G, Dolan M, Sakai H, Lu G, et al. The eight amino-acid differences within three leucine-rich repeats between Pi2 and Piz-t resistance proteins determine the resistance specificity to Magnaporthe grisea. Mol Plant Microbe Interact. 2006; 19: 1216–28. PMID: 17073304

42. Zhang Z-M, Zhang X, Zhou Z-R, Hu H-Y, Liu M, Zhou B, et al. Solution structure of the Magnaporthe oryzae avirulence protein AvrPiz-t. J Biomol NMR. 2013; 55: 219–23. doi: 10.1007/s10858-012-9695-5 PMID: 23334361

43. Nyarko A, Singarapu KK, Figueroa M, Manning V a., Pandelova I, Wolpert TJ, et al. Solution NMR Structures of Pyrenophora tritici-repentis ToxB and Its Inactive Homolog Reveal Potential Determinants of Toxin Activity. J Biol Chem. 2014; 289: 25946–25956. doi: 10.1074/jbc.M114.569103 PMID: 25063993

44. Barthe P, Ropars V, Roumestand C. DYNAMOF: a program for the dynamics analysis of relaxation data obtained at multiple magnetic fields. Comptes Rendus Chim. 2006; 9: 503–513.

45. Holm L, Rosenström P. Dali server: Conservation mapping in 3D. Nucleic Acids Res. 2010; 38: 545–549.

46. Ciuffetti LM, Manning V a., Pandelova I, Betts MF, Martinez JP. Host-selective toxins, Ptr ToxA and Ptr ToxB, as necrotrophic effectors in the Pyrenophora tritici-repentis-wheat interaction. New Phytol. 2010; 187: 911–919. doi: 10.1111/j.1469-8137.2010.03362.x PMID: 20646221

47. Chiapello H, Mallet L, Guérin C, Aguileta G, Amselem J, Kroj T, et al. Deciphering genome content and evolutionary realtionships of isolates from the fungus Magnaporthe oryzae attacking different hosts. Genome Biol Evol. 2015;in press.

48. Klaubauf S, Tharreau D, Fournier E, Groenewald JZ, Crous PW, de Vries RP, et al. Resolving the polyphyletic nature of Pyricularia (Pyriculariaceae). Stud Mycol. ELSEVIER B.V; 2014; 79: 85–120. doi: 10.1016/j.simyco.2014.09.004 PMID: 25492987

49. Césari S, Kanzaki H, Fujiwara T, Bernoux M, Chalvon V, Kawano Y, et al. The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease resistance. EMBO J. 2014; 33: 1941–1959. doi: 10.15252/embj.201487923 PMID: 25024433

50. Césari S, Bernoux M, Moncuquet P, Kroj T, Dodds PN. A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. Front Plant Sci. 2014; 5: 606. doi: 10.3389/fpls.2014.00606 PMID: 25506347

51. Pedersen C, Themaat V, Ver E, McGuffin L, Abbott JC, Burgis TA, et al. Structure and evolution of barley powdery mildew effector candidates. BMC Genomics. 2012

52. Chuma I, Isobe C, Hotta Y, Ibaragi K, Futamata N, Kusaba M, et al. Multiple translocation of the AVR-Pita effector gene among chromosomes of the rice blast fungus Magnaporthe oryzae and related species. PLoS Pathog. 2011; 7: e1002147. doi: 10.1371/journal.ppat.1002147 PMID: 21829350

53. Rouxel T, Grandaubert J, Hane JK, Hoede C, van de Wouw AP, Couloux A, et al. Effector diversification within compartments of the Leptosphaeria maculans genome affected by Repeat-Induced Point mutations. Nat Commun. 2011; 2: 202. doi: 10.1038/ncomms1189 PMID: 21326234

54. Ma L-J, van der Does HC, Borkovich K a, Coleman JJ, Daboussi M-J, Di Pietro A, et al. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature. 2010; 464: 367–73. doi: 10.1038/nature08850 PMID: 20237561

55. Jonge R De, Bolton MD, Kombrink A, De Jonge R, Bolton MD, Kombrink A, et al. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Res. 2013; 23: 1271–1282. doi: 10.1101/gr.152660.112 PMID: 23685541

56. Win J, Krasileva K V., Kamoun S, Shirasu K, Staskawicz BJ, Banfield MJ. Sequence Divergent RXLR Effectors Share a Structural Fold Conserved across Plant Pathogenic Oomycete Species. Heitman J, editor. PLoS Pathog. 2012; 8: e1002400. doi: 10.1371/journal.ppat.1002400 PMID: 22253591

57. Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, et al. Signatures of adaptation to obligate biotrophy in the Hyaloperonospora arabidopsidis genome. Science. 2010; 330: 1549–1551. doi: 10.1126/science.1195203 PMID: 21148394

58. Raffaele S, Farrer R a, Cano LM, Studholme DJ, MacLean D, Thines M, et al. Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science. 2010; 330: 1540–3. doi: 10.1126/science.1193070 PMID: 21148391

59. Dou D, Kale SD, Wang X, Chen Y, Wang Q, Wang X, et al. Conserved C-terminal motifs required for avirulence and suppression of cell death by Phytophthora sojae effector Avr1b. Plant Cell. 2008; 20: 1118–33. doi: 10.1105/tpc.107.057067 PMID: 18390593

60. Boutemy LS, King SRF, Win J, Hughes RK, Clarke T a, Blumenschein TM a, et al. Structures of Phytophthora RXLR effector proteins: a conserved but adaptable fold underpins functional diversity. J Biol Chem. 2011; 286: 35834–42. doi: 10.1074/jbc.M111.262303 PMID: 21813644

61. Chou S, Krasileva K V, Holton JM, Steinbrenner AD, Alber T, Staskawicz BJ. Hyaloperonospora arabidopsidis ATR1 effector is a repeat protein with distributed recognition surfaces. Proc Natl Acad Sci U S A. 2011; 108: 13323–8. doi: 10.1073/pnas.1109791108 PMID: 21788488

62. Yaeno T, Li H, Chaparro-Garcia A, Schornack S, Koshiba S, Watanabe S, et al. Phosphatidylinositol monophosphate-binding interface in the oomycete RXLR effector AVR3a is required for its stability in host cells to modulate plant immunity. Proc Natl Acad Sci U S A. 2011; 108: 14682–7. doi: 10.1073/pnas.1106002108 PMID: 21821794

63. Bayry J, Aimanianda V, Guijarro JI, Sunde M, Latgé JP. Hydrophobins-unique fungal proteins. PLoS Pathog. 2012; 8: 6–9.

64. Kubicek CP, Baker S, Gamauf C, Kenerley CM, Druzhinina IS. Purifying selection and birth-and-death evolution in the class II hydrophobin gene families of the ascomycete Trichoderma/Hypocrea. BMC Evol Biol. 2008; 8: 4. doi: 10.1186/1471-2148-8-4 PMID: 18186925

65. Wosten H a. H YDROPHOBINS: Multipurpose Proteins. Annu Rev Microbiol. 2001; 55: 625–46. PMID: 11544369

66. Kwan a HY, Winefield RD, Sunde M, Matthews JM, Haverkamp RG, Templeton MD, et al. Structural basis for rodlet assembly in fungal hydrophobins. Proc Natl Acad Sci U S A. 2006; 103: 3621–3626. PMID: 16537446

67. Stergiopoulos I, Kourmpetis Y a I, Slot JC, Bakker FT, De Wit PJGM, Rokas A. In silico characterization and molecular evolutionary analysis of a novel superfamily of fungal effector proteins. Mol Biol Evol. 2012; 29: 3371–84. doi: 10.1093/molbev/mss143 PMID: 22628532

68. Ve T, Williams SJ, Catanzariti A-M, Rafiqi M, Rahman M, Ellis JG, et al. Structures of the flax-rust effector AvrM reveal insights into the molecular basis of plant-cell entry and effector-triggered immunity. Proc Natl Acad Sci U S A. 2013

69. Wang C-IAC-I a, Guncar G, Forwood JK, Teh T, Catanzariti A-MA-M, Lawrence GJ, et al. Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. Plant Cell. 2007; 19: 2898–912. PMID: 17873095

70. Sarma GN, Manning VA, Ciuffetti LM, Karplus PA. Structure of Ptr ToxA: An RGD-Containing Host-Selective Toxin from Pyrenophora tritici-repentis. 2005;17: 3190–3202.

71. Studier FW. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif. 2005; 41: 207–234. PMID: 15915565

72. Habeeb AFSA. [37] Reaction of protein sulfhydryl groups with Ellman's reagent. In: Hirs C. H. W. SNT, editor. Methods in Enzymology. Academic Press; 1972. pp. 457–464. doi: 10.1016/S0076-6879(72)25041-8

73. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, et al. 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J Biomol NMR. 1995; 6: 135–140. PMID: 8589602

74. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, et al. The CCPN data model for NMR spectroscopy: Development of a software pipeline. Proteins Struct Funct Bioinforma. 2005; 59: 687–696.

75. Carr HY, Purcell EM. Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. Phys Rev. 1954; 94: 630–632.

76. Meiboom S, Gill D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. Rev Sci Instrum. 1958; 29: 688.

77. Kay LE, Torchia DA, Bax A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry. 1989; 28: 8972–8979. PMID: 2690953

78. Güntert P. Automated NMR structure calculation with CYANA. Methods Mol Biol. 2004; 278: 353–378. PMID: 15318003

79. Brunger AT. Version 1.2 of the Crystallography and NMR system. Nat Protoc. 2007; 2: 2728–2733. PMID: 18007608

80. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR. 2009; 44: 213–223. doi: 10.1007/s10858-009-9333-z PMID: 19548092

81. Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk CAEM, Nabuurs SB, et al. RECOORD: A recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. Proteins Struct Funct Bioinforma. 2005; 59: 662–672.

82. Laskowski RA, Moss DS, Thornton JM. Main-chain bond lengths and bond angles in protein structures. J Mol Biol. 1993; 231: 1049–1067. PMID: 8515464

83. Suzek BE, Huang H, McGarvey P, Mazumder R, Wu CH. UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics. 2007; 23: 1282–1288. PMID: 17379688

84. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011; 8: 785–786. doi: 10.1038/nmeth.1701 PMID: 21959131

85. Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Res. 2005; 33: 2302–2309. PMID: 15849316

86. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 2011; 39: W29–W37. doi: 10.1093/nar/gkr367 PMID: 21593126

87. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32: 1792–1797. PMID: 15034147

88. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004; 14: 1188–1190. PMID: 15173120

89. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30: 2725–2729. doi: 10.1093/molbev/mst197 PMID: 24132122

90. Berruyer R, Adreit H, Milazzo J, Gaillard S, Berger a., Dioh W, et al. Identification and fine mapping of Pi33, the rice resistance gene corresponding to the Magnaporthe grisea avirulence gene ACE1. Theor Appl Genet. 2003; 107: 1139–1147. PMID: 12838393

91. Faivre-Rampant O, Thomas J, Allègre M, Morel J-B, Tharreau D, Nottéghem J-L, et al. Characterization of the model system rice-Magnaporthe for the study of nonhost resistance in cereals. New Phytol. 2008; 180: 899–910. doi: 10.1111/j.1469-8137.2008.02621.x PMID: 19138233

92. Delteil A, Blein M, Faivre-rampant O, Guellim A, Estevan J, Hirsch J, et al. Building a mutant resource for the study of disease resistance in rice reveals the pivotal role of several genes involved in defence. 2012;13: 72–82.

93. Magnan CN, Baldi P. SSpro/ACCpro 5: almost perfect prediction of protein secondary structure and relative solvent accessibility using profiles, machine learning and structural similarity. Bioinformatics. 2014; 30: 2592–2597. doi: 10.1093/bioinformatics/btu352 PMID: 24860169

Supplemental Table S1. *NMR experiments acquired for structure calculations and chemical shift assignments*

a) AVR-Pia

		Size			Sweep width (ppm)						
Experiments	nuclei	F3	F2	F1	F3	F2	F1	Mix(ms)	NS	D1(s)	B_o(MHz)
^{15}N-HSQC	^{1}H, ^{15}N	1024	128	-	-	-	-	-	16	1	700
HNCO(*)	^{1}H,^{15}N,^{13}C	1024	64	64	12.98	36	17.67	-	8	0.001	500
HNCA(*)	^{1}H,^{15}N,^{13}C	1024	64	80	12.98	36	31.81	-	16	0.1	500
HNCOCACB(*)	^{1}H,^{15}N,^{13}C	1024	64	128	12.98	36	75.02	-	32	0.2	500
HNCACO(*)	^{1}H,^{15}N,^{13}C	1024	64	80	12.98	36	17.67	-	64	0.001	500
HNCACB(*)	^{1}H,^{15}N,^{13}C	1024	64	128	12.98	36	75.02	-	32	0.2	500
CON(**)	^{13}C, ^{15}N	1024	200	-	30	36	-	-	4	1	500
CACO(**)	^{13}C, ^{13}C	1024	256	-	30	50.32	-	-	8	1	500
^{15}N-NOESY-HSQC	^{1}H,^{15}N, ^{1}H	1500	76	360	14	35	14	120	8	1	700
^{15}N-TOCSY-HSQC	^{1}H,^{15}N, ^{1}H	1500	70	320	14	35	14	56	8	1	700
NOESY (D_2O)	^{1}H, ^{1}H	2048 4096	512 512	-	15.00 12.01	15.00 12.01	- -	100 150	128 128	1 1	700 700
TOCSY (D_2O)	^{1}H, ^{1}H	4096	512	-	12.01	12.01	-	39.6	64	1	700
COSY-dqf (D_2O)	^{1}H, ^{1}H	4096	800		12.01	12.01	-	-	64	1	700
^{1}H,^{15}N NOE	^{1}H, ^{15}N	1024	128	-	14	35	-	Sat. 3s	64	6	500
R_1	^{1}H, ^{15}N	1024	128	-	14	35	-		16	2.5	500
R_2	^{1}H, ^{15}N	1024	128	-	14	35	-		16	2.5	500

b) AVR1-CO39

		Size			Sweep width (ppm)						
Experiments	nuclei	F3	F2	F1	F3	F2	F1	Mix(ms)	NS	D1(s)	B_o(MHz)
^{15}N-HSQC	^{1}H, ^{15}N	1500	160	-	12.02	28	-	-	16	1	700
1^{3}C-HSQC	^{1}H, ^{13}C	2048	256	-	15.00	80	-	-	512	1.5	700
HNCO(*)	^{1}H,^{15}N,^{13}C	1024	64	64	12.02	28	17.67	-	8	0.001	500
HNCA(*)	^{1}H,^{15}N,^{13}C	1024	64	80	18.02	28	31.81	-	16	0.1	500
HNCOCACB(*)	^{1}H,^{15}N,^{13}C	1024	64	128	18.02	28	75.02	-	32	0.2	500
HNCACO(*)	^{1}H,^{15}N,^{13}C	1024	64	80	18.02	28	17.67	-	64	0.001	500
HNCACB(*)	^{1}H,^{15}N,^{13}C	1024	64	128	18.02	28	75.02	-	32	0.2	500
CON(**)	^{13}C, ^{15}N	1024	200	-	30	28	-	-	4	1	500
CACO(**)	^{13}C, ^{13}C	1024	256	-	30	50.32	-	-	2	1	500
^{15}N-NOESY-HSQC	^{1}H,^{15}N, ^{1}H	1500	76	360	12	28	12	120	8	1	700
^{15}N-TOCSY-HSQC	^{1}H,^{15}N, ^{1}H	1024	72	320	14	35	14	56	8	1	700
NOESY (D_2O)	^{1}H, ^{1}H	2048 2048	512 512	-	15.00 15.00	15.00 15.00	- -	100 200	64 32	1 1	700 700
TOCSY	^{1}H, ^{1}H	2048	512	-	15.00	15.00	-	40	32	1	700
COSY-dqf	^{1}H, ^{1}H	2048	800		15.00	15.00	-	-	32	1	700
^{1}H,^{15}N NOE	^{1}H, ^{15}N	1024	128	-	14	35	-	Sat. 3s	64	6	500
R_1	^{1}H, ^{15}N	1024	128	-	14	35	-		16	2.5	500
R_2	^{1}H, ^{15}N	1024	128	-	14	35	-		16	2.5	500

Experiments were recorded using the TOPSPIN Library (v. 2.1) at 305 K.
(*) Best pulse sequences
(**)IPAP scheme for virtual decoupling

Supporting information

S2 Table. DALI Statistics for structural alignment of AVR-Pia, AVR1-CO39, AVRPiz-t and ToxB.

RMSD (Å) / Z-score	AVR-Pia	AVR1-CO39	AVRPiz-t	ToxB
AVR-Pia		2.3	2.8	2.2
AVR1-CO39	2.9		3.0	2.2
AVRPiz-t	3.3	3.1		2.2
ToxB	4.4	4.2	5.4	

The rmsd (upper right) and the Z-score (lower left) are from the pairwise superposition. The low RMSDs and the high Z-scores indicate that the ToxB(b) structure is the closest to the other three structures.

S3 Table MAX-effector candidates identified by psi-Blast in the genomes of the *M. oyzae* isolates 7015 and TH16 and the *M. grisea* isolate BR29.

geneID	localization	AVR-Pia*	AVR1-CO39*	ToxB*
M_BR29_EuGene_00004921	scaffold00001	9,0E-12	-	6,0E-12
M_BR29_EuGene_00041131	scaffold00007	-	9,0E-06	-
M_BR29_EuGene_00043011	scaffold00008	-	9,0E-06	-
M_BR29_EuGene_00060181	scaffold00013	-	4,0E-08	-
M_BR29_EuGene_00081821	scaffold00023	-	3,0E-07	-
M_BR29_EuGene_00082031	scaffold00023	-	1,0E-05	-
M_BR29_EuGene_00085071	scaffold00025	-	-	3,0E-15
M_BR29_EuGene_00087671	scaffold00027	4,0E-13	4,0E-12	8,0E-11
M_BR29_EuGene_00088411	scaffold00027	-	-	5,0E-12
M_BR29_EuGene_00091361	scaffold00030	3,0E-12	-	-
M_BR29_EuGene_00091681	scaffold00031	-	2,0E-05	-
M_BR29_EuGene_00095641	scaffold00035	-	3,0E-13	5,0E-13
M_BR29_EuGene_00106461	scaffold00049	7,0E-23	-	-
M_BR29_EuGene_00107481	scaffold00051	-	2,0E-06	-
M_BR29_EuGene_00112111	scaffold00059	-	6,0E-09	-
M_BR29_EuGene_00113041	scaffold00061	-	7,0E-08	-
M_BR29_EuGene_00118801	scaffold00076	-	2,0E-05	-
M_BR29_EuGene_00119491	scaffold00079	-	3,0E-22	-
M_BR29_EuGene_00119511	scaffold00079	-	9,0E-17	-
M_BR29_EuGene_00121691	scaffold00087	-	1,0E-13	-
M_BR29_EuGene_00125811	scaffold00145	-	7,0E-08	-
M_BR29_EuGene_00126081	scaffold00163	-	-	2,0E-20
MGG_00821	Chromosome_8.5	-	6,0E-14	-
MGG_04384	Chromosome_8.2	-	2,0E-04	-
MGG_08482	Chromosome_8.4	-	1,0E-06	-
MGG_08944	Chromosome_8.2	-	-	5,0E-09
MGG_10120	Chromosome_8.4	-	5,0E-04	-
MGG_14793	Chromosome_8.2	4,0E-19	3,0E-06	9,0E-15
MGG_14834	Chromosome_8.4	-	2,0E-05	-
MGG_15207	Chromosome_8.3	-	5,0E-08	-
MGG_15459	Chromosome_8.1	-	8,0E-08	-
MGG_16058	Chromosome_8.1	-	1,0E-05	-
MGG_16113	Chromosome_8.1	-	3,0E-09	-
MGG_16175	Chromosome_8.1	-	5,0E-08	-
MGG_16619	Chromosome_8.3	-	3,0E-06	-
MGG_17132	Chromosome_8.4	2,0E-06	7,0E-04	9,0E-10
MGG_17255	Chromosome_8.4	-	1,0E-06	-
MGG_18019	Chromosome_8.7	-	7,0E-05	8,0E-11
MGG_18060	Chromosome_8.7	5,0E-14	5,0E-14	4,0E-06
M_TH16_EuGene_00000541	scaffold00001	-	1,0E-14	-
M_TH16_EuGene_00027191	scaffold00004	4,0E-10	-	-
M_TH16_EuGene_00027411	scaffold00004	4,0E-12	-	-
M_TH16_EuGene_00034081	scaffold00004	-	3,0E-08	-
M_TH16_EuGene_00040131	scaffold00005	2,0E-25	-	-
M_TH16_EuGene_00045871	scaffold00007	-	3,0E-12	-
M_TH16_EuGene_00079081	scaffold00016	6,0E-10	-	-
M_TH16_EuGene_00079311	scaffold00016	-	2,0E-08	-
M_TH16_EuGene_00099371	scaffold00026	-	3,0E-08	-
M_TH16_EuGene_00101881	scaffold00028	-	2,0E-05	2,0E-08
M_TH16_EuGene_00106621	scaffold00033	-	-	5,0E-14
M_TH16_EuGene_00120731	scaffold00052	3,0E-06	-	-
M_TH16_EuGene_00124981	scaffold00063	-	7,0E-11	-
M_TH16_EuGene_00127871	scaffold00072	-	2,0E-21	-
M_TH16_EuGene_00134971	scaffold00110	-	9,0E-14	1,0E-14
M_TH16_EuGene_00135161	scaffold00112	-	-	2,0E-19

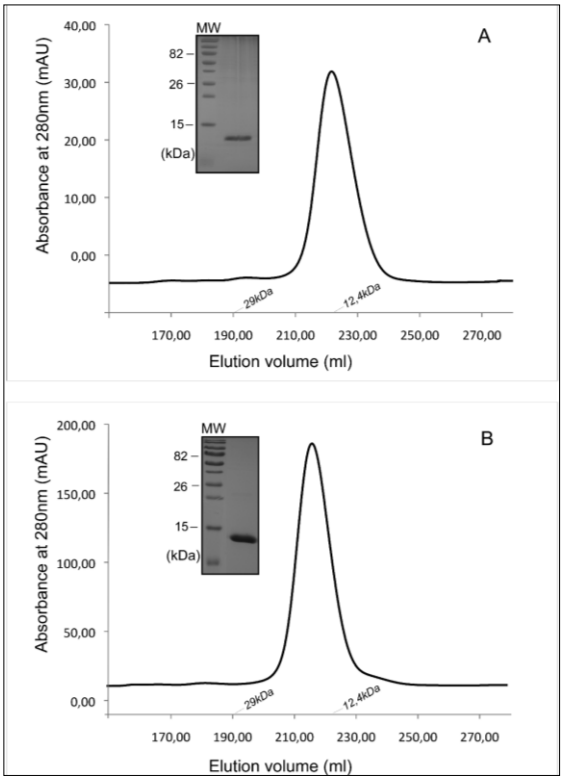
Orthologs of 7015 proteins in TH16 are not listed.

* Values correpond to e-values in psi-blast search with the corresponding protein

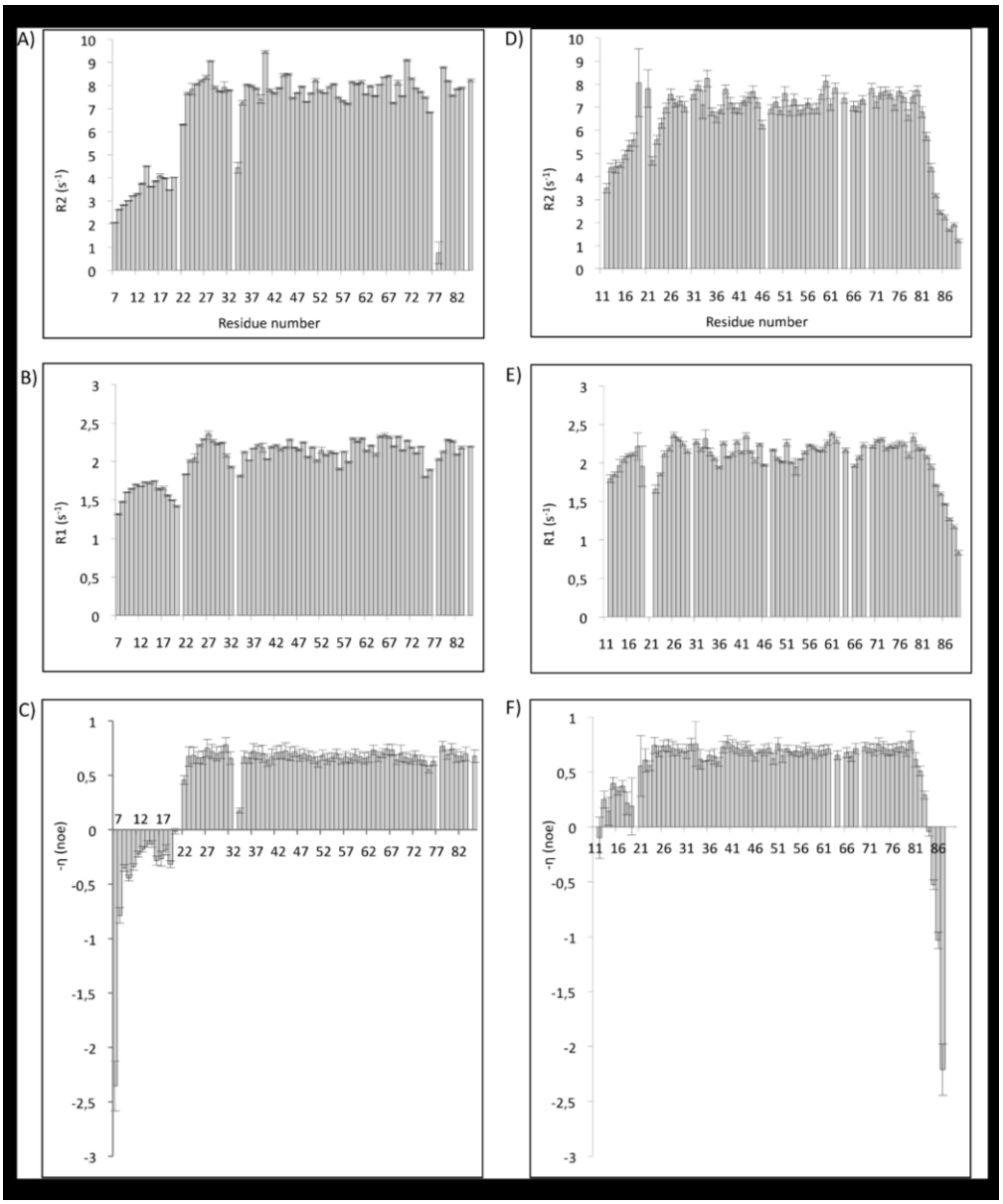
S 4 Table Primers used for qRT-PCR

GENE	Forward	Reverse
Infection specific expression		
MGG_00954	GTACCCAAGCCCTGGAAAGT	CCAGGACAATGCTCAAATCA
MGG_02207	GTGCCAATTGGTGGCTTACT	TCTTTCGCTTAGCGTCGTAAC
MGG_02546	AAGCTACGCCGTTTGTGAAC	CGGGTAATTTCTCGCTGTCA
MGG_04735	GCAAAGTCGTCACCGAGATT	CGTCCATTGTCCCTTCAATTA
MGG_05896	GACGTTTGCCTGGGTTAAAG	GGCAAGTGTAGCCGTTGATAA
MGG_08414	CGTCGTGGGTGATTTGAAC	GTACCAATTCCACGCCAGTC
MGG_08482	ACCATGCAAGGTCACGATTT	TGATATGTGTGGCCGGTAGA
MGG_08944	CCTGACCAATGGCAACAAG	AAATTCGCACTTGGAGTTCA
MGG_08992	AAGAGACGGGAACGTTGATG	ATCCACGTCGTACCGCTTAT
MGG_09675	CCGTCAAAGAGGAAGGATTC	TTTGGCAGAATGCAATTCAC
MGG_11967	ATGGGCTCAAAGCTCTTGTC	TGGATTTCTTGGTCCCAGTT
MGG_13133	AGTCGAAAGGGTTGGGAGAT	ATCGGATTGCACTCCTTGTC
MGG_15207	CCCACCTCAATCCTGAGAGA	TATCTCGCTCCCACGGTTAC
MGG_15625	CAGCAAAGAAGGTGGGAAGA	GCCGATACACATTTGCACTG
MGG_15972	TCGAGAGCGAGACCCTAACT	CCCAGTTTGGACCAACTTTC
MGG_16175	CGATTGCACACTTGGTTGTAA	TGCATCCTCTTCCGATCTTT
MGG_16619	GAAGCTAGGCGATCGTTCTG	GCAATCCTGTTCCACATGAA
MGG_16939	ATTGCTGGCCAAGGTTGAG	CGGCAGTTTCTGAACATGAG
MGG_17266	AAACACCAGCTCCCATCAAC	GCTGGAACTAGGCGACTCTG
MGG_18019	TGCCTTCCAGCCTACAGAGT	CTTTCGGAACGCTCTTCTTG
MGG_18062	CAACAGCAGCTTCGATATCAT	CCAACTTTGGCAAGCAATTC
M_TH16_00000541	GGAAATCAGCGGTTTCAAGT	CAAACTTCCAGCTCGTGACA
M_TH16_00027411	AGAGTCGAAGCAGGCGAAAT	TTCGCGGAGACTGTACATGA
M_TH16_00034081	AATCAGACGCCTGAGCAAGT	GCGATCGCAGGTAATGTAGC
M_TH16_00040131	GCTTGGCTGGGTTAAGATTG	TTTGCAGTTGGTGTCCTCTG
M_TH16_00079081	GGTTGGTGCAAGGTCAAAGT	CCGTCGAGATTCCAATGAGT
M_TH16_00079311	AAAGCTGCATGGCAAGACT	AACCTCCGATTGTTCCACTG
M_TH16_00104561	CGTCGTCGATTCAGCAAATA	GCTGCTTCCTTATCCCACAA
M_TH16_00119711	CAGCGGTGTAGTGGCAATAA	TCCTTCCCAGCTTTCACAGT
M_TH16_00120731	ACCGTGGGAATGGGTTACTT	TAAACTCCCGGCTGGATAAG
M_TH16_00124981	GCTTGGGTTGCACCTACAGT	TTCAAACTCGCAGGTGTCTG
M_TH16_00127871	GTTGCCTGTGCAGCCTATCT	GATTCGGCTGGAATCATCTC
M_TH16_00136331	AAGGAAGAAGTCGAGGGTGAG	AACGATATCCCAACCGTCAG
Constitutive expression and controls		
MGG_14793	ACAGCCGTCTGCGACTTTAT	GTCCGCTCAGGCTAAGTTTG
MGG_07184	AGCGTGTCTGCAAAGCTGTA	AGACCTCCCAACGGTTCTCT
Actin, MGG_03982	TCTTCGAGACCTTCAACGCC	ACCGGAGTCGAGCACGATAC
EF1α, MGG_03641	GCC CGG TAT GGT CGT TAC CT	AGC TGC TGG TGG TGC ATC TC
ORF3, MGG_08381	GGTGAGGGTGTTGGAGGTAGTG	TGGAGCTGCCCAACATGTC
MGG_01147	CGACGACCTACCTGCTGACT	CAATGCTCCTCTTCCTGGAG
BAS3, MGG_11610	CCCGTGTTTGAGGAATTGTG	CTTGAGGTTGTCGGTGCTCT
No expression		
MGG_00821	GCGGCTACACGATTGAGAC	**ACCTCGTGGCTATTCCTGAC**
MGG_08469	CCAGACGTACCTCCCACTTG	**AGGCCCTCCAATTTCTTCTC**
MGG_08607	ACGCAACAACCGAAGAAATC	**GGTCCAACCGGAACTGATAA**
MGG_09425	TAGCCAGGAAGGCACTTACA	**GCGATCAAATCGAACCTTGA**
MGG_10335	ATCCTTGTAGGCGCAAAGTG	**GCAGGTTAATGCCTTTGACA**
MGG_12426	GGCAGGAGAGTCACCTTCAG	**CGGAGATGATGCAACGTTTA**
MGG_15106	TTTGGATCACCGGGAAATAC	**CTCGCCGCCAATAAACTTTC**
MGG_15575	AGGACTTGCCAAAGGGATTT	**GCGCCAAGAACTCATCAATC**
MGG_16113	GGGTTATGGCACGGTTGTTA	**GCCTGAGGCAGTTTCAGTTC**
MGG_16416	GCAACTCATTCGCCACAGTA	**GCAATGCCCATAACCAGTCT**
MGG_17132	AAATGCAATTCAAATCCCTCTT	**ACCATCTCCTCCGGATTTCT**
MGG_17255	AGTCGTGTCACTGGCATCAG	**GCCACCAATTGACGTTCTTT**
MGG_17799	TTGGCACGAAATGTGCTATT	**TCGTAGTCCAACCGGACTGT**
MGG_18041	GGGTTGGGCCGTTAGATTT	**CCAAATTGGTGCATTTCAGA**
MGG_18060	GGGTGTCCATATTCCGTGAG	**TGATCGGCTGAGATTGTAGC**
M_TH16_EuGene_00099371	TCCTGGCTTCGTTCAATAAGA	**TCGAATGATAGTGCGGGTAAC**
M_TH16_EuGene_00101881	GTGCAAGGTCGAGATGTTTG	**CGGACCAGTTCTTGGGTAAG**
M_TH16_EuGene_00106631	TCGAAATTCAACGTCGTGAA	**CTCAAAGCGCTCACGGTAAG**

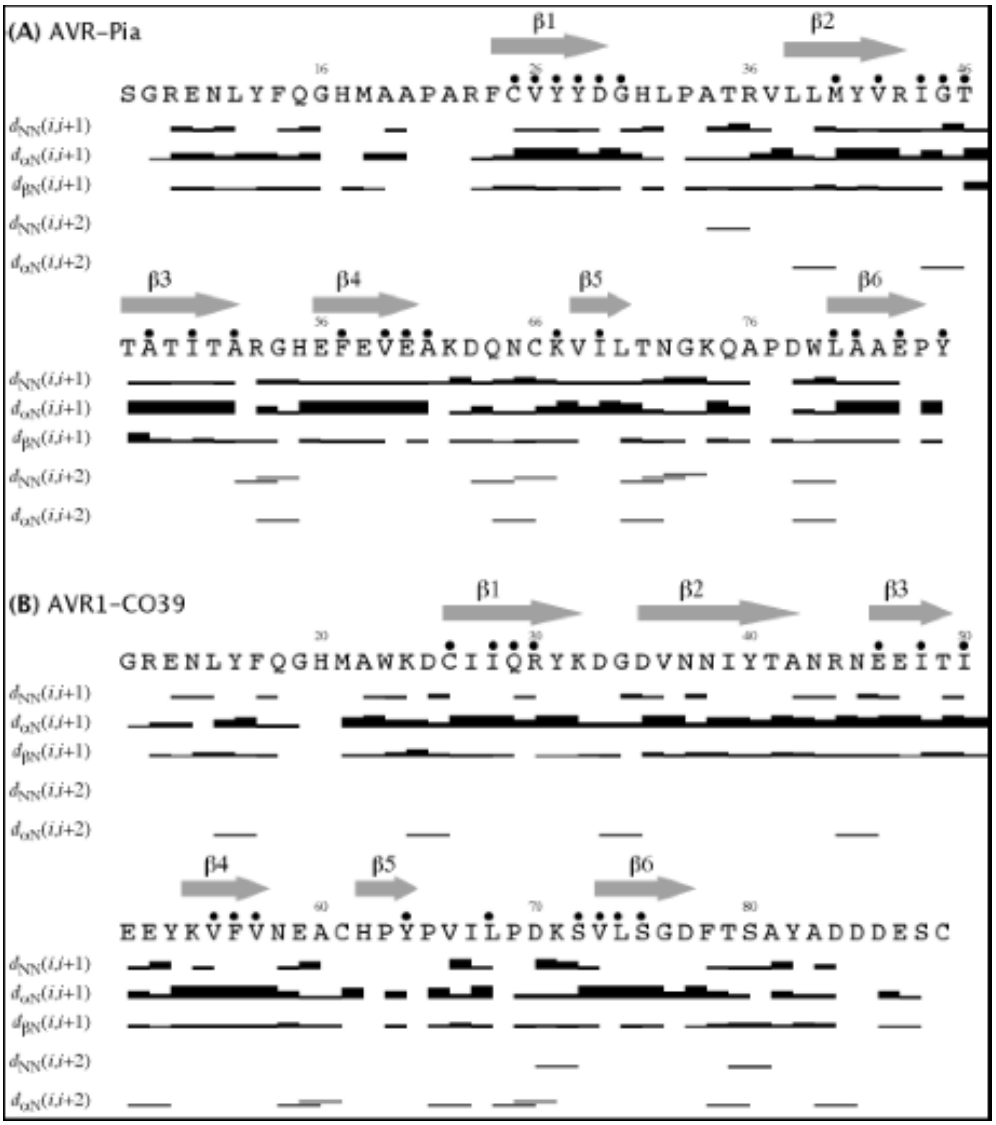
S1 Fig. Gel filtration profile and SDS-PAGE analysis of purified AVR-Pia (A) and AVR1-CO39 (B) proteins.



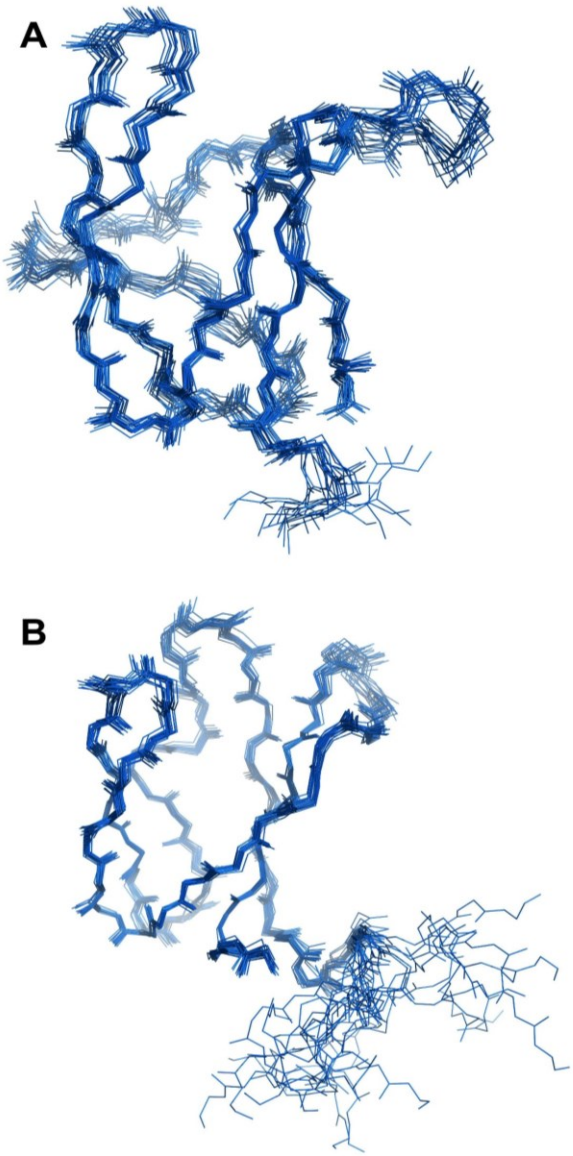
S2 Fig. 15N Relaxation data at 500 MHz and 305K for AVR-Pia (panels A, B and C) and AVR1-CO39 (panels D, E and F).



S3 Fig. Backbone sequential and medium range NOEs observed for (A) AVR-Pia and (B) AVR1-CO39. The line width is proportional to the NOE intensity. The dots (•) indicate slow exchange NH observed in 2D-NOESY in D2O. Grey arrows indicate the ß-strands determined from the structure analysis (vide infra).

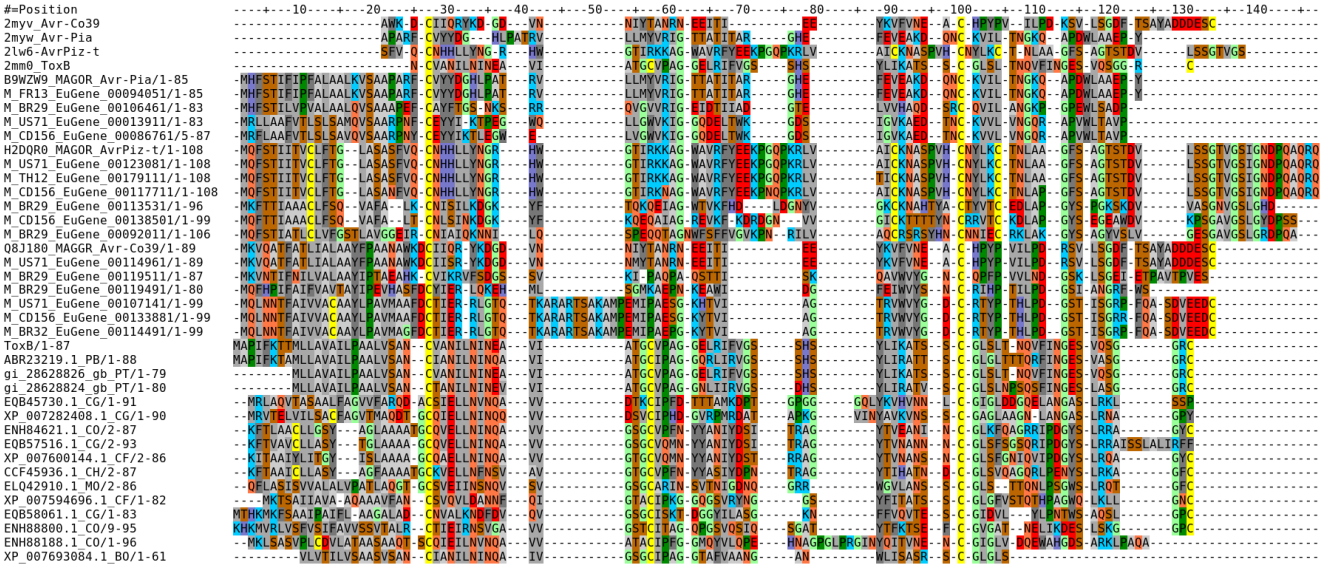


S4 Fig. Solution structures of (A) AVR-Pia and (B) AVR1-CO39. Superposition of the backbone atoms of the 20 lowest energy conformers used to calculate the final structures. Only mature chains are shown, from residues Ala20 and Trp23 for AVR-Pia and AVR1-CO39, respectively.



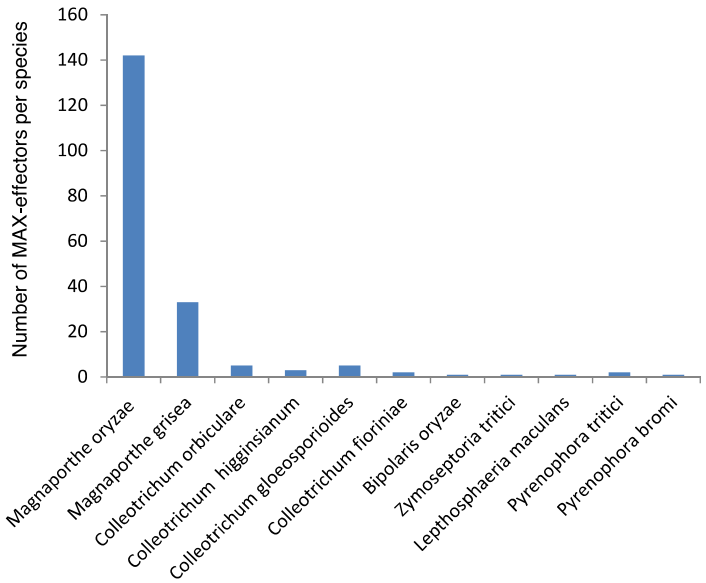
S5 Fig. Structure-guided alignment and diversity of MAX-effector homologs identified by Blast. A) Homologs of AVR1-CO39, AvrPiz-t and AVR-Pia identified by Blast in *M. oryzae* and *M. grisea* genomes and ToxB homologs identified by Psi-Blast in the GeneBank database were aligned to the structural alignment of mature ToxB, AVR1-CO39, AvrPiz-t and AVR-Pia. (B) A diversity tree was constructed by the neighbor-joining method using the alignment in (A). It highlights the high diversity of MAX-effector homologs. Branch supports are based on 1000 bootstraps and horizontal branch length reflects sequence divergence. Accession numbers of non-Magnaporthe sequences were completed by a 2 letter identifier for the species: BO for *Bipolaris oryzae*, CF is for *Colletotrichum fioriniae*, CH for *C. higgensianum*, CG for *C. gloeosporioides*, CO for *C. rbiculare*, LM for *Lepthosphaeria maculans*, PT for *Pyrenophora triticirepent* is and PB for *Pyrenophora bromi*.

A



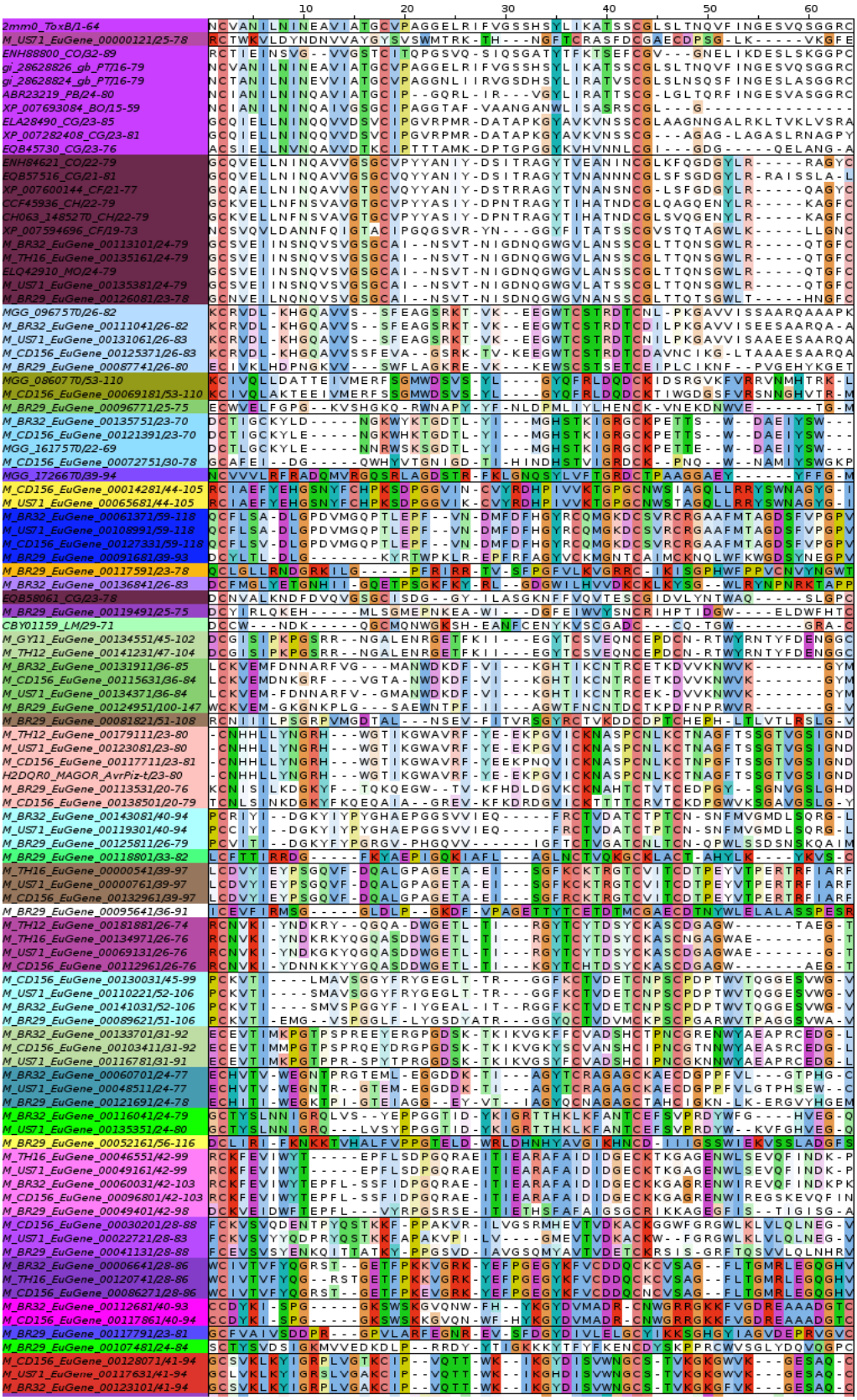
S6 Fig. MAX-effector homologs identified by a high stringency HMM search. (A) Histogram showing the numbers of MAX-effectors identified by an HMM pattern search in a nonredundant database comprising the small secreted proteins of 25 ascomycete fungi and of 8 additional *M. oryzae* and one *M. gisea* isolate. (B) MAX-effectors were aligned to the structural alignment of mature ToxB, AVR1-CO39, AvrPiz-t and AVR-Pia and gaps were removed. (C) A diversity tree was constructed by the neighbor-joining method using the alignment in (B). Branch supports are based on 1000 bootstraps and horizontal branch length reflects sequence divergence. Accession numbers of non-Magnaporthe sequences were completed by a 2 letter identifier for the species: BO for *Bipolaris oryzae*, CF for *Colletotrichum fioriniae*, CH for *C. higgensianum*, CG for *C. gloeosporioides*, CO for *C. orbiculare*, LM for *Lepthosphaeria maculans*, PT for *Pyrenophora tritici-repentis* PB for *Pyrenophora bromi* and ZT for *Zymoseptoria tritici*.

A

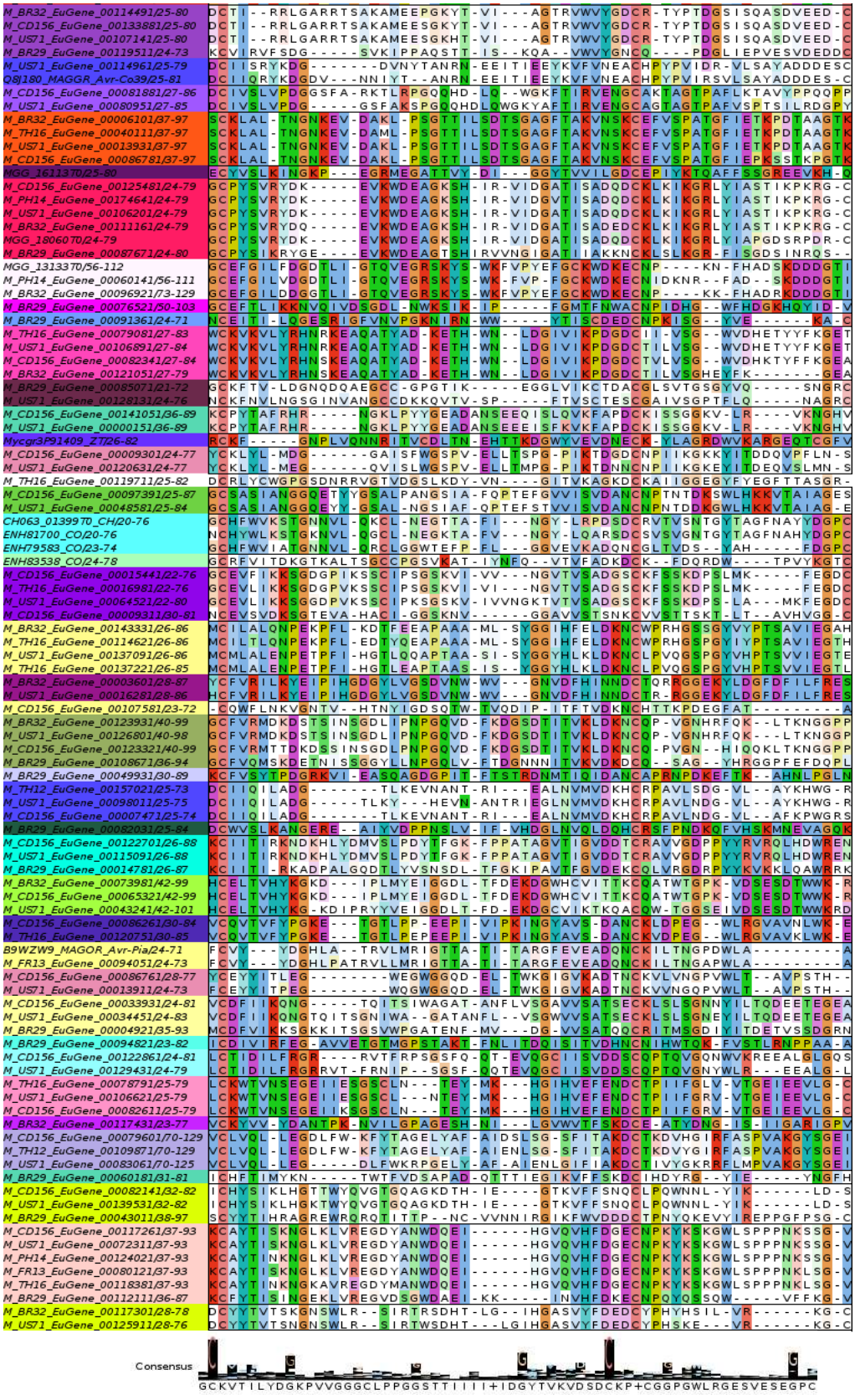


S6 Fig. continued

B

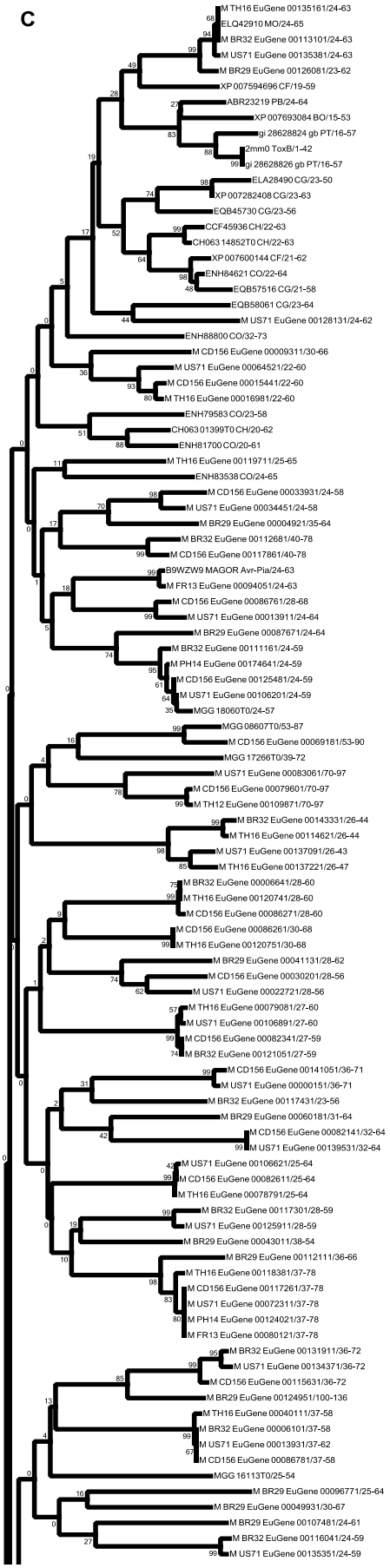


S6 Fig.B continued

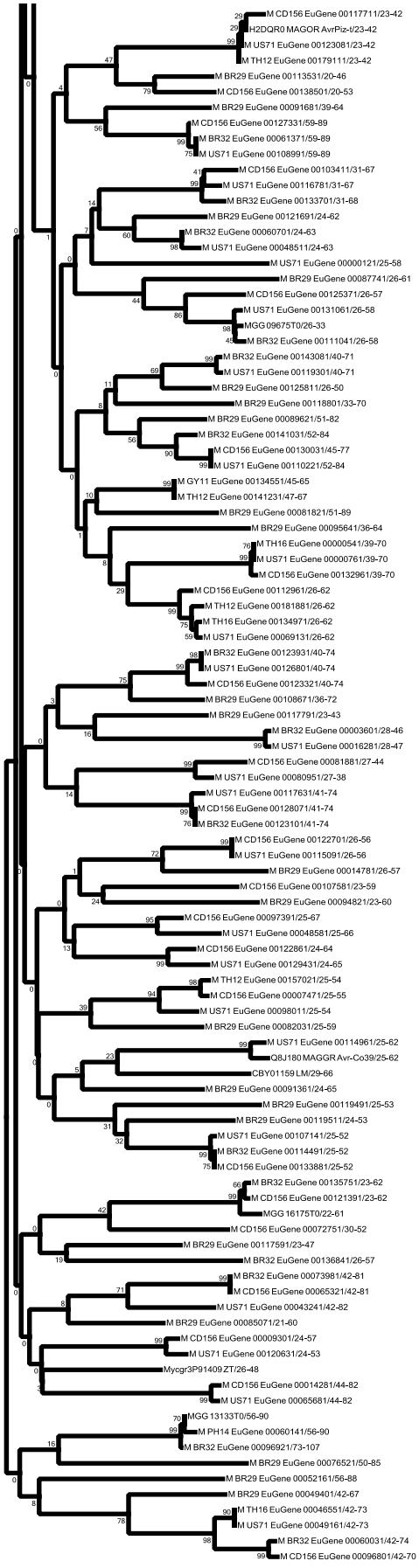


S6 Fig. continued

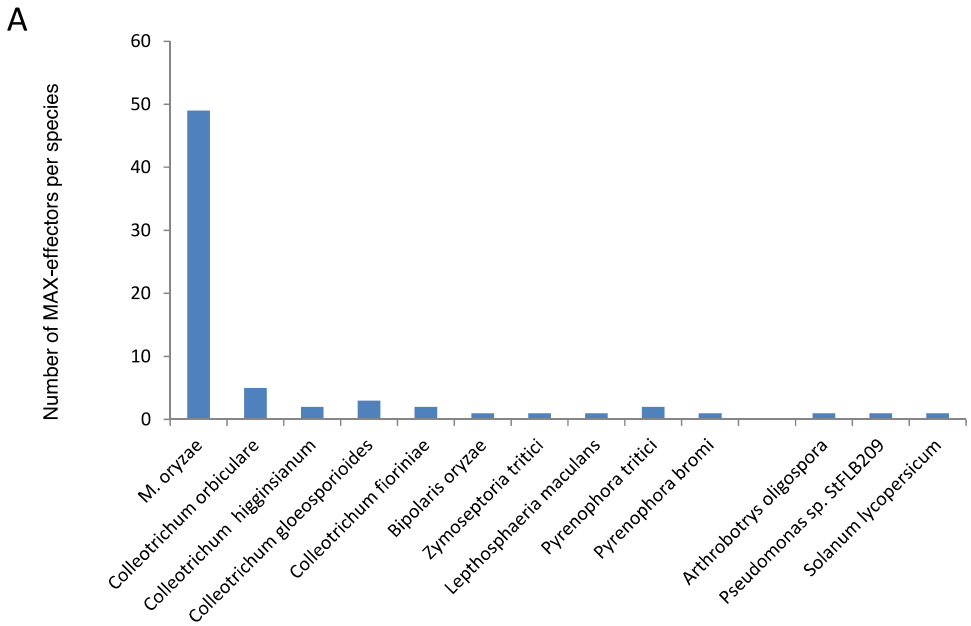
C



S6 Fig.C continued

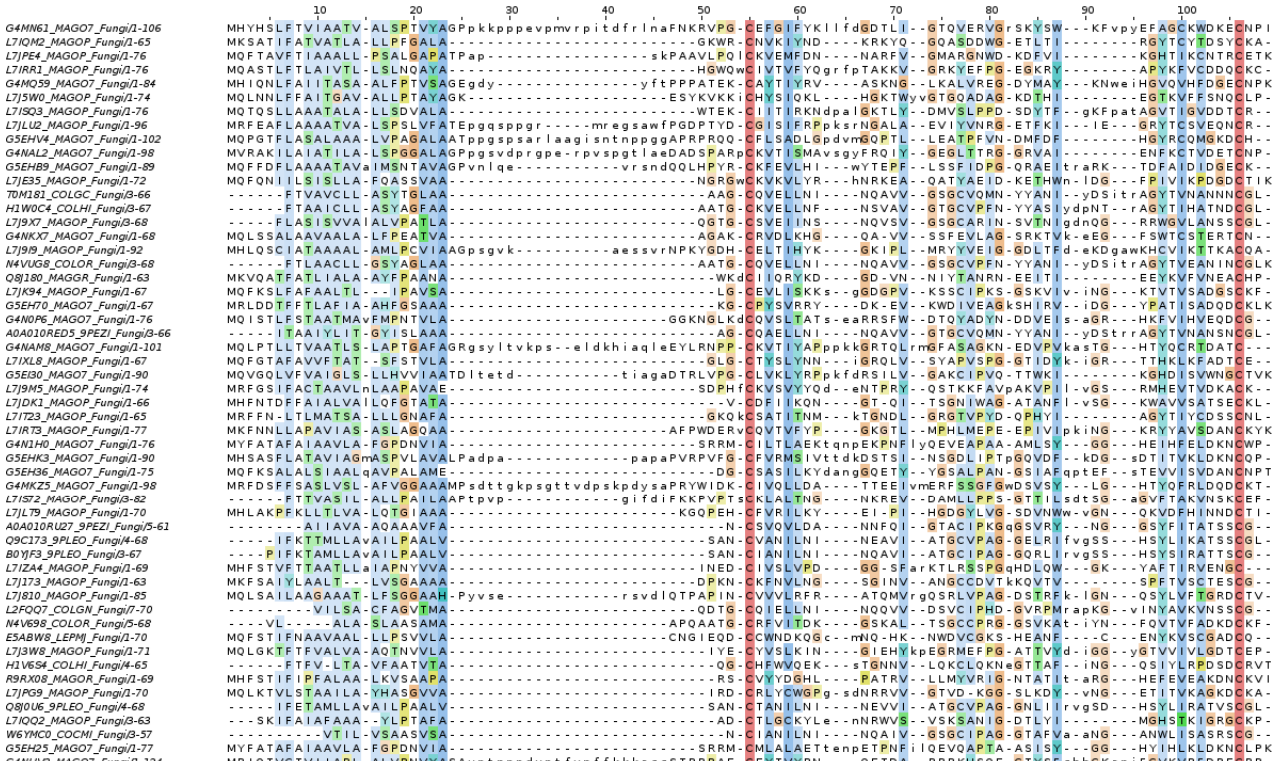


S7 Fig. MAX-effector homologs identified in the UniRef90 database by a low stringencyHMM search. (A) Histogram showing the numbers of MAX-effectors identified by an HMM pattern search in a non-redundant UniRef90 database. (B) MAX-effectors identified by HMMpattern search were aligned to the structural alignment of mature ToxB, AVR1-CO39, AvrPizt and AVR-Pia. (C) A diversity tree was constructed by the neighbor-joining method using the alignment in (B). This highlights the high diversity of MAX-effector homologs. Branch supports are based on 1000 bootstraps and horizontal branch length reflects sequence divergence. Accession numbers contain the following information on the species: MAGGR, MAGO7, MAGOP and MAGOR are from *M . oryzae* , COLGC and COLGN from *C . gloeosporioides* , COLHI from *Colletotrichum higginsianum* , 9PEZI from *C . fioriniae* and COLOR from *Colletotrichum orbiculare* , 9PLEO from *P . tritici-repentis* or *P . bromi* , ARTOA from *Arthrobotrys oligospora* , COCMI from *Bipolaris oryzae* , LEPMJ from *Leptosphaeria maculans* , MYCGM from *Zymoseptoria tritici* , 9PSED from *Pseudomonas sp . StFLB209* and SOLLC from *Solanum lycopersicum*.



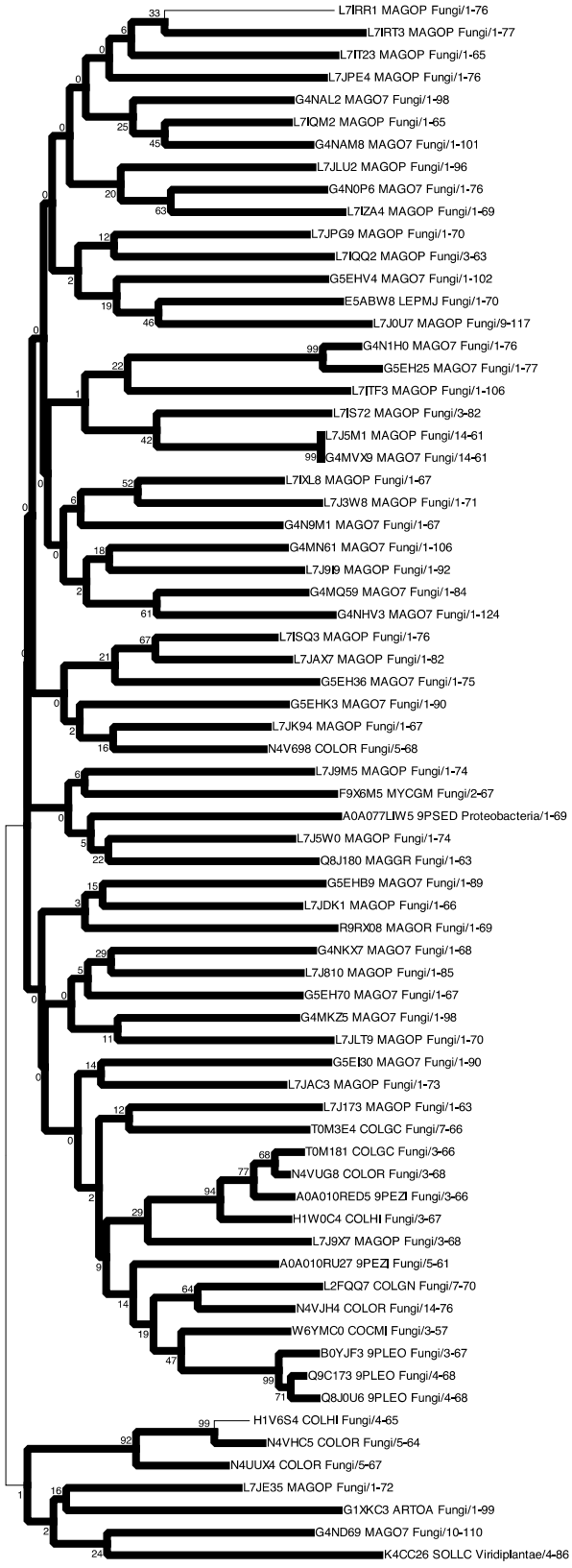
S7 Fig. continued

B



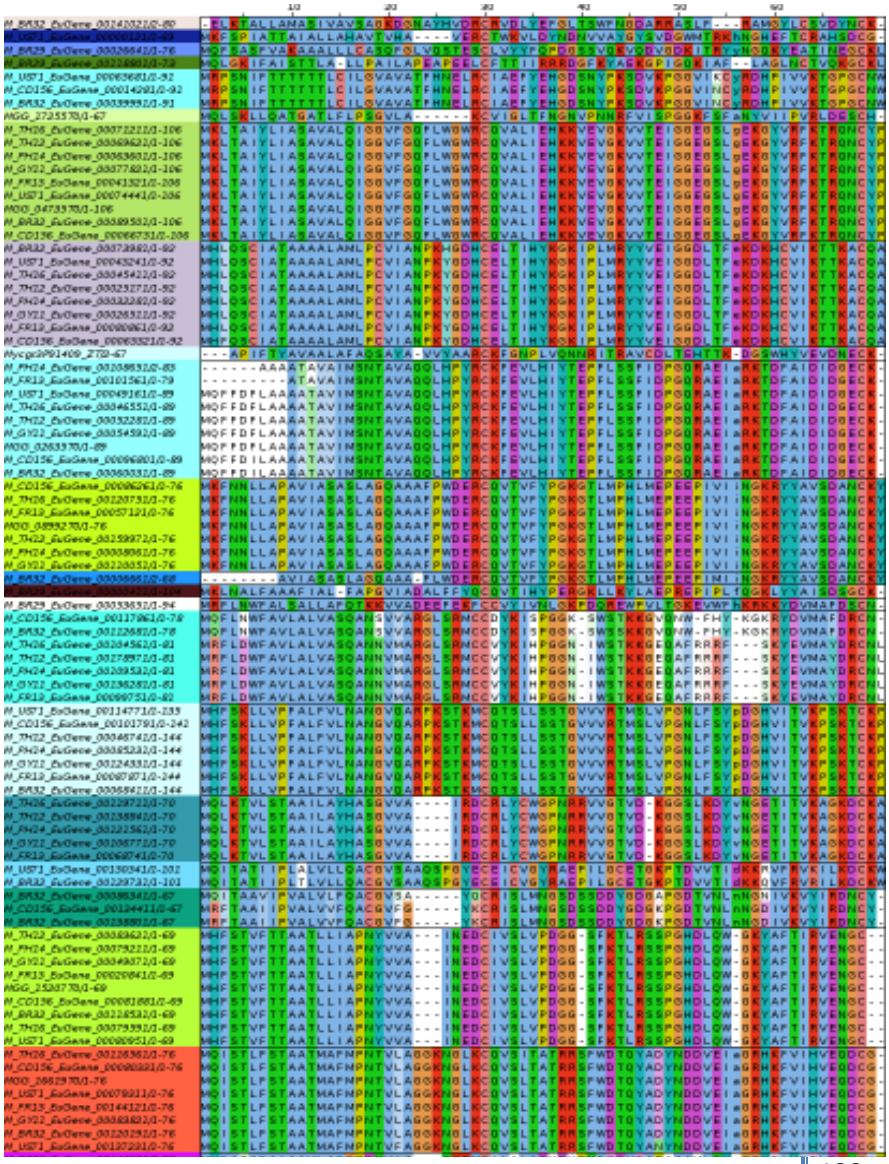
S7 Fig. continued

C

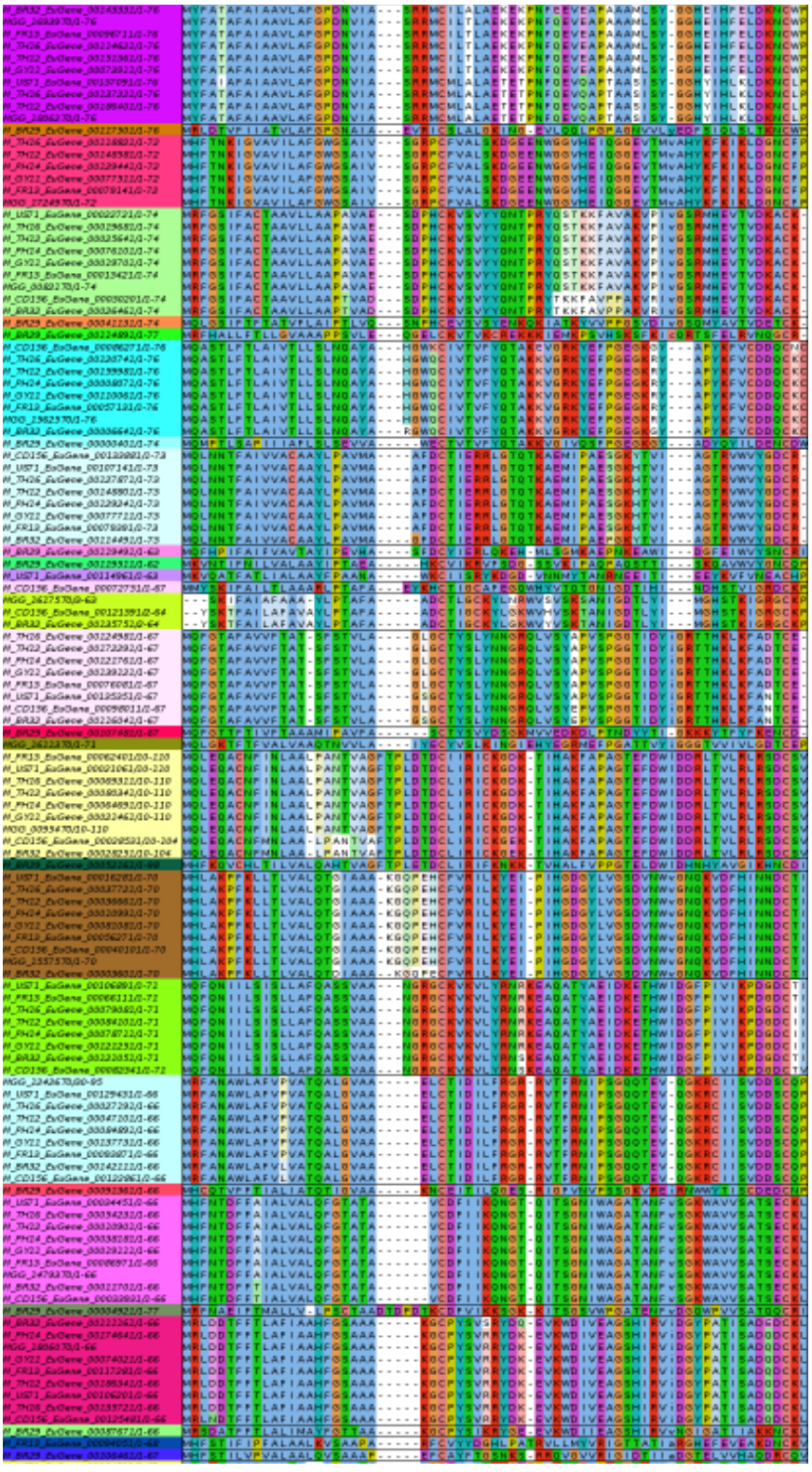


S8 Fig. MAX-effector homologs identified by a low stringency HMM search in the fungal genomes database. (A) MAX-effectors identified by an HMM pattern search in a redundant database comprising the small secreted proteins of 25 ascomycete fungi, 8 additional M. oryzae and one M. gisea isolate were aligned to the structural alignment of mature ToxB, AVR1-CO39, AvrPiz-t and AVR-Pia and gaps were removed. (B) A diversity tree was constructed by the neighbor-joining method using the alignment in (A). Branch supports are based on 1000 bootstraps and horizontal branch length reflects sequence divergence. Accession numbers of non-Magnaporthe sequences were completed by a 2 letter identifier for the species: BO for Bipolaris oryzae, CF for Colletotrichum fioriniae, CH for C. higgensianum, CG for C. gloeosporioides, CO for C. orbiculare, GF for Fusarium fujcuroi, LM for Lepthosphaeria maculans, PT for Pyrenophora tritici-repentis, PB for Pyrenophora bromi and ZT for Zymoseptoria tritici.

A

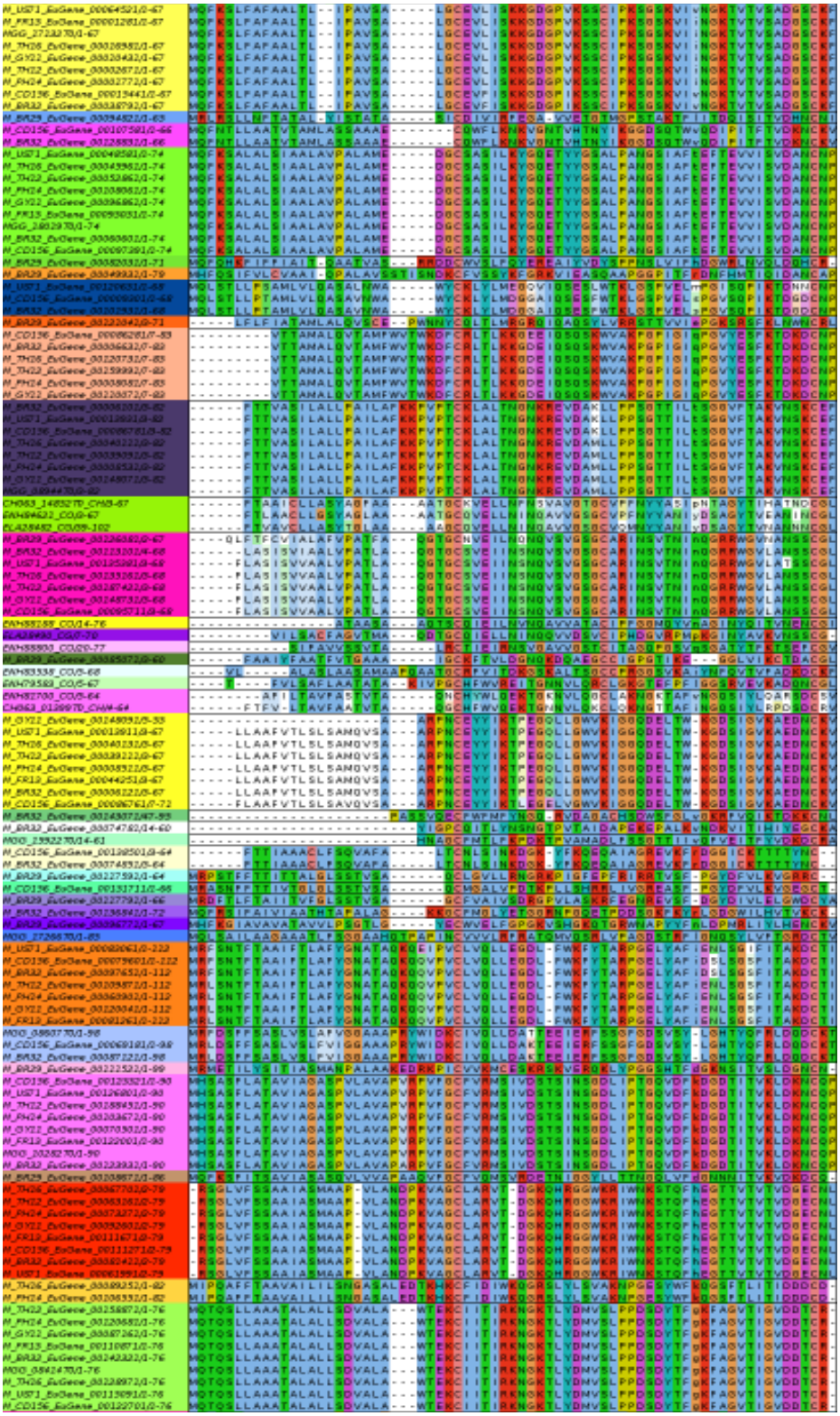


S8 Fig.A continued

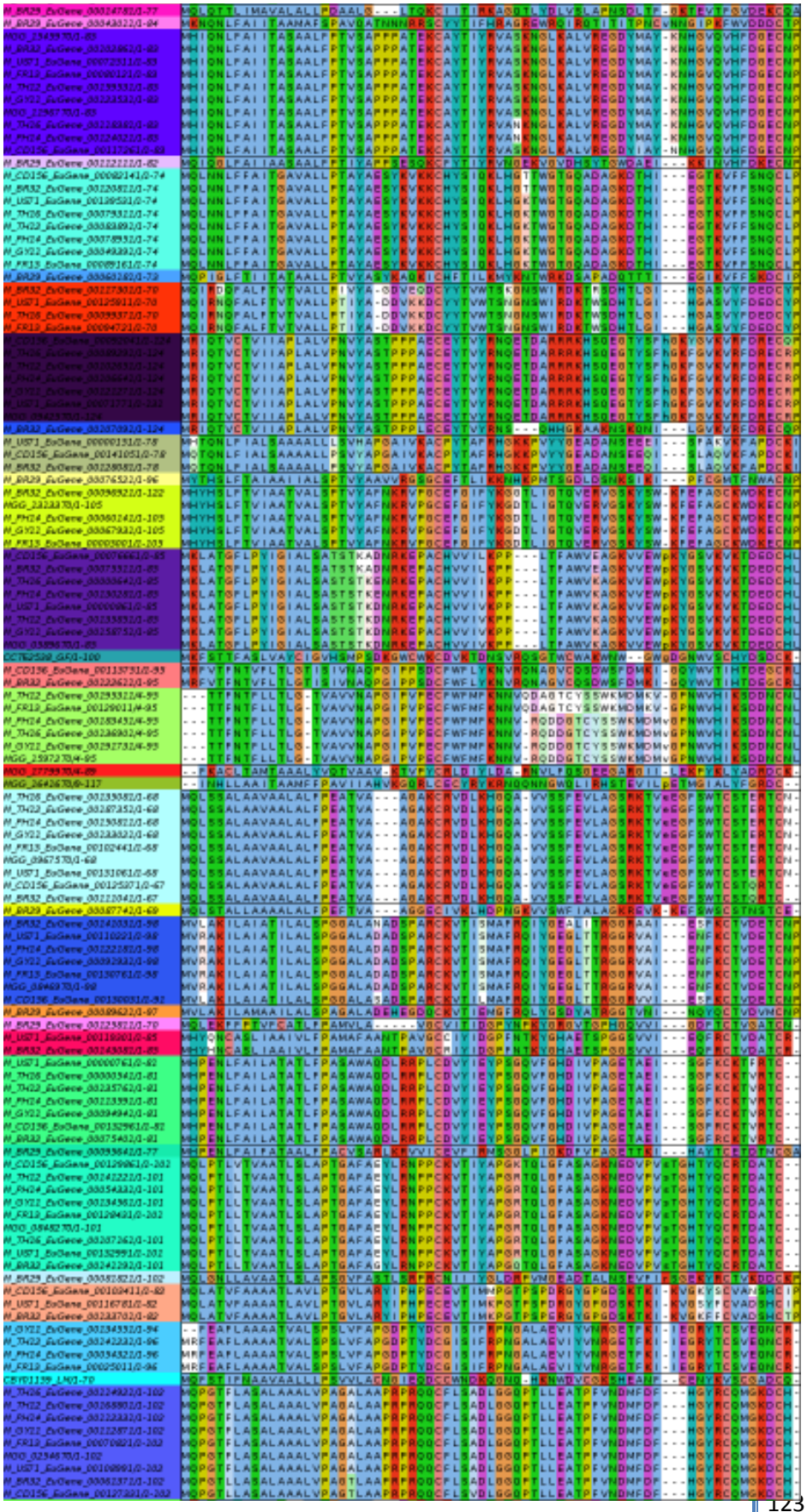


.

S8 Fig.A continued

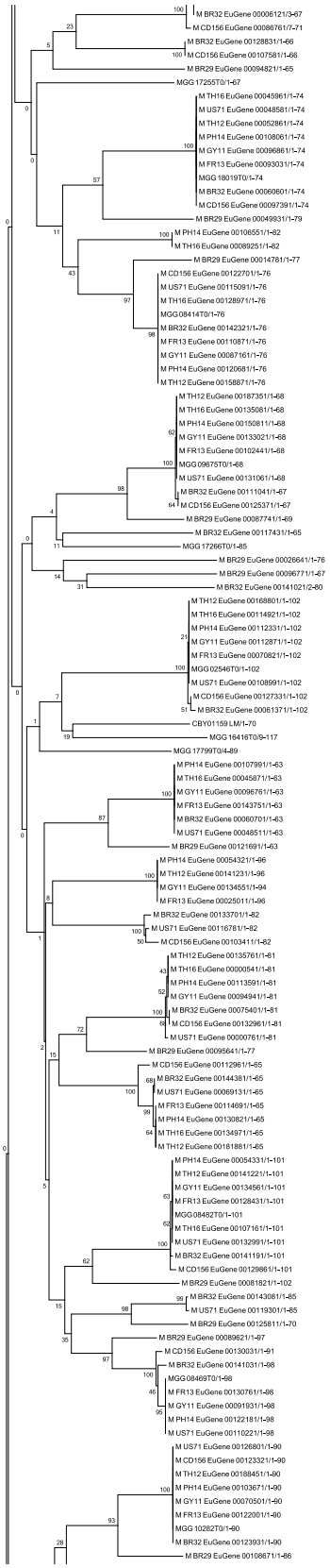


S8 Fig.A continued

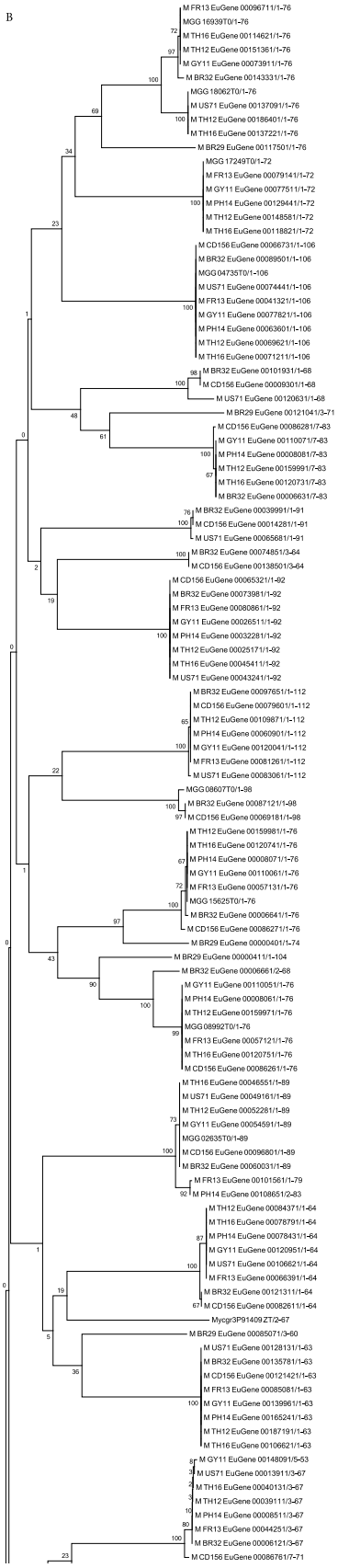


S8 Fig.A continued

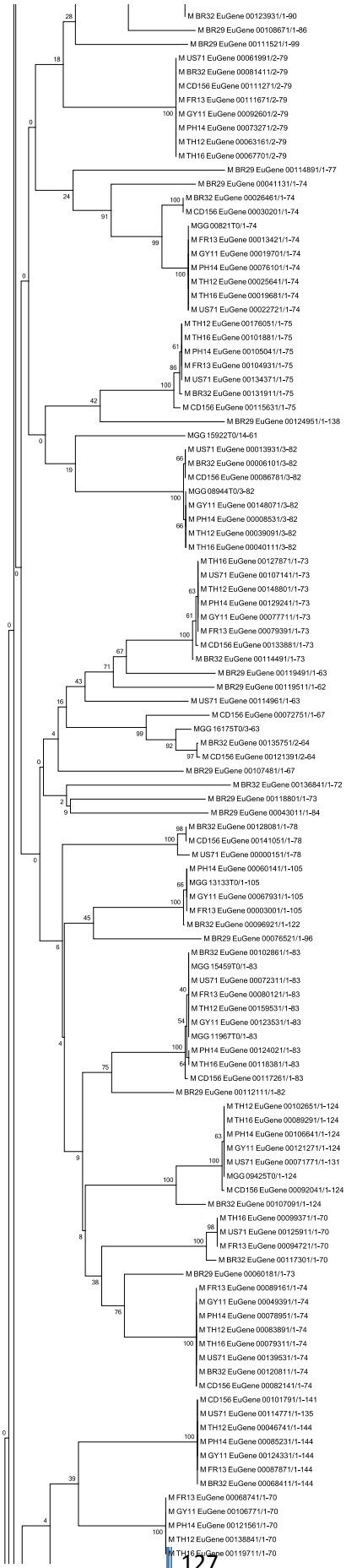
S8 Fig.C continued



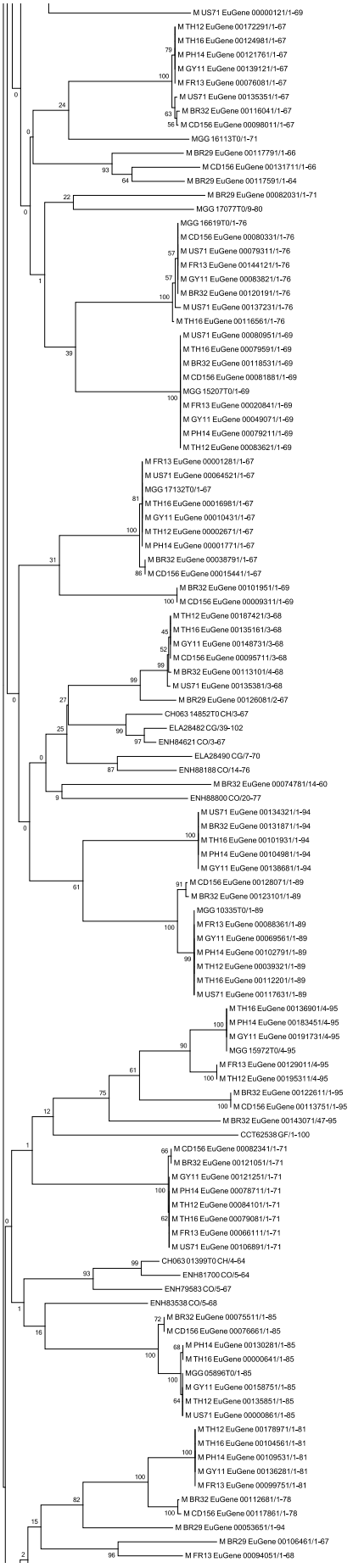
S8 Fig.C continued



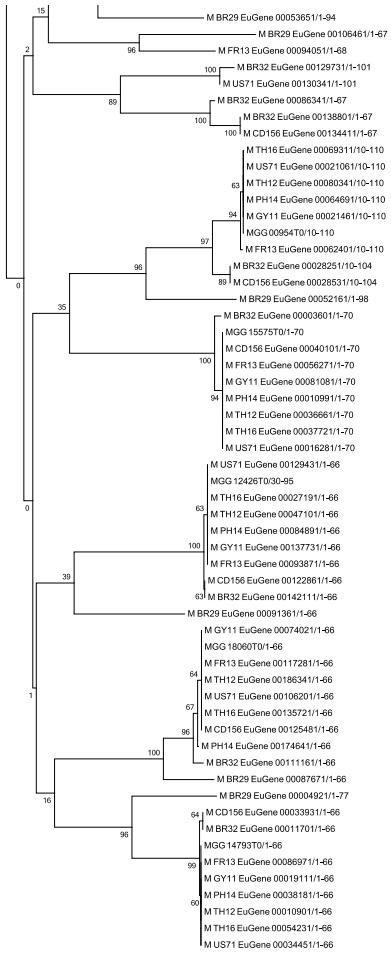
S8 Fig.C continued



S8 Fig.C continued

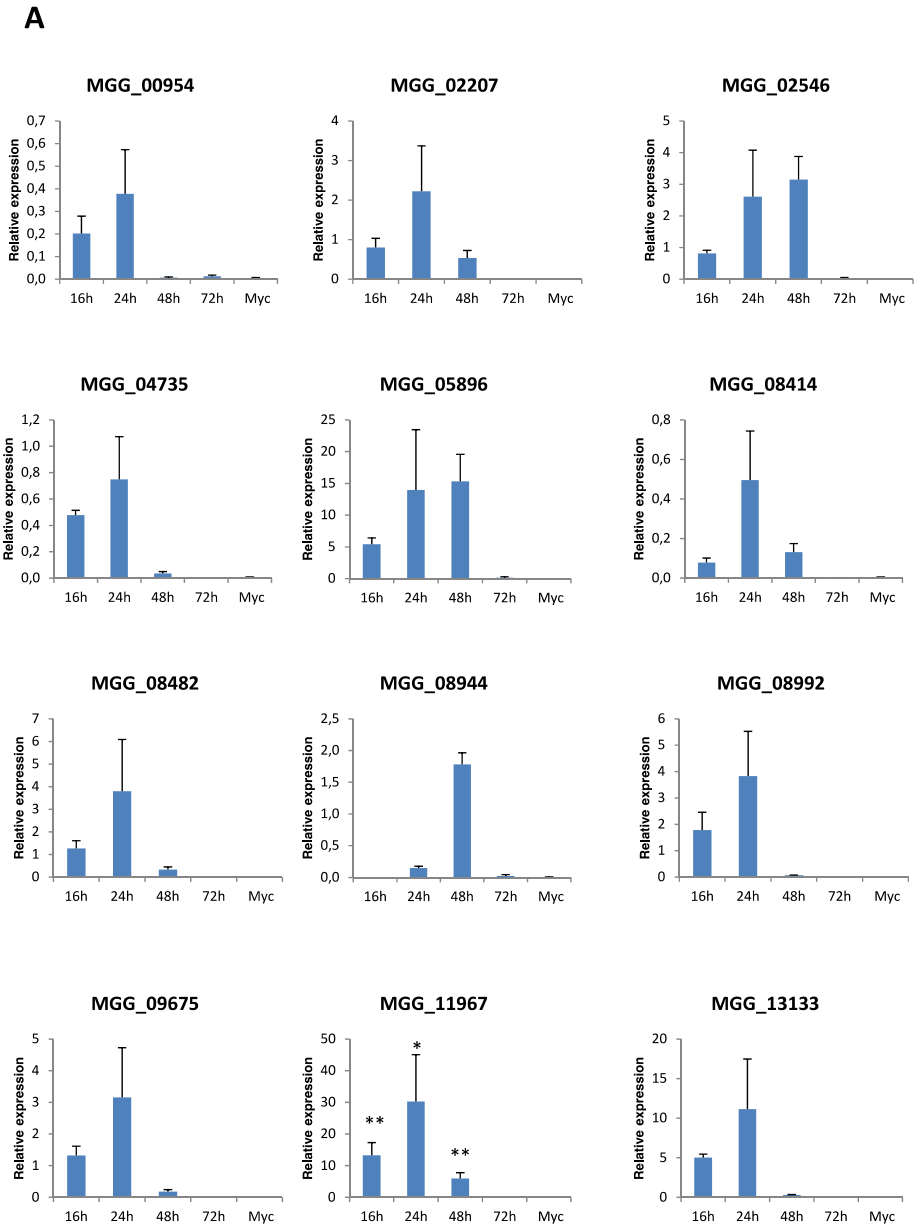


S8 Fig.C continued

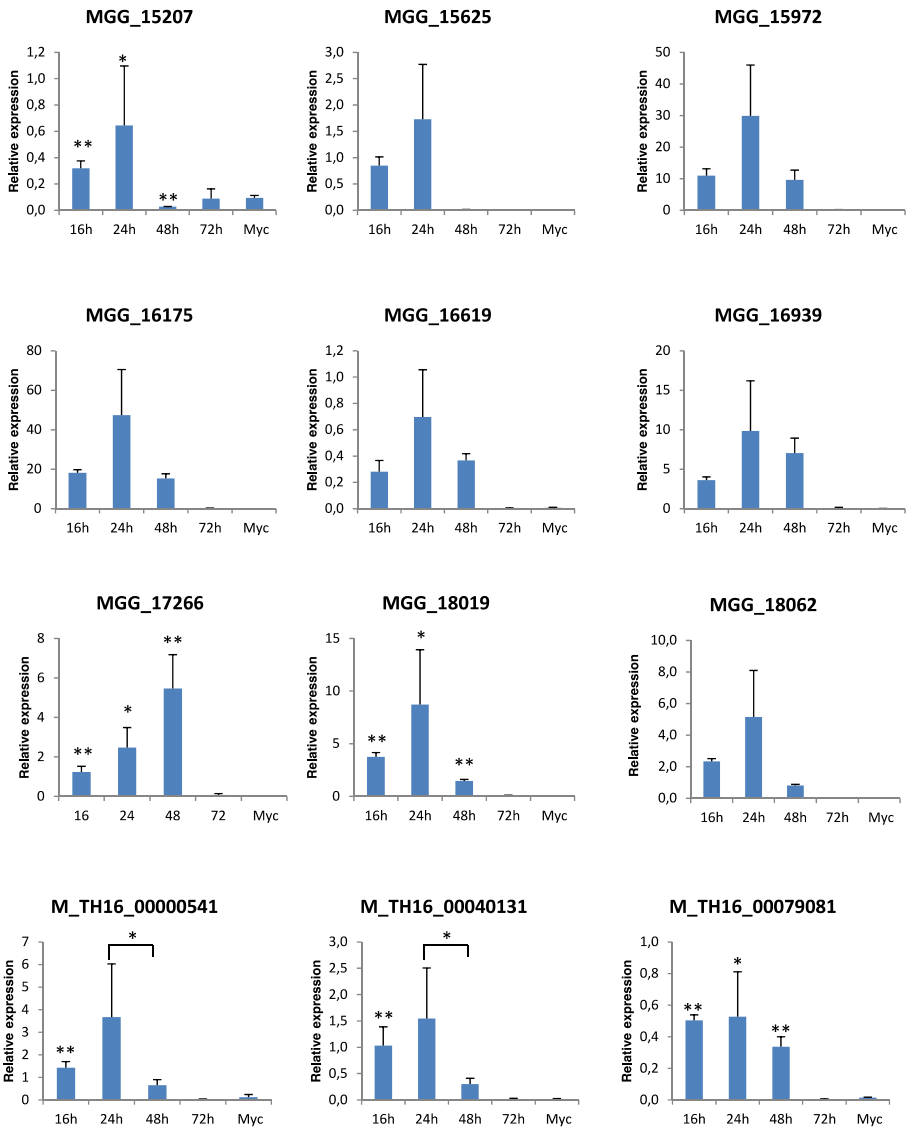


S9 Fig. Expression of *M. oryzae* MAX-effector candidates and marker genes during rice infection and in in vitro grown mycelium. mRNA levels of *M. oryzae* genes coding for MAX-effectors (A, B and C) and marker genes (D) was determined by q-RT-PCR in rice leaf samples harvested 16, 24, 48 or 72 h after inoculation and mycelium grown liquid medium for 72 hours. (A) Infection specific MAX-effectors identified in the HMM search, (B) infection specific MAX-effectors identified in the Psi Blast search but nor in the HMM search, (C) constitutively expressed MAX-effectors identified in the HMM search and (D) marker genes for appressorium and very early infection (*ORF3* of the ACE1 cluster, *MGG_08381*), biotrophic infection (*BAS3, MGG_11610*), late infection (MGG_01147), constitutive expression (*EF1α, MGG_03641*). Relative expression levels were calculated by using expression of a constitutively expressed Actin (MGG_03982) as a reference. Mean values and standard deviation were calculated from three independent biological samples.The analyzed genes, were in most cases not or extremely weakly expressed in the mycelium. For genes with significant expression in the mycelium (ratio gene versus actine > 0,01) a T-test was performed to determine if in planta expression was significantly different from expression in the mycelium. In these cases (MGG_11967, MGG_14793, MGG_15207, MGG_17266, MGG_18019, M_TH16_00000541, M_TH16_00040131, M_TH16_00079081, M_TH16_00104561, M_TH16_00120731, M_TH16_00124981), a star or two stars (* or **) mark conditions where the expression was different from expression in the mycelium at respectively $p<0,05$ or $p<0,005$.

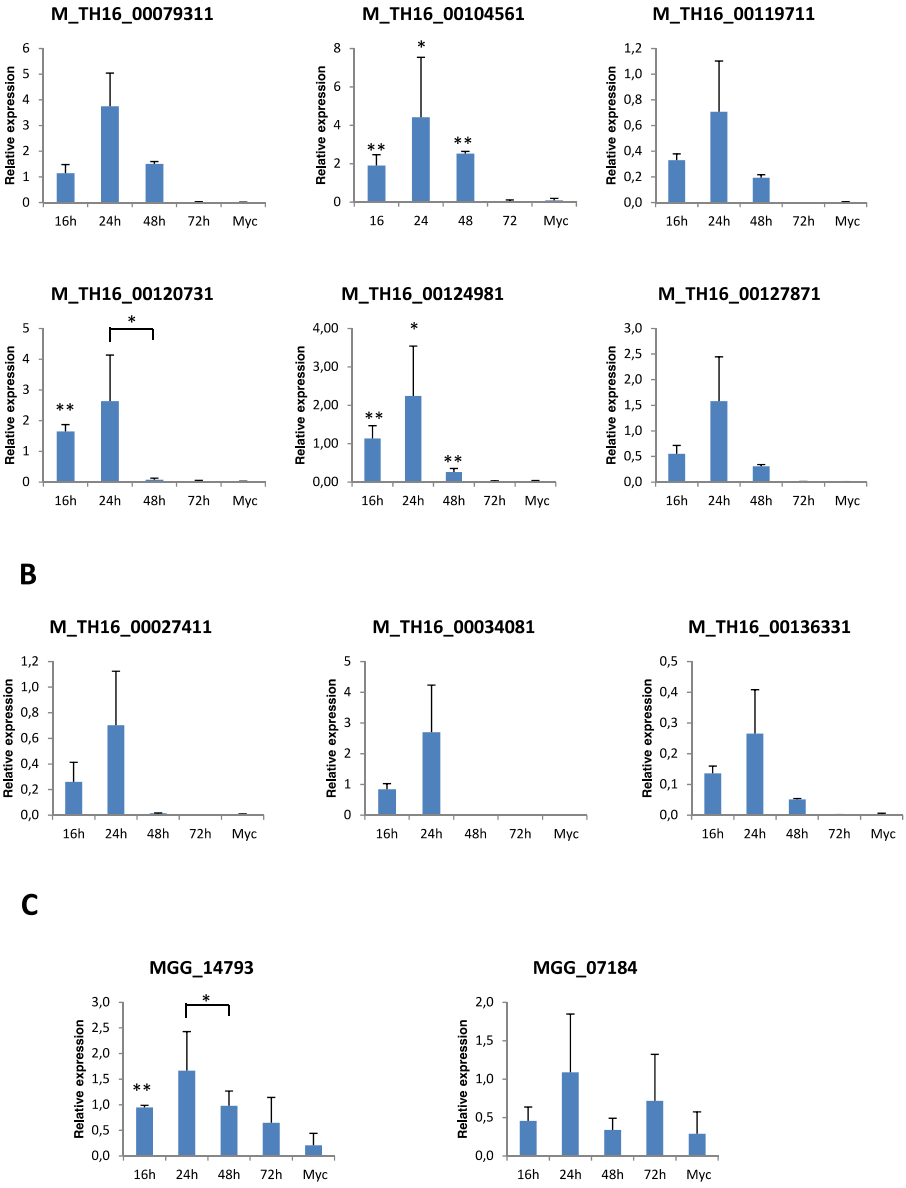
S9 Fig. continued



S9 Fig. continued

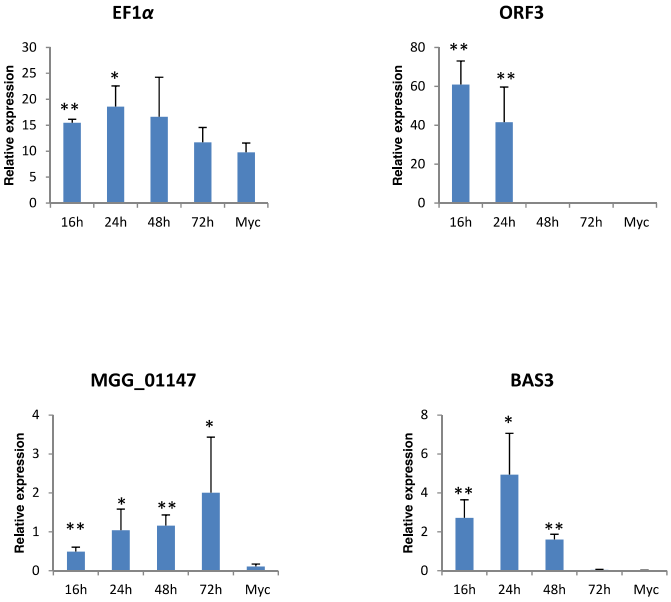


S9 Fig. continued

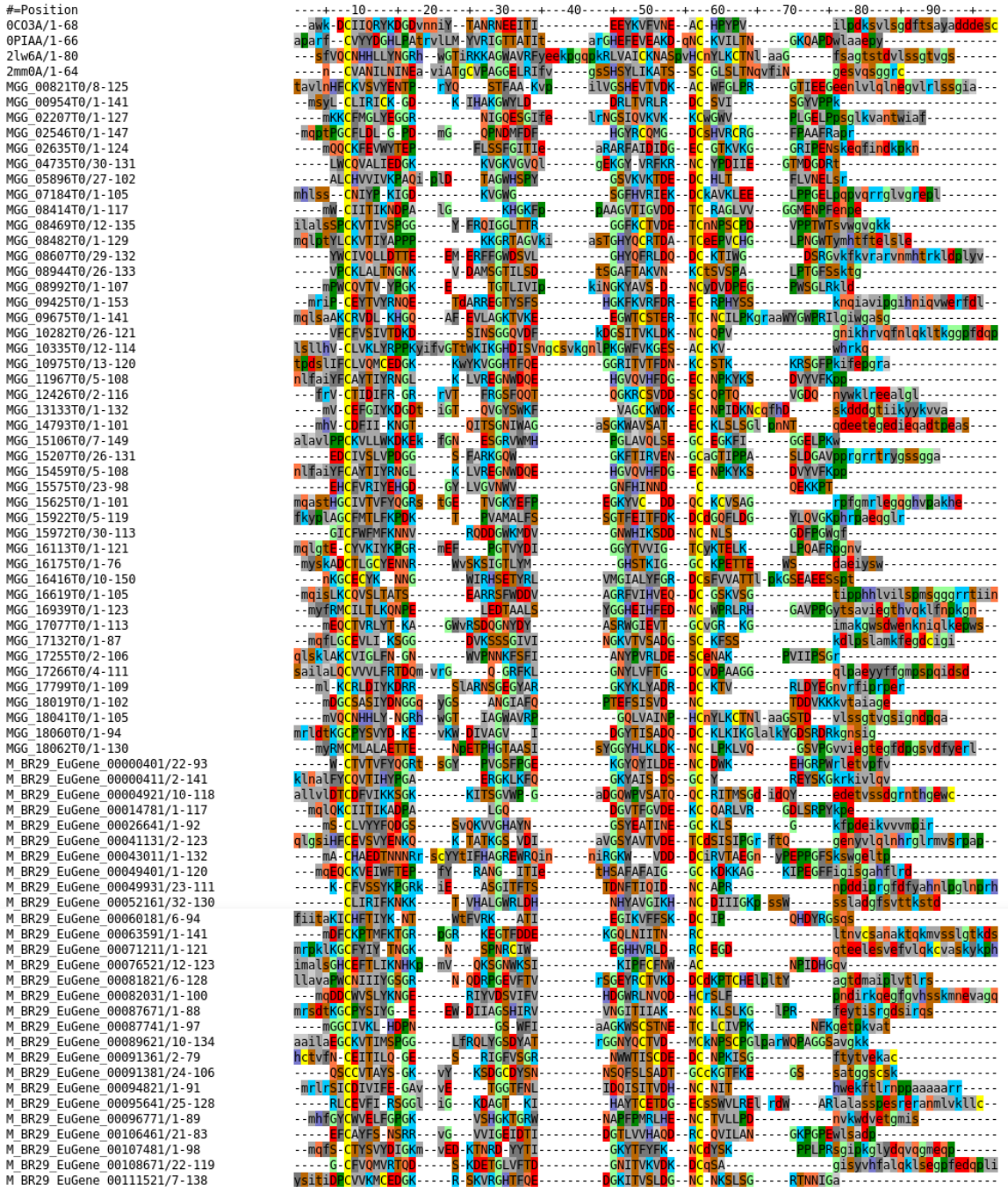


S9 Fig. continued

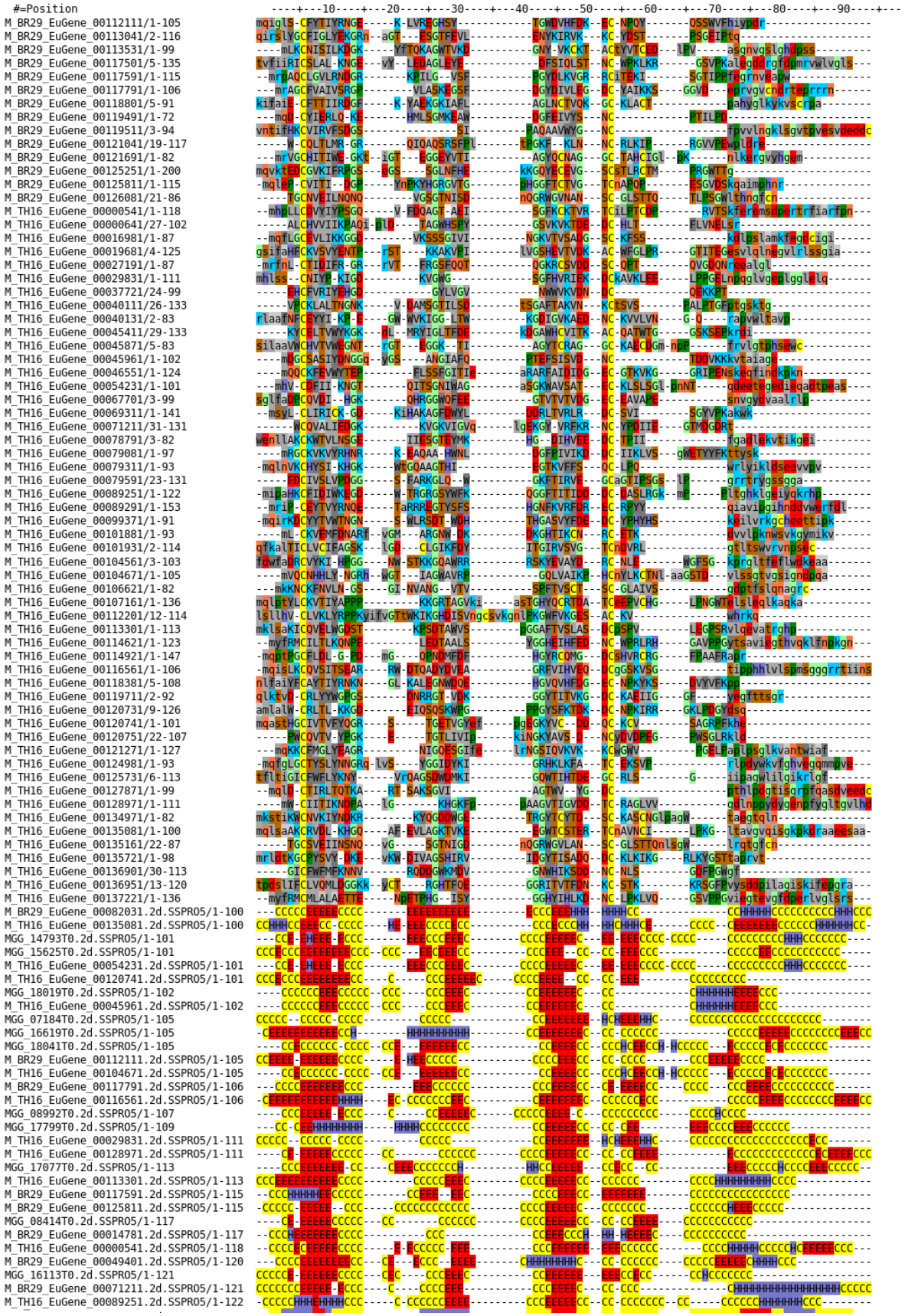
D



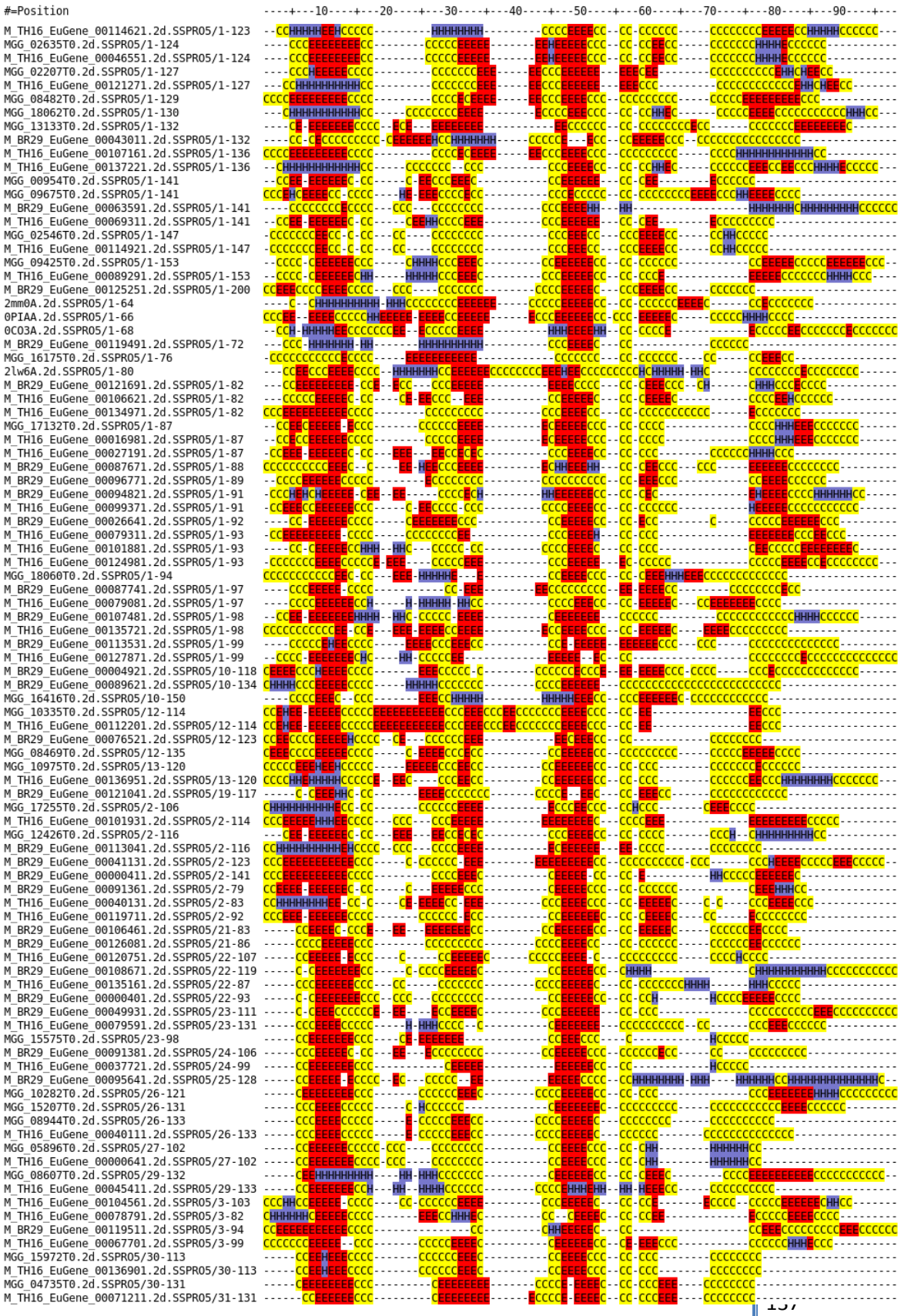
S10 Fig. Prediction of the secondary structure of M. oryzae MAX effectors. The secondary structures of the MAX-effectors from the 70–15, TH16 and BR29 genomes was predicted with SSPRO5 [93].The predictions are shown at the bottom of the figure and are aligned onto the corresponding primary sequence alignment shown at the top of the figure. Sequence identifiers for the secondary structure predictions are suffixed with « .2d.SSPRO5 ». Blue »H », red « E » and yellow « C » correspond respectively to helix, extended sheet and coil predictions. The sequences of the 4 MAX effectors with experimentally determined structures are displayed at the top of the multiple sequence alignment and, for clarity, the alignment positions corresponding to shared gaps in the known structures were removed.



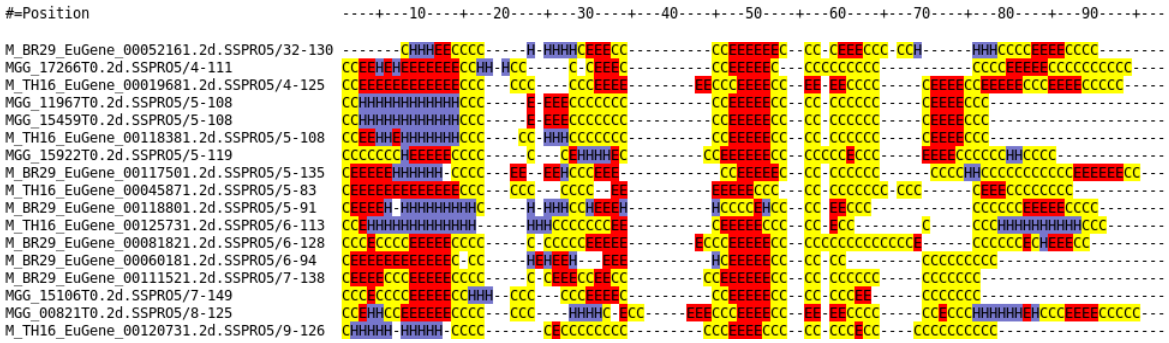
S10 Fig. continued



S10 Fig. continued



S10 Fig. continued

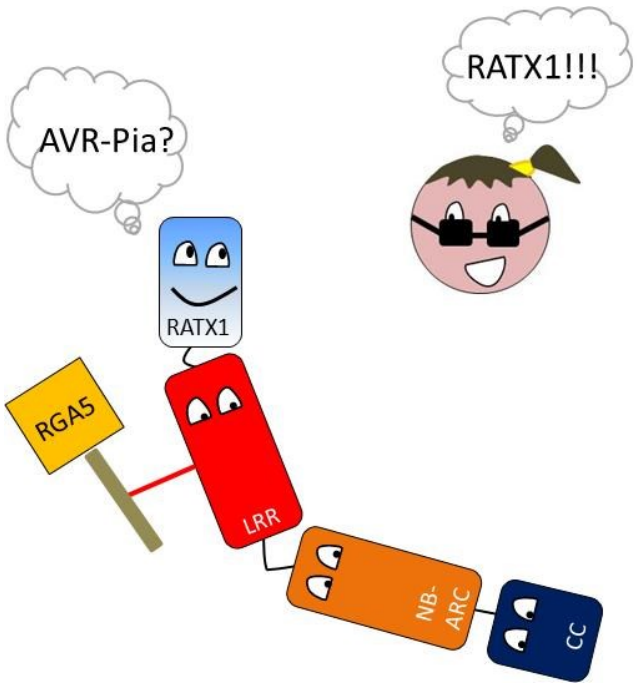


SUMMARY POINTS (CHAPTER I)

- ✓ Structural insights into effector proteins start to reveal common patterns hidden in highly diverse amino acid sequences.
- ✓ MAX-effectors are a large family of effectors characterized by a common β-sandwich structure and unrelated sequences.
- ✓ The common β-sandwich structure of MAX-effectors is specifically found in effectors from Ascomycete fungi and has massively expanded and diversified only in *M. oryzae* and *M. grisea*.
- ✓ MAX-effectors seem to evolve under diversifying selection.
- ✓ MAX-effector candidates could act as virulence effectors since the majority of them are express specifically during biotrophic infection.
- ✓ MAX-effectors present very different surface properties and activities despite their similar topology.
- ✓ The conserved architecture found in MAX-effectors allows effector proteins to change and adapt to new functions as well as to face general constrains such as stability in fungus-host interfaces.

Chapter II

AVR-Pia recognition by the decoy domain of rice the NLR immune receptor RGA5



INTRODUCTION

Since the initial cloning of genes conferring disease resistance in plants and coding for NLR immune receptors, tremendous efforts have been made to better understand the mechanistic basis of their function. One long-term goal in this research is the engineering of NLRs with defined novel specificities that confer durable and broad spectrum disease resistance in economically important crop plants. Unfortunately, until now attempts for resistance engineering have generally failed owing to the complexity of effector recognition as well as NLR signaling. To date, only a few NLRs that recognize a broader range of alleles of the same pathogen effector have been successfully engineered through *in vitro* evolution and rational design. However, the development of NLRs with completely new specificities has not yet been achieved.

In many cases NLRs monitor host proteins, to sense pathogen attack. Hence, an alternative approach in which the host target protein of the pathogen effector is engineered rather than the NLR itself has recently been successfully implemented (Kim et al., 2016). PBS1 from *Arabidopsis thaliana* is a "decoy" kinase protein targeted by the protease effector AvrPphB. The proteolytic cleavage of PBS1 by AvrPphB results in the activation of the NLR protein RPS5. A three amino-acid insertion at the cleavage site of PBS1 protein also results in the activation of RPS5 in absence of effector proteins showing that RPS5 can sense different PBS1 modifications. Based on these findings, the AvrPphB cleavage site in PBS1 was substituted with new sequences that corresponded to the cleavage sites of two different proteases: AvrRpt2 from *Pseudomonas syringae* or NIa from tobacco etch virus (TEV). In both cases, the cleavage of PBS1 resulted in RPS5-mediated resistance to the corresponding pathogen, showing that decoy proteins can be engineered to function in a very specific manner as a target of effector proteins from very different classes of pathogens (Kim et al., 2016).

Work on the rice NLR pair RGA4/RGA5 lead to the development of the integrated decoy model and the demonstration that the integration of potential decoy domains is frequent (Cesari et al., 2014). Key hypothesis of the integrated decoy model were confirmed by functional studies

with the NRLs RRS1 and Pik-1 and large scale searches of NLRs with integrated domains (NLR-IDs) revealed that NLRs occur in roughly 5% of the NLRs of land plants (Sarris et al., 2016; Kroj et al., 2016). Based on these findings, it was proposed that the modular architecture of NLR-IDs presents an alternative and extremely promising opportunity for the engineering of NLRs with novel specificities. Indeed, by changing the integrated decoy and introducing a novel effector trap, an NLR-ID with a novel recognition specificity may be generated. (Cesari et al., 2014; Ellis, 2016). However, to really reach this goal, a much better understanding of the mode of function of NLR-IDs is required.

Previous work had shown that the RATX1/HMA domain in RGA5 establishs physical interactions with the effectors AVR1-CO39 and AVR-Pia *in planta* and in yeast two hybrid assays (Cesari et al. 2013). Furthermore, it was suggested that the physical interaction between the effectors and $RGA5_{RATX1}$ is required for recognition. Indeed, AVR-PiaH3, an AVR-Pia allele that is not recognized in resistant rice plants, does not interact with the RATX1 domain. These results prompted us to go forward in the molecular characterization of the interaction of AVR-Pia and $RGA5_{RATX1}$.

In this chapter we present new insights into the role of integrated decoys in the recognition of effector proteins. In particular, we delimit the interface with which AVR-Pia binds the RGA5 RATX1 domain and we propose a model that explains the role of AVR-Pia-$RGA5_{RATX1}$ binding in RGA5/RGA4-mediated effector recognition.

ARTICLE 2

Recognition of the *Magnaporthe oryzae* effector AVR-Pia by the decoy domain of the rice NLR immune receptor RGA5

Diana Ortiz[a*], Karine de Guillen[b*], Stella Cesari[c], Véronique Chalvon[a], Jérome Gracy[b], André Padilla[b], Thomas Kroj[a,1]

[a] INRA, BGPI, Biology and Genetics of Plant-Pathogen Interactions, Campus International de Baillarguet, Montpellier, France

[b] CNRS UMR 5048, INSERM U1054, Centre de Biochimie Structurale, Université Montpellier,France

[c] CSIRO Agriculture Flagship, GPO Box 1600, Canberra ACT 2601, Australia

* These authors contributed equally to the work

[1] **Corresponding author:** Thomas Kroj, thomas.kroj@supagro.inra.fr

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Thomas Kroj (thomas.kroj@supagro.inra.fr).

Short title: Recognition of the effector AVR-Pia

SUMMARY

Nucleotide-binding domain and leucine-rich repeat proteins (NLRs) are important receptors in plant immunity that allow recognition of pathogen effectors. The rice NLR RGA5 recognizes the *Magnaporthe oryzae* effector AVR-Pia through direct interaction. Here, we gained detailed insights into the molecular and structural bases of AVR-Pia-RGA5 interaction and the role of the RATX1 decoy domain of RGA5. NMR titration combined with *in vitro* and *in vivo* protein-protein interaction analyses identified the AVR-Pia interaction surface that binds to the RATX1 domain. Structure-informed AVR-Pia mutants showed that, although AVR-Pia associates with additional sites in RGA5, binding to the RATX1 domain is necessary for pathogen recognition, but can be of moderate affinity. Therefore, RGA5-mediated resistance is highly resilient to mutations in the effector. We propose a model that explains such robust effector recognition as a consequence, and an advantage, of the combination of integrated decoy domains with additional independent effector-NLR interactions.

INTRODUCTION

Plant disease resistance largely relies on inducible immune responses that are triggered upon receptor-mediated recognition of pathogen molecules and that often involve a localized programmed cell death called the hypersensitive response (HR). Particularly important are cytoplasmic nucleotide-binding oligomerization domain-like receptors (NLR) that present a multi-domain architecture composed of a C-terminal leucine-rich repeat (LRR) domain and a central nucleotide-binding (NB-ARC) domain (Takken and Goverse, 2012; Jacob et al., 2013; Qi and Innes, 2013). In most cases, they carry in addition an N-terminal coiled-coil (CC) or TOLL/interleukin-1 receptor (TIR) domain that have both been reported to mediate NLR homo complex formation and to be crucial for the activation of downstream signalling (Bernoux et al., 2011; Maekawa et al., 2011). Plant NLR proteins specifically recognize pathogen-derived effectors that act inside plant cells (Cui et al., 2015).

Traditionally, both effector recognition and activation of resistance signalling are thought to be mediated by single plant NLRs but recent studies revealed an increasing number of cases where different NLRs cooperate in pathogen recognition and resistance (Eitas and Dangl, 2010; Césari et al., 2014a). Frequently, the genes coding for these paired NLRs occur in a paired, inverted tandem arrangement in the genome. In the few cases that have been investigated in more detail, the NLR pairs seem to act as hetero-complexes where only one of the paired NLRs acts directly in effector recognition while the other is crucial for the activation of downstream signalling (Williams et al., 2014; Césari et al., 2014b). In other cases, helper NLRs that act downstream of several NLRs with different recognition specificities were shown to be required for resistance and pathogen detection (Gabriëls et al., 2007; Bonardi et al., 2011; Wu et al., 2015).

In certain cases, plant NLRs recognize effectors in an indirect manner. They detect either a modification of the effector's host target protein called 'guardee', or modifications of a host protein that mimics the effector target and is called a 'decoy' (van der Hoorn and Kamoun, 2008). In other cases, effectors are recognized in a direct manner by binding, either alone, or in complex with a cofactor that may be a guardee or a decoy, to the NLRs (Takken and Goverse, 2012; Collier and Moffett, 2009). In these cases, the LRR domain plays a crucial role in recognition specificity and was frequently shown to mediate direct effector binding (Ellis et al. 2007; Krasileva et al., 2010; Jia et al., 2000). Alternatively, direct effector recognition can also be mediated by non-canonical domains integrated into NLRs at low frequencies (Kanzaki et al., 2012; Sarris et al., 2015; Maqbool et al., 2015; Le Roux et al., 2015; Césari et al., 2013). Recent work lead to the hypothesis that these highly diverse integrated domains are mimics of effector targets and can therefore be considered as integrated decoy domains (Le Roux et al., 2015; Sarris et al., 2015; Césari et al., 2014a). However, the molecular mechanisms of effector recognition by integrated domains and eventual advantages of this mode of recognition remain largely unknown.

Rice blast, caused by the fungus *Magnaporthe oryzae* is a highly destructive crop disease and a serious threat for food security (Pennisi, 2010; Dean et al., 2012; Skamnioti and Gurr, 2009). NLR-mediated pathogen recognition is the major mechanism in rice blast resistance. Among 25 different blast resistance genes cloned over the last 20 years, 24 code for NLRs (Liu et al., 2014).

A particular feature is that blast-resistance is frequently conferred by paired-NLRs with clustered tandem organisation in the genome. Among them, the NLR pair RGA4/RGA5 conferring *Pi-CO39* and *Pia* resistances has been developed as a model for the molecular understanding of paired NLRs (Okuyama et al., 2011; Césari et al., 2013, 2014b). In this pair, RGA4 acts as a constitutively active disease resistance and cell death inducer that is repressed by RGA5 in the absence of pathogen (Césari et al., 2014b). In addition to its repressor function, RGA5 acts as a receptor for the *M. oryzae* effectors AVR1-CO39 and AVR-Pia. Direct binding of RGA5 to these effectors results in de-repression of RGA4 and activation of resistance signalling. Effector binding involves the unconventional C-terminal domain of RGA5 that is related to ATX1 (RATX1 domain), a heavy metal-associated (HMA) domain protein from *Saccharomyces cerevisiae,* that acts as a cytoplasmic copper chaperone (Césari et al., 2013). Interestingly, the RATX1 domain of RGA5 is dispensable for RGA4 repression and seems exclusively dedicated to effector binding (Césari et al., 2014b). Since the rice RATX1/HMA protein Pi21 is a blast susceptibility factor required for full disease development (Fukuoka et al., 2009), it has been hypothesised that AVR1-CO39 and AVR-Pia target RATX1/HMA proteins for disease development and that the RATX1 domain is an integrated decoy domain (Césari et al., 2014b, 2013). Interestingly, an HMA domain 53% identical to the RGA5 RATX1 domain is also present in another rice NLR, Pik-1 that acts together with the NLR Pik-2 in the specific recognition of the *M.oryzae* effector AVR-Pik. Like in RGA5, this domain acts by binding directly the effector and is crucial for its recognition (Kanzaki et al., 2012). However, on the contrary to the C-terminal RATX1 domain of RGA5, the HMA domain of Pik-1 is located between the CC and NB domains indicating independent integration of the same domains in the two unrelated NLRs (Césari et al., 2013). Recently, the determination of the crystal structure of the AVR-PikD/Pikp-1_{HMA} domain complex allowed the precise identification of the AVR-PikD surface mediating binding to the Pikp-1_{HMA} domain (Maqbool et al., 2015).

Although AVR-Pik, AVR-Pia and AVR1-CO39 do not share sequence similarities, they share a highly similar three dimensional structure characterized by a six β-sandwich fold also present in two other effectors: AvrPiz-t from *M. oryzae* and ToxB from the wheat pathogenic fungus *Pyrenophora tritici repentis* (Zhang et al., 2013; Guillen et al., 2015; Maqbool et al., 2015; Nyarko et al., 2014). The corresponding, structurally related *Magnaporthe* Avr and ToxB effectors were termed MAX effectors. MAX effectors are present in other, sometimes only distantly related phytopathogenic fungi and the MAX effector family underwent strong expansion in *M. oryzae* where it accounts for roughly 10% of the effectors (Guillen et al., 2015). In the present study, we investigated the molecular and structural bases of AVR-Pia recognition by RGA5 with a focus on the role of the RATX1 domain in effector-binding and recognition. We show that AVR-Pia interacts with $RGA5_{RATX1}$ domain through a precise surface that shares some similarity, but also important differences, with the HMA-binding surface of AVR-Pik. We demonstrate that binding to the RATX1 domain is required for effector recognition but that strong reduction in binding strength is tolerated. We also provide evidence that the RATX1 domain is not required for association of AVR-Pia with RGA5 and that it interacts with additional sites in the NLR which could explain the high tolerance of recognition to reduced AVR-Pia-$_{RATX1}$ binding strength. Based on our results, we propose a model illustrating advantages of effector recognition by integrated decoy domains as well as additional simultaneously occurring interactions with NLR receptors.

RESULTS

The F24S and T46N substitutions in the non-recognized AVR-Pia-H3 allele affect surface properties but not structure

We previously described the naturally occurring *AVR-Pia* allele *AVR-Pia-H3* that carries two non-synonymous polymorphisms leading to the F24S and T46N substitutions (Césari et al., 2013). *M. oryzae* isolates carrying the *AVR-Pia-H3* allele are virulent on rice varieties carrying the *Pia* resistance locus and AVR-Pia-H3 does not interact in Y2H assays with the C-terminal part of the rice NLR immune receptor RGA5 containing the RATX1 domain ($RGA5_{C\text{-}ter}$). The NMR structure of AVR-Pia showed that both the F24 and T46 residues are surface-exposed and suggested that the corresponding substitutions only affect AVR-Pia surface properties without major structural

rearrangements (de Guillen et al., 2015). To test this hypothesis, the structures of AVR-Pia-H3 and the single mutants AVR-Pia^{F24S} or AVR-Pia^{T46N} were analyzed by NMR spectroscopy. We performed sequential assignments using ^{15}N-labeled AVR-Pia samples and the ^{13}Cα and ^{13}Cβ assignments were performed using ^{13}C-^{1}H 2D experiments with a ^{13}C-natural abundance simple in D_2O (Supplemental Table 1).

When compared to AVR-Pia wild type, ^{1}H-^{15}N chemical shifts differed more in AVR-Pia-H3 than in AVR-Pia^{F24S} to AVR-Pia^{T46N} single mutants (Figure 1A). The NMR structure of AVR-Pia-H3 proved to be very similar to the structure of AVR-Pia (PDB code 5JHJ) (Figure 1B, supplemental Table 2). The backbone RMSD for superposition of the AVR-Pia and AVR-Pia-H3 structures is 1.53 Å and drops to 0.93 Å when the ß1-ß2 loop is excluded and the superposition starts at residue R23. Like the AVR-Pia wild type protein, AVR-Pia-H3 shows the MAX-effector topology characterized by 6 antiparallel β-strands (Figure 1C). The ^{1}H-^{15}N chemical shift data for AVR-Pia^{F24S} and AVR-Pia^{T46N} indicate that both single mutants probably also keep the MAX-effector fold (Figure 1A). The analysis of the relaxation data indicates that both AVR-Pia^{wt} and AVR-Pia-H3 have rigid structures with average S2 values of 0.8 and similar S2 profiles indicating similar protein dynamics (Supplemental Figure 1). The 3D structure of AVR-Pia-H3 therefore supports that the F24S and T46N substitutions do not result in conformational changes but rather alter AVR-Pia surface properties required for strong interaction with the RATX1 domain of RGA5 ($RGA5_{RATX1}$) and disease resistance activation.

AVR-Pia binds $RGA5_{RATX1}$ with intermediate affinity

To characterize the AVR-Pia/$RGA5_{RATX1}$ interaction, *in vitro* binding assays with recombinant $RGA5_{RATX1}$ and AVR-Pia or AVR-Pia-H3 were performed using isothermal calorimetry (ITC). For AVR-Pia, specific and direct binding to $RGA5_{RATX1}$ with a one site model and a Kd of 7,8 µM was detected (Figure 2A). For AVR-Pia-H3, no binding was detected under identical conditions, suggesting that its affinity to $RGA5_{RATX1}$ is at least 10 fold smaller than that of AVR-Pia.

β-strands 2 and 3 and residues R23, F24, E56 and E58 constitute a candidate $RGA5_{RATX1}$-interaction surface in AVR-Pia

To test the hypothesis that the residues F24 and T46 are part of the AVR-Pia surface mediating

direct contacts to $RGA5_{RATX1}$ and to identify other residues in direct contact to $RGA5_{RATX1}$ or located in the close vicinity of the binding interface, NMR-titration experiments were performed. This technique consists in recording the ^{1}H-^{15}N-HSQC NMR spectra of ^{15}N-labelled AVR-Pia in the presence of increasing amounts of unlabeled $RGA5_{RATX1}$. When protein-protein binding occurs, it modifies the chemical environment of the amino acids located on the binding surface and eventually mediating the binding. This results in a change of the chemical shift in NMR experiments. Depending on the transition rate of the complex formation, expressed by the exchange rate constant kex, and the chemical shift difference Δω between the unbound and bound states (Δω difference between the resonance frequencies of the exchanging sites), different exchange regimes occur. NMR titration shows that the AVR-Pia-$RGA5_{RATX1}$ complex is in slow exchange with kex << Δω since separate resonances appear for individual species (bound and unbound states) (Figure S2A). Residues with important chemical shift changes between free AVR-Pia (R=0) and AVR-Pia bound to $RGA5_{RATX1}$ (molar ratio R=2) were almost exclusively surface-exposed and located in a region formed essentially by β-strands 2 and 3 and including, in addition, residues R23 and F24 from β-strand 1 as well as E56 and E58 from β-strand 4 (Figures 2B and C). No peaks were observed for residues Y27, V37, Y41, I44 and T51 in the complex. This candidate interaction surface largely overlaps with an extended, solvent-exposed patch of hydrophobic residues formed by F24, V26 and Y28 in β1, V37, L38 and Y41 in β2, and Y85 in β6. The residues on the other side of the AVR-Pia structure were not shifted in the NMR titration and seem therefore not involved in the interaction with $RGA5_{RATX1}$ (Figure 2D).

Two exceptions are E83, which probably senses a perturbation of the residue Y41 that is close in space, and the I69 residue, which may be involved in local conformational rearrangement of the short β5 strand. $RGA5_{RATX1}$-titration experiments were also performed with ^{15}N-labelled AVR-Pia-H3 that shows no binding in Y2H (Césari et al., 2013) and ITC analysis (Figure 2A). Spectral perturbations were strongly reduced and only few and limited changes of chemical shifts occurred when titrating AVR-Pia-H3 with $RGA5_{RATX1}$ (Figure 2B and E and Figure S2B). Signals for the R23, S24, V42, R43 and E83 residues were still observed at the end of the titration while they were mostly lost at a molar ratio of 0.5 in the case of AVR-Pia (Figure S2). Similarly, signals for E58, V59 and T47 were much less perturbed. Nevertheless, the peaks for Y41, N46 and T51

were also perturbed indicating a weak residual interaction between $RGA5_{RATX1}$ and AVR-Pia-H3 (Figure 2B and E and Supplemental Figure 2).

In summary, NMR titration identified a candidate interaction surface formed by β-strands 2 and 3 and including, in addition, residues R23, F24, E56 and E58 (Figure 2B). This surface overlaps extensively with an extended hydrophobic patch on the AVR-Pia surface that contains F24 and has T46 on its border and that may be crucial for $RGA5_{RATX1}$-binding.

Y2H experiments with structure-informed AVR-Pia mutants confirm an important role of the candidate interaction surface in RGA5C-ter-binding

To test whether the AVR-Pia candidate interaction surface identified *in vitro* mediates binding to RGA5C-ter *in vivo*, we performed yeast two-hybrid (Y2H) assays using AVR-Pia variants bearing point mutations in critical residues identified by NMR titration. Individual surface-exposed hydrophobic (M40, Y41, Y85) or charged (R23, D29, R36, E56, E58) amino acids, located in or at the border of the candidate interaction surface, were replaced by alanine. In addition, naturally occurring AVR-Pia polymorphisms located within the candidate interaction surface were tested: F24S and T46N from AVR-Pia-H3 and R43G from AVR-Pia-H2 identified in *M. oryzae* isolates pathogenic on *Setaria* species (Supplemental Figure 3A) (Césari et al., 2013). As controls, mutants where surface-exposed charged residues located outside the candidate interaction surface are replaced by alanine (D63A, K67A, K74A, D78A) were generated (Figure 3A).

As previously reported, yeasts co-expressing BD-AVR-Pia and AD-RGA5C-ter or AD-RGA5C-ter and BD-RGA5C-ter grew on selective medium indicating physical binding between AVR-Pia and RGA5C- ter, and homo-interaction of the RGA5C-ter domain (Figure 3A) (Cesari et al., 2013; Cesari et al.,2014). Yeasts co-expressing AD-$RGA5_{C\text{-}ter}$ and AVR-Pia^{F24S}, AVR-Pia^{R43G}, AVR-Pia^{R36A} fused to the BD domain did not grow on selective medium indicating that these mutations abolish binding to RGA5C-ter. Isolates expressing BD-fusions of AVR-Pia variant carrying the mutation R23A, D29A, T46N, E58A or D63A showed reduced growth compared to wild type BD-AVR-Pia indicating that these mutations also affect AVR-Pia-RGA5C-ter interaction. AVR-Pia^{Y41A}, AVR-Pia^{E56A}, AVR-Pia^{K67A}, AVR-Pia^{K74A}, AVR-Pia^{D78A}, AVR-Pia^{M40A} and AVR-Pia^{Y85A} isolates showed stronger growth. All BD-AVR-Pia variants were expressed at similar levels as the wild-

type BD-AVR-Pia (Figure 3B). Taken together, these Y2H data show that the replacement of all charged amino acids in the interaction surface, with the exception of E56, either abolish or reduce binding of AVR-Pia to RGA5$^{C\text{-}ter}$ while exchanging hydrophobic residues within the interaction surface seems to abolish the interaction in the case of F24S, or to increase the binding in the case of M40A and Y85A.

To rule out that reduced binding of AVR-Pia mutants to RGA5$^{C\text{-}ter}$ is due to major changes in protein structure, the AVR-Pia mutants R23A, D29A, R36A, R43G and E58A were expressed in *E. coli*, purified to homogeneity and analyzed by ^{1}H-1D-NMR experiments (Supplemental Figure 3B). All mutant proteins showed similar spectra as AVR-Pia wild-type, indicating that they are well structured and only locally disturbed. Recombinant AVR-Pia^{D63A} could not be expressed. Taken together, these results suggest that most residues of the AVR-Pia interaction surface identified by NMR titration play an important role in RGA5$^{C\text{-}ter}$-binding.

Co-IP experiments identify key residues in the AVR-Pia interaction surface that are crucial for RGA5$_{RATX1}$-binding *in planta*

To investigate the role of the AVR-Pia interaction surface in *in planta* binding to RGA5$^{C\text{-}ter}$, co-immunoprecipitation (co-IP) experiments were performed. HA-tagged RGA5$^{C\text{-}ter}$ and YFP-tagged AVR-Pia mutants with reduced binding in Y2H were co-expressed in *N. benthamiana* by *Agrobacterium tumefaciens*-mediated transient transformation. In addition, AVR-Pia^{M40A} that, according to Y2H experiments, has increased affinity for RGA5$^{C\text{-}ter}$ was also analyzed. As a negative control, a YFP fusion of the cytoplasmic *M. oryzae* effector PWL2 was used (Khang et al., 2010). Immunoblotting using anti-GFP and anti-HA antibodies showed proper expression of all proteins (Figure 4). However, AVR-Pia mutants with reduced binding to RGA5$^{C\text{-}ter}$ in Y2H reproducibly accumulated at lower levels than AVR-Pia^{wt} while AVR-Pia^{M40A} was expressed at similar levels (Figure 4). All YFP fusion proteins were efficiently precipitated with anti-GFP antibodies but only AVR-Pia^{M40A} co-precipitated RGA5$^{C\text{-}ter}$ as strong as AVR-Pia^{wt}. The other mutants showed various degrees of impairment ranging from slightly (AVR-Pia^{R23A} AVR-Pia^{E58A}, AVR-Pia^{D63A}) to strongly reduced (AVR-Pia^{D29A} AVR-Pia^{R36A}, AVR-Pia^{R43G}) or even completely abolished RGA5$^{C\text{-}ter}$ co-precipitation (AVR-Pia^{F24S}) (Figure 4A). The specificity of the interactions was confirmed with PWL2 that does not interact with RGA5$^{C\text{-}ter}$.

It has previously been shown that the interaction of AVR-Pia with RGA5$_{C\text{-}ter}$ relies on interaction with the RATX1 domain (RGA5$_{RATX1}$) (César et al., 2013). To verify that interaction specificities of the AVR-Pia mutants with RGA5$_{C\text{-}ter}$ correlates with interaction strength with RGA5$_{RATX1}$, co-IP experiments were performed using HA-tagged RGA5$_{RATX1}$. AVR-Pia^{wt} and AVR-Pia^{M40A} strongly co-precipitated HA-RGA5$_{RATX1}$, while the other mutants showed reduced (R23A and D63A), strongly reduced (D29A, R36A and E58A) or no co-precipitation of RGA5$_{RATX1}$ (F24S and R43G) (Figure 4B). Taken together, these data indicate that AVR-Pia^{wt} and AVR-Pia^{M40A} RGA5$_{C\text{-}ter}$ and RGA5$_{RATX1}$, while mutants affected in direct binding to RGA5$_{C\text{-}ter}$ in Y2H showed reduced association with RGA5$_{C\text{-}ter}$ and RGA5$_{RATX1}$ *in planta*. Complete absence of association with RGA5$_{RATX1}$ for AVR-Pia^{F24S} and AVR-Pia^{R43G}, both *in planta* and in Y2H, indicates a crucial role of these residues in the binding interface and suggests that they are pivotal for AVR-Pia recognition.

Direct binding to the RATX1 domain is required for AVR-Pia recognition

To determine the role of the RATX1-binding surface of AVR-Pia in specific recognition by the RGA4/RGA5 pair, AVR-Pia mutants were co-expressed in *N. benthamiana* with RGA4/RGA5 and cell death activation was monitored. Since tagged versions of AVR-Pia proved inactive in this assay, un-tagged AVR-Pia mutants were used. AVR-Pia mutants with wildtype binding to RGA5$_{RATX1}$ induced cell death indicating they are recognized by RGA5/RGA4 (Supplemental Figure 4A and B).

Weakly or non-binding mutants lost cell death inducing activity but were also less abundant than AVR-Pia^{wt} or recognized AVR-Pia mutants (Supplemental Figure 4A, B and C). Therefore, no clear conclusions can be drawn for these mutants since lack of recognition may be due to low protein abundance. Similar differences in the protein level of AVR-Pia mutants were previously observed with YFP-tagged variants expressed in *N. benthamiana* (Figure 4) but not upon expression in *E. coli* or yeast (Figure 3B). Therefore, differences in the accumulation of AVR-Pia variants seem not related to an intrinsic destabilization of these proteins but rather result from reduced stability in *N. benthamiana*.

Since transient heterologous experiments failed to determine the importance of the binding of AVR-Pia to $RGA5_{RATX1}$ for recognition and disease resistance, the biological activity of AVR-Pia mutants was assayed in the homologous rice/*M. oryzae* system. Transgenic *M. oryzae* isolates were generated that carried the different mutant alleles under the control of the constitutive *RP27* promoter ($RP27_{Pro}$) (Bourett et al., 2002). As a control, transgenic Guy11 isolates carrying a $RP27_{Pro}$*:mRFP* construct were generated and proved to be fully virulent (Figures 5 and Supplemental Figure 5B). For three different PCR-validated transgenic isolates per construct, the accumulation of AVR-Pia variants was verified in culture filtrates by immunoblotting with anti AVR-Pia antibodies (Supplemental Figure 5A). All AVR-Pia mutants were detected in at least one transgenic isolate except AVR-Pia^{D63A} that may be instable in *M. oryzae.* For AVR-Pia^{D29A} and AVR-Pia^{E58A}, only two and one isolate expressed the mutant protein (Figure S5A).

The transgenic isolates were analyzed on the rice cultivars Kitaake carrying the *Pia* locus and Maratelli lacking *Pia*. All isolates were highly virulent on Maratelli indicating that they were not affected in virulence (Supplemental Figure 5B). On Kitaake plants, the isolates expressing AVR-Pia^{wt}, AVR-Pia^{R23A}, AVR-Pia^{D29A}, AVR-Pia^{R36A} or AVR-Pia^{E58A} were completely avirulent and produced either no symptoms or small HR lesions characteristic of resistance (Figure 5 and Supplemental Figure 5B). This indicates that these AVR-Pia variants are fully active and recognized by RGA4/RGA5. Consistent with the absence of protein expression, AVR-Pia^{D63A} isolates did not induce resistance and were fully virulent on Kitaake plants. Isolates producing AVR-Pia^{R43G} were partially virulent and formed disease lesions characterized by a grey center that were however smaller and less frequent than those provoked by the control mRFP isolates. Isolates expressing AVR-Pia^{F24S} were highly virulent on Kitaake and produced large numbers of disease lesions (Figures 5 and Supplemental Figure5B).

Taken together, these results indicate that interaction of AVR-Pia with the $RGA5_{RATX1}$ domain is required for recognition but that a reduction of this interaction as in AVR-Pia^{R23A}, AVR-Pia^{D29A}, or AVR-Pia^{E58A} does not impair recognition. Only the R43G and F24S polymorphisms that abolished $RGA5_{RATX1}$ interaction both *in planta* and in yeast affected AVR-Pia recognition, with AVR-Pia^{F24S} being completely inactive.

AVR-Pia interacts with RGA5 outside of the RATX1 domain

The high resilience of *RGA4/RGA5*-mediated AVR-Pia recognition to reduction of AVR-Pia-$RGA5_{RATX1}$ interaction strength suggested that AVR-Pia might interact with additional sites in RGA5. To test this hypothesis, *in planta* association of the AVR-Pia mutants with the RGA5 full-length protein was assayed by co-IP. All AVR-Pia mutants, including AVR-Pia^{F24S} and AVR-Pia^{R43G} co-precipitated RGA5 as efficiently as AVR-Pia^{wt} (Figure 6A). This indicates that lack of binding to $RGA5_{RATX1}$ does not abolish association with RGA5. To test whether binding of AVR-Pia to RGA5 is truly independent of the RATX1 domain, association of AVR-Pia with an RGA5 construct lacking the RATX1 domain ($RGA5_{\Delta RATX1}$) was tested by co-IP. All AVR-Pia variants co-precipitated $RGA5_{\Delta RATX1}$ (Figure 6B) and AVR-Pia mutants with reduced or no binding to $RGA5_{RATX1}$ interacted as strongly with $RGA5_{\Delta RATX1}$ as AVR-Pia^{wt} demonstrating that the RATX1 domain is not necessary for formation of RGA5/AVR-Pia complexes. These results suggest that AVR-Pia interacts with additional sites in RGA5 outside of the RATX1 domain and that the region of AVR-Pia that mediates interaction with $RGA5_{\Delta RATX1}$ lies outside of the RATX1-binding surface.

It was previously shown that $RGA5_{\Delta RATX1}$ inhibits RGA4-triggered cell death and that therefore the RATX1 domain is not required for RGA5-mediated repression of RGA4 (Césari et al., 2014b). Since AVR-Pia still associates with $RGA5_{\Delta RATX1}$ *in planta*, we tested whether AVR-Pia would be recognized by $RGA5_{\Delta RATX1}$/RGA4 and trigger cell death independently of the RATX1 domain. Neither co-expression of RGA4, $RGA5_{\Delta RATX1}$ and AVR-Pia, nor expression of these three proteins together with the isolated RATX1 domain triggered cell death (Figure S6). This indicates that interaction of AVR-Pia with regions outside of the RATX1 domain is not sufficient to release RGA5-mediated RGA4 repression and further confirms that binding of AVR-Pia to the $RGA5_{RATX1}$ is required for de-repression of RGA4. In addition, these results suggest that AVR-Pia has to interact with the RATX1 domain in the context of the full length RGA5 protein since an isolated RATX1 domain does not complement $RGA5_{\Delta RATX1}$ for AVR-Pia recognition

DISCUSSION

Identification of a RGA5$_{RATX1}$-binding surface in AVR-Pia

In this study, we provide evidence that AVR-Pia interacts with the RATX1 domain of RGA5 through a precise binding surface consisting of β-strands 2 and 3, residues R23 and F24 from β- strand 1, and residues E56 and E58 from β-strand 4 (Figure 2C). This interaction surface, identified by NMR titration experiments with recombinant AVR-Pia and the RATX1 domain, was confirmed by mutant analysis. Indeed, replacement of residues R23, F24, D29, R43, T46 or E58 strongly reduced or abolished binding to RGA5$_{RATX1}$ in Y2H and Co-IP experiments while replacement of residues M40 and Y85 increased interaction in Y2H (Figures 3 and 4). Substitutions outside of the candidate interaction surface had no impact on binding with the exception of residues R36 and D63. Residue R36 is located in the loop joining β1 and β2 and might also be involved in RATX1-binding. Alternatively, it may play an important role in defining the positions of β- strands 1, 2 and 6 through the salt bridge it forms with residue E83 in β6. The D63A polymorphism seems to destabilize the overall structure since AVR-Pia^{D63A} could not be expressed in *E. coli* or *M. oryzae*. Actually, D63 seems important to structure the loop ß4-ß5 as its side-chain carboxyl group forms a hydrogen bond with the side-chain amid group of N65. This may be required to position properly cysteine C66 that forms a disulfide bridge with C25 linking the two β sheets β1, β2, β6 and β3, β4, β5 and thereby impact global folding.

AVR-Pia and AVR-PikD have distinct RATX1/HMA-binding surfaces that are situated at similar positions

The 3 dimensional structure of a AVR-PikD–Pikp-1$_{HMA}$ complex was determined by crystallography and showed that, like the formation of the AVR-Pia-RGA5$_{RATX1}$ complex, binding of AVR-PikD to Pikp-1$_{HMA}$ involves β-strands 2 and 3 (Supplemental Figure 7A) (Maqbool et al., 2015). However, in AVR-PikD, the residues of β-strand 2 that are crucial for Pikp-1$_{HMA}$-binding, R64 and D66, are charged and establish, respectively, hydrogen bonds and salt bridge interactions. In contrast, in AVR-Pia, surface-exposed residues of β-strand 2 are hydrophobic and probably establish hydrophobic interactions (Figure 2C). In addition, unlike AVR-Pia, AVR- PikD possesses an N-terminal extension of 32 amino acids that is crucial for physical binding to Pikp-1$_{HMA}$ and

recognition by Pikp-1/Pikp-2 (Supplemental Figure7B). In particular, residue H46 from this extension establishes important interactions with matching residues in Pikp-1_{HMA}. These interactions are necessary for binding and, together with the neighboring residues P47 and G48, for matching specificities with alleles of Pikp-1 (Kanzaki et al., 2012; Maqbool et al., 2015). These residues are missing in AVR-Pia but similarly important interactions are established with the amino acid F24 from the very hydrophobic β-strand 1. Therefore, recognition of the sequence-unrelated, but structurally similar, effectors AVR-Pia and AVR-PikD seems to involve similar structural elements but relies on distinct and highly specific mechanisms.

Whether the effector interaction surfaces of the RATX1/HMA domains of RGA5 and Pikp-1 are similar or completely different remains an open question. Crystal structures show that Pikp- 1_{HMA}, has a typical HMA α/β- sandwich fold composed of two α-helices and a 4-stranded anti- parallel β-sheet, that mediates interaction with AVR-PikD (Supplemental Figure 7C) (Maqbool et al., 2015). We used molecular modeling to evaluate whether AVR-Pia-binding may involve similar regions in $RGA5_{RATX1}$ but no consensus docking model could be generated for the AVR-Pia-$RGA5_{RATX1}$ complex (Supplemental Figure 7C). Interestingly, none of the docking models predicted an interaction surface in $RGA5_{RATX1}$ similar to the effector-binding surface of Pikp- 1_{HMA}. This suggests that the $RGA5_{RATX1}$–AVR-Pia complex differs significantly from the Pikp-1_{HMA}–AVR-Pik-D complex.

Taken together, recognition of the structurally similar MAX effectors AVR-Pia and AVR-Pik by independently acquired NLR-integrated HMA domains seems to rely on distinct molecular mechanisms. Future work is required to test this hypothesis through functional studies of the Pikp-1_{HMA} interaction surface identified by structural analysis and the identification of the surface that mediates effector binding in $RGA5_{RATX1}$.

Binding of AVR-Pia to the integrated RATX1 domain is required for recognition but of moderate affinity

The mutants, AVR-Pia^{F24S} and AVR-Pia^{R43A}, showed drastically reduced $RGA5_{RATX1}$-binding and triggered, respectively, no or reduced resistance indicating that the AVR-Pia-$RGA5_{RATX1}$ interaction is required for RGA4/RGA5-mediated recognition. The presence of these

polymorphisms in naturally occurring AVR-Pia alleles (Ribot et al., 2013) suggests that *AVR-Pia* is undergoing selection for mutations in the RATX1-interaction surface and escape from RGA4/RGA5-mediated recognition.

These results therefore provide further support for a crucial role of non-conventional, integrated decoy domains in effector recognition and NLR specificity. However, we also found high resilience of AVR-Pia recognition to a reduction in $RGA5_{RATX1}$- binding strength since weakly-binding AVR-Pia mutants were still able to trigger resistance. Similar observation was made in the case of AVR-PikD, the only other case where the affinity of an effector to the integrated decoy domain of its NLR receptor has been determined (Maqbool et al., 2015). Indeed, $AVR\text{-}PikD^{A67D}$ and $AVR\text{-}PikD^{P47A\text{-}G48D}$ mutants showed drastically reduced binding to $Pikp\text{-}1_{HMA}$ but were nevertheless perfectly well recognized by Pik-1/Pik-2.

A possible explanation for this tolerance to a reduction in the affinity between effectors and integrated decoys could be that effectors interact with multiple independent sites in NLR receptors. Indeed, our study suggests that, besides the RATX1 domain, AVR-Pia interacts with other, not yet defined, regions in RGA5. In the simplest case, this interaction relies on direct physical binding, but since it was solely detected by co-IP experiments, it cannot be excluded to be indirect and to involve additional co-factors. This interaction seems mediated by other AVR-Pia surfaces than those involved in $RGA5_{RATX1}$-binding since mutants with reduced binding to $RGA5_{RATX1}$ are not affected in interaction with $RGA5_{\Delta RATX1}$. Since the RATX1 domain is covalently linked to the rest of the RGA5 receptor, AVR-Pia-binding to these other sites has the potential to increase the overall effector binding affinity to RGA5 despite the low affinity binding to $RGA5_{RATX1}$ (K_d=7µM). In this context, further mutation-induced reduction of AVR-Pia-affinity toward the RATX1 domain may not have a dramatic effect unless it completely abolishes AVR- Pia/$RGA5_{RATX1}$ interaction. This situation highlights an advantage of the integration of the decoy domain into the NLR receptor over a situation where the decoy is a separate molecule and has to bind to the effector before subsequent binding to the NLR receptor. In the latter case, low affinity of the effector-decoy interaction leads to drastically reduced receptor occupancy and renders the corresponding resistance very vulnerable to effector mutations affecting decoy binding.

Interaction of effectors with multiple independent sites is a hallmark of NLR receptor activation

Effector recognition by RGA4/RGA5 differs from other well-studied NLR models. Indeed, RGA5 has no inherent signaling activity and functions, on the one hand, by repressing RGA4 signaling activity and, on the other, by releasing repression upon AVR-Pia binding (Césari et al., 2014a). Interestingly, the RATX1 domain is only required for de-repression and not for repression (Supplemental Figure 6) (Césari et al., 2014).

Providing the RATX1 domain separately in the presence of $RGA5_{\Delta RATX1}$ and AVR-Pia does not relieve repression despite the fact that AVR-Pia interacts with the separate partners, $RGA5_{\Delta RATX1}$ and $RGA5_{RATX1}$. To explain this result, we propose the hypothesis that simultaneous binding of AVR-Pia to different sites in RGA5, including the RATX1 domain, is required to trap RGA5 in a conformation unable to repress RGA4 (Figure 7). That effectors have to establish simultaneously several independent interactions with NLRs or NLRS and co-factors to be recognized and trigger resistance has been frequently observed with effectors from various origins (Collier and Moffett, 2009). In the *Pseudomonas syringae* effector AvrRPS4 two different surface areas on opposite and distant sites of the molecule are required for recognition by the RRS1/RPS4 pair (Sohn et al., 2012). One of these sites is crucial for binding to the integrated WRKY decoy domain of RRS1 while the other seems to interact with other not yet identified regions in RRS1 (Sarris et al., 2015). Similarly, recognition of the *Hyaloperonospora arabidopsidis* effector ATR1-EMOY2 by the NLRs RPP1-NdA or RPP1-WsB from *A. thaliana* relies on two different surface areas from two different domains and on opposite sides of the molecule suggesting simultaneous interaction with independent binding sides in RPP1-NdA and RPP1WsB (Chou et al., 2011; Steinbrenner et al., 2015). Also in NLRs that recognize effector-co- factor complexes, simultaneous binding of these complexes to different parts of the NLR, generally involving the N-terminus and the LRR have been frequently described (Collier and Moffett, 2009). Therefore, we propose the hypothesis that effectors or effector-co-factor complexes forcing or trapping NLRs in an activated state by simultaneously binding to multiple binding sites and inducing or stabilizing by this major conformational changes is a widespread mechanism in NLR activation and in particular NLRs with integrated domains. Future, structural and functional analysis will be necessary to test this model and elucidate in more detail how activation occurs at the molecular level.

METHODS

Growth conditions of plants and fungi and infection assays

N. benthamiana plants were grown in a growth chamber at 22°C with a 16-h light period. Rice plants (Oryza sativa) were grown as described (Faivre-Rampant et al., 2008). Transgenic *M. oryzae* GUY 11 strains were grown at 25°C during 5 days on rice flour agar for spore production (Berruyer et al., 2003) and in Tanaka complete culture medium (Villalba et al., 2008) agitated at 60 rpm and 25°C during 5 days for liquid culture.

For the analysis of interaction phenotypes, a suspension of *M. oryzae* conidiospores in water with 0.1% of gelatin and adjusted to 5 x 104 spores ml^{-1} was sprayed on the leaves of 3-week- old rice plants (Berruyer et al., 2003). Symptoms were analyzed 7 days after inoculation on the youngest leave that was fully expanded at the time of inoculation. For quantitative analysis leasons were classified and counted; resistant lesions, visible as small brown spots (type 1); weakly susceptible/partially resistant lesions characterized by a pronounced brown border and a small grey centre (type 2); fully susceptible lesions characterized by a large grey centre (type 3).

Constructs

Plasmids were generated by Gateway cloning (ThermoFisher, Waltham, USA), restriction/ligation, site-directed mutagenesis using the QuickChange Lightning kit (Agilent, Santa Clara, USA) or gap-repair cloning in yeast (Bruno et al., 2004). Gateway entry clones were generated using the pDONR207 plasmid (ThermoFisher, Waltham, USA). Gateway destination vectors were modified pBIN19 plasmids for expression of tagged proteins in *N. benthamiana* (Césari et al. 2013) or modified pGAD-T7 or pGBK-T7 plasmids (Clontech, Mountain View, USA) for yeast two hybrid experiments (Bernoux et al., 2011). For protein expression, the pET-15b vector (Merck-Millipore, Darmstadt Germany) was used. For *M. oryzae* transformation constructs were based on the pDL02 plasmid (Bruno et al., 2004). For details on PCR and mutagenesis primers and generation of plasmid refer to Supplemental Tables 3 and 4.

Nuclear magnetic resonance spectroscopy

Spectra were acquired on a 700 MHz Avance Bruker spectrometer equipped with triple- resonance

(^{1}H, ^{15}N, ^{13}C) z-gradient cryo-probe at 305 K. Experiments were recorded using the TOPSPIN pulse sequence library (v. 2.1) (Supplemental Table 1). All spectra are referenced to the internal reference DSS for the ^{1}H dimension and indirectly referenced for the ^{15}N and ^{13}C dimensions (Wishart et al., 1995). Spectra were processed using Topspin (v. 3.2) and analyzed using strip-plots with Cindy in house software and CCPN (Vranken et al., 2005) [analysis v 2.3]. The ^{15}N and ^{13}C assignments were derived from the 2D and 3D spectra at 700 MHz listed in Table S1.

NMR titration

For the assignments, protein samples (1mM) in 20 mM potassium-sodium phosphate, pH 5.4 and 150 mM NaCl, were used. For the titrations of ^{15}N-labeled AVR-Pia proteins, different samples with constant concentrations of AVR-Pia WT or H3 (50µM) and various concentrations of unlabeled RATX1 (ratios 2:1, 1:1, 0.5:1, 0.25:1, 0:1 for the reference) were prepared. HSQC spectra were recorded at 305K on a Bruker Avance 700 MHz spectrometer. Chemical shift differences were measured from the HSQC spectra of AVR-Pia or AVR-H3 alone and the AVR-RATX1 complex at R = 2. They are reported as Hamming distance weighted by the magnetogyric ratios (Schumann et al., 2007).

Coimmunoprecipitation and Yeast two hybrid interaction assays

Protein-protein interaction analyses by co-immunoprecipitation were performed with protein extracts from *N.benthamina* leaf discs harvested 2 days after *Agrobacterium* infiltration (Césari et al., 2013). For the interaction of AVR-Pia variants with RGA5$_{C\text{-}ter}$ and RGA5$_{RATX1}$, 5 leaf disks per sample were homogenized in extraction buffer (25mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 10 mM DTT, 1 mM PMSF, 0.1% IGEPAL CA-630 [NP-40]), supplemented with complete protease inhibitor cocktail (Roche) and polyvinylpolypyrrolidone (PVPP 0.5%). After 2 centrifugations (30 minutes, 15000g) 5 uL magnetic GFP-trap_M beads ((Chromotek) per sample washed two times with protein extraction buffer (without PVPP) were added to 500 µL protein extract and incubated with gentle rotation for 2 h at 4°C. Beads were separated and washed three times with 600 µL of protein extraction buffer (without PVVP).

For the interaction of AVR-Pia variants and RGA5 or RGA5$_{\Delta RATX1}$ a modified protein extraction buffer was used (50mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 10 mM DTT, 1 mM PMSF,

1.0% IGEPAL CA-630 [NP-40], 0.5% sodium deoxycholate, and 0.1% SDS, supplemented with a complete protease inhibitor cocktail (Roche) and polyvinylpolypyrrolidone 0.5%). And co-IP was performed with 8 uL of agarose GFP_trap_A suspension (Chromotek) and four washes with the modified protein extraction buffer.

Bound proteins were eluted by boiling for 10 min at 70°C in 50 uL of Nupage sample buffer, separated by polyacrylamide gel electrophoresis using NuPAGE 4-12% gels (Invitrogen, Carlsbad, CA, USA), transferred to nitrocellulose membrane (Millipore), and analyzed by immunoblotting. For immunodetection of proteins, rat anti-HA-horseradish peroxidase (clone 3F10 ; Roche) or mouse anti-GFP (Roche) and goat anti-mouse-horseradishperoxidase (Sigma-Aldrich) were used in combination with the Immobilon western kit (Millipore).

Binding domain (BD) fusions of AVR-Pia variants in pGBKT7-53 and activation domain (AD) fusions of RGA5C-ter in pGADT7 were transformed in Gold and Y187 yeast strain respectively. Interactions assays were performed according to the Matchmaker Gold yeast two-hybrid system protocol (Clontech).

Transient Protein Expression and HR assays in *N. benthamiana*

For agroinfiltration in *N. benthamiana* pBIN19 binary vectors containing either AVR-Pia, PWL2 or RGA5 variants were transformed into *Agrobacterium tumefaciens* strains GV3101 by electroporation. Individual clones were selected and grown in Luria-Bertani liquid medium containing 50 mg ml-1 rifampicin, 15 mg ml-1 gentamycin, and 50 mg ml-1 kanamycin at 28°C for 24h before agroinfiltration. Co-inoculation mixtures adjusted to an OD600 of 1.0 were infiltrated in 4 weeks old *N. benthamiana* plants. The infiltrated plants were incubated for 48 or 96 h in growth chambers under controlled conditions for coimmunoprecipitations or cell death assays respectively. Three days post infiltration, *N. benthamiana* leaves were scanned using a Typhoon FLA9000 fluorescence scanner (GE Healthcare) with excitation at 635 nm and a long- pass red filter (LPR-665 nm) to evaluate the HR response as a lack of red chlorophyll fluorescence.

Accesion Numbers

Sequence data from this article correspond to those previously published by (Cesari, et al., 2013) and can be found in the the GenBank/EMBL databases under the following accession numbers: AVR-Pia (AB498873), AVR-Pia-H3 (KC777366), PWL2 (U26313), RGA4 (AB604622), Sasanishiki RGA5-A (AB604627), Sasanishiki RGA5-B (KC777365). The PDB accession number for the AVR-Pia_H3 structure is 5JHJ.

Author contributions

Conceptualization, T.K and A.P.; Investigation, D.O., J.G., K.G, V.C. and A.P.; Writing – Original Draft, D.O, A.P and T.K.; Writing – Review & Editing, D.O, A.P, S.C. and T.K..; Funding Acquisition, T.K and A.P.; Resources, D.O. and S.C.; Supervision, T.K. and A.P. T.K and A.P designed the research. D.O, S.C, K.G, V.C and A.P performed the research. A.P and K.G performed NMR analyses. D.O and S.C generated vectors. D.O and V.C produced transgenic *M. oryzae* strains. D.O performed protein-protein interaction analyses. D.O, J.G, K.G, A.P and T.K analyzed data. D.O, TK and AP wrote the article

Acknowledgments

This work was supported by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-0 and the ANR project Immunereceptor (ANR-15-CE20-0007). D. Ortiz-Vallejo was supported by a PhD grant from the ministry of research of Colombia (Colciencias). This work benefited from interactions promoted by COST Action FA 1208 https://www.cost-sustain.org. Christian Roumestand is acknowledged for fruitful discussions.

REFERENCES

Bernoux, M., Ve, T., Williams, S., Warren, C., Hatters, D., Valkov, E., Zhang, X., Ellis, J.G., Kobe, B., and Dodds, P.N. (2011). Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. Cell Host Microbe **9**: 200–11.

Bonardi, V., Tang, S., Stallmann, A., Roberts, M., Cherkis, K., and Dangl, J.L. (2011). Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. Proc. Natl. Acad. Sci. U. S. A. **108**: 16463–8.

Bourett, T.M., Sweigard, J. a, Czymmek, K.J., Carroll, A., and Howard, R.J. (2002). Reef coral fluorescent proteins for visualizing fungal pathogens. Fungal Genet. Biol. **37**: 211–20.

Césari, S. et al. (2013). The rice resistance protein pair RGA4/RGA5 recognizes the Magnaporthe oryzae effectors AVR-Pia and AVR1-CO39 by direct binding. Plant Cell **25**: 1463–81.

Césari, S., Bernoux, M., Moncuquet, P., Kroj, T., and Dodds, P.N. (2014a). A novel conserved mechanism for plant NLR protein pairs: the “integrated decoy” hypothesis. Plant-Microbe Interact. **5**: 606.

Césari, S., Kanzaki, H., Fujiwara, T., Bernoux, M., Chalvon, V., Kawano, Y., Shimamoto, K., Dodds, P., Terauchi, R., and Kroj, T. (2014b). The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease resistance. EMBO J. **33**: 1941–1959.

Chou, S., Krasileva, K. V, Holton, J.M., Steinbrenner, A.D., Alber, T., and Staskawicz, B.J. (2011). Hyaloperonospora arabidopsidis ATR1 effector is a repeat protein with distributed recognition surfaces. Proc. Natl. Acad. Sci. U. S. A. **108**: 13323–8.

Collier, S.M. and Moffett, P. (2009). NB-LRRs work a “bait and switch” on pathogens. Trends Plant Sci. **14**: 521–9.

Cui, H., Tsuda, K., and Parker, J.E. (2015). Effector-Triggered Immunity: From Pathogen Perception to Robust Defense. Annu. Rev. Plant Biol. **66**: 487–511.

Dean, R., Kan, J.A.N.A.L.V.A.N., Pretorius, Z.A., Hammond-kosack, K.I.M.E., Pietro, A.D.I., Spanu, P.D., Rudd, J.J., Dickman, M., Kahmann, R., Ellis, J., and Foster, G.D. (2012). The Top 10 fungal pathogens in molecular plant pathology. **13**: 414–430.

Eitas, T.K. and Dangl, J.L. (2010). NB-LRR proteins: pairs, pieces, perception, partners, and pathways. Curr. Opin. Plant Biol. **13**: 472–7.

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.

Fukuoka, S., Saka, N., Koga, H., Ono, K., Shimizu, T., Ebana, K., Hayashi, N., Takahashi, A., Hirochika, H., Okuno, K., and Yano, M. (2009). Loss of function of a proline-containing protein confers durable disease resistance in rice. Science **325**: 998–1001.

Gabriëls, S.H.E.J., Vossen, J.H., Ekengren, S.K., van Ooijen, G., Abd-El-Haliem, A.M., van den Berg, G.C.M., Rainey, D.Y., Martin, G.B., Takken, F.L.W., de Wit, P.J.G.M., and Joosten, M.H. a J. (2007). An NB-LRR protein required for HR signalling mediated by both extra- and intracellular resistance proteins. Plant J. **50**: 14–28.

Guillen, K., Ortiz-Vallejo, D., Gracy, J., Fournier, E., Kroj, T., and Padilla, A. (2015). Structure analysis uncovers a highly diverse but structurally conserved effector family in phytopathogenic fungi. PLoS Pathog. **accepted f**.

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.

Jacob, F., Vernaldi, S., and Maekawa, T. (2013). Evolution and Conservation of Plant NLR Functions. Front. Immunol. **4**: 297.

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 J. **19**: 4004–14.

Kanzaki, H. et al. (2012). Arms race co-evolution of Magnaporthe oryzae AVR-Pik and rice Pik genes driven by their physical interactions. Plant J.: 894–907.

Khang, C.H., Berruyer, R., Giraldo, M.C., Kankanala, P., Park, S.-Y., Czymmek, K., Kang, S., and

Valent, B. (2010). Translocation of Magnaporthe oryzae effectors into rice cells and their subsequent cell-to-cell movement. Plant Cell **22**: 1388–403.

Krasileva, K. V, Dahlbeck, D., and Staskawicz, B.J. (2010). Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. Plant Cell **22**: 2444–58.

Liu, W., Liu, J., Triplett, L., Leach, J.E., and Wang, G.-L. (2014). Novel Insights into Rice Innate Immunity Against Bacterial and Fungal Pathogens.

Maekawa, T. et al. (2011). Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. Cell Host Microbe **9**: 187–99.

Maqbool, a, Saitoh, H., Franceschetti, M., Stevenson, C., Uemura, a, Kanzaki, H., Kamoun, S., Terauchi, R., and Banfield, M. (2015). Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. Elife **4**: 1–24.

Nyarko, A., Singarapu, K.K., Figueroa, M., Manning, V. a., Pandelova, I., Wolpert, T.J., Ciuffetti, L.M., and Barbar, E. (2014). Solution NMR Structures of Pyrenophora tritici-repentis ToxB and Its Inactive Homolog Reveal Potential Determinants of Toxin Activity. J. Biol. Chem. **289**: 25946–25956.

Okuyama, Y. et al. (2011). A multifaceted genomics approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS-LRR protein genes. Plant J. **66**: 467–79.

Pennisi, E. (2010). Armed and Dangerous. Science (80-.). **327**: 804–805.

Qi, D. and Innes, R. (2013). Recent Advances in Plant NLR Structure, Function, Localization, and Signaling. Front. Immunol.: 221.

Ribot, C., Césari, S., Abidi, I., Chalvon, V., Bournaud, C., Vallet, J., Lebrun, M.-H., Morel, J.-B., and Kroj, T. (2013). The Magnaporthe oryzae effector AVR1-CO39 is translocated into rice cells independently of a fungal-derived machinery. Plant J. **74**: 1–12.

Le Roux, C. et al. (2015). A Receptor Pair with an Integrated Decoy Converts Pathogen Disabling

of Transcription Factors to Immunity. Cell **161**: 1074–1088.

Sarris, P.F.F. et al. (2015). A Plant Immune Receptor Detects Pathogen Effectors that Target WRKY Transcription Factors. Cell **161**: 1089–1100.

Schumann, F.H., Riepl, H., Maurer, T., Gronwald, W., Neidig, K.P., and Kalbitzer, H.R. (2007). Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. J. Biomol. NMR **39**: 275–289.

Skamnioti, P. and Gurr, S.J. (2009). Against the grain: safeguarding rice from rice blast disease. Trends Biotechnol. **27**: 141–50.

Sohn, K.H., Hughes, R.K., Piquerez, S.J., Jones, J.D.G., and Banfield, M.J. (2012). Distinct regions of the Pseudomonas syringae coiled-coil effector AvrRps4 are required for activation of immunity. Proc. Natl. Acad. Sci. U. S. A. **109**: 16371–6.

Steinbrenner, A.D., Goritschnig, S., and Staskawicz, B.J. (2015). Recognition and Activation Domains Contribute to Allele-Specific Responses of an Arabidopsis NLR Receptor to an Oomycete Effector Protein. PLOS Pathog. **11**: e1004665.

Takken, F.L.W. and Goverse, A. (2012). How to build a pathogen detector: structural basis of NB-LRR function. Curr. Opin. Plant Biol. **15**: 375–384.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: Development of a software pipeline. Proteins Struct. Funct. Bioinforma. **59**: 687–696.

Williams, S.J. et al. (2014). Structural basis for assembly and function of a heterodimeric plant immune receptor. Science **344**: 299–303.

Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L., and Sykes, B.D. (1995). 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J. Biomol. NMR **6**: 135–140.

Wu, C.-H., Belhaj, K., Bozkurt, T.O., Birk, M.S., and Kamoun, S. (2015). Helper NLR proteins NRC2a/b and NRC3 but not NRC1 are required for Pto-mediated cell death and resistance

in Nicotiana benthamiana. New Phytol.: doi: http://dx.doi.org/10.1101/015479.

Zhang, Z. M., Zhang, X., Zhou, Z.-R., Hu, H.-Y., Liu, M., Zhou, B., and Zhou, J. (2013). Solution structure of the Magnaporthe oryzae avirulence protein AvrPiz-t. J. Biomol. NMR **55**: 219–23.

FIGURE LEGENDS

Figure 1. The AVR-Pia-H3 NMR structure is similar to the structure of wild-type AVR-Pia

(A) Chemical shift differences (ΔδNH) from the comparison of ^{15}N-HSQC of AVR-Pia wild-type and mutants F24S, T46N or F24S T46N (AVR-Pia-H3). The ß-strand assignments from the AVR-Pia^{wt} structure are indicated on the top and polymorphic residues by (*).(B) Structure overlay of AVR-Pia (blue) and AVR-Pia-H3 (orange). (C) Topology of the AVR-Pia-H3 structure.

Figure 2. AVR-Pia binds RGA5$_{RATX1}$ with intermediate affinity and a well-defined interaction surface

(A) ITC curves for the titration of the RGA5$_{RATX1}$ domain by AVR-Pia^{wt} (□) and AVR-Pia-H3 (□) at 25°C. For AVR-Pia^{wt} the fitting parameters were: N = 0.994 ± 0.004, K_a = 1.28 ± 0.04 10^{-5} mol^{-1}, ΔH = -8179 ± 47.95 cal.mol-1, ΔS = -4.06 cal.K-1.mol-1. The red line shows a simulated curve for a 10x lower affinity (K_a = 1.28 10^{-4} mol^{-1}).

(B) NMR titration and surface mapping. Plot of the chemical shift differences (Δppm) between unbound and bound AVR-Pia (blue) or AVR-Pia-H3 (red). Chemical shift differences were calculated as the Hamming distance (Schumann et al., 2007), Δδ (ppm) = $|\Delta\delta(^1H)_{ij}|$ + 0.102 * $|\Delta\delta(^{15}N)_{ij}|$, where $\Delta\delta(^1H)_{ij}$ and $\Delta\delta(^{15}N)_{ij}$ are the chemical shift differences observed at R=0 and R=2, respectively.

Structures of AVR-Pia (C and D) and AVR-Pia-H3 (E) with color-coded surfaces showing the differences in chemical shifts in the NMR titration (difference between free (R=0) and RGA5$_{RATX1}$-bound AVR-Pia or AVR-PiaH3 (R=2)). Surfaces of residues with chemical shift differences Δδ(ppm) ≥ 0.2 are shown in dark blue (residues in white letters), and in light blue for 0.2 > Δδ(ppm) ≥ 0.1 ppm (residues in black letters). Surfaces of residues not observed in the AVR-Pia-RGA5$_{RATX1}$ complex (R=2) HSQC are reported in grey (residues in red letters), and not perturbed residues are not highlighted (residues are not indicated). The view in D is the opposite face of C, which has been rotated 180° from the vertical axis.

Figure 3. Mutations in the binding surface of AVR-Pia affect binding to $RGA5_{C\text{-}ter}$ in yeast two hybrid assays.

(A) The interaction between AVR-Pia mutants (BD fusion) and $RGA5_{C\text{-}ter}$ (AD fusion) was assayed by a yeast two-hybrid experiment. Three dilutions (1/10, 1/100, 1/1000) of yeast cultures adjusted to an OD of 0.2 were spotted on synthetic double drop out (DDO) medium (-Trp/-Leu) to control proper growth and on synthetic TDO (-Trp/-Leu/-His) either without or supplemented with 3-amino-1,2,4-triazole (3AT) to test for interaction. Yeast transformations and interaction analyses were performed twice with identical results. Photos were taken after 4 days of growth.

(B) Equal production of AVR-Pia mutant proteins was determined by immunoblotting with anti-AVR-Pia antibodies.

Figure 4. AVR-Pia mutants with reduced $RGA5_{C\text{-}ter}$ binding in yeast are also impaired in binding to $RGA5_{C\text{-}ter}$ and $RGA5_{RATX1}$ *in planta*.

HA:$RGA5_{C\text{-}ter}$ (A) or HA:$RGA5_{RATX1}$ (B) were transiently expressed with YFP:AVR-Pia^{WT}, YFP:AVR-Pia mutants or YFP:PWL2 in *N. benthamiana*. Protein extracts were analysed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analysed by immunoblotting with α-GFP for the detection of immunoprecipitated AVR-Pia variants. Co-precipitated HA:$RGA5_{C\text{-}ter}$ (A) or HA:$RGA5_{RATX1}$ (B) proteins were detected using α-HA antibody.

Figure 5. Effector recognition by RGA5 requires binding to the RATX1 domain

Transgenic *M. oryzae* isolates were analysed for the production of the AVR-Pia protein by immunoblotting using culture filtrate and α-AVR-Pia antibodies (A lower panel) and were sprayed on 3-week-old plants of the rice cultivar Kitaake possessing *Pia* resistance. 7 days after inoculation, leaves were scanned (A lower panel) and 3 different types of lesions (1= fully resistant, 2=partially resistant/weakly susceptible 3=fully susceptible) were counted on leaves from 10 different plants per isolate to determine mean symptom scores and significantly different classes of isolates using Kruskal-Wallis analysis of variance combined with a multi- comparison Dunn test for non-parametric data (B). The AVR-Pia variants grouped with respect to their avirulence activity in three

significantly different classes: a=inactive; b=partially active; c=active. Similar results were obtained in two independent experiments and with additional transgenic isolates.

Figure 6. AVR-Pia interacts with RGA5 outside the RATX1 domain

HA:RGA5 (A) and HA:RGA5$_{\Delta RATX1}$ (B) were expressed with YFP:AVR-Pia^{WT}, YFP:AVR-Pia mutants and YFP:PWL2 in *N. benthamiana*. Protein extracts were analysed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analysed by immunoblotting with α-GFP for the detection of immunoprecipitated AVR-Pia variants. Co-precipitated RGA5 (A) or HA:RGA5$_{\Delta RATX1}$ (B) were detected using α-HA antibody.

Figure 7. Model of AVR-Pia recognition by the RGA4/RGA5 receptor complex

AVR-Pia binds to the RATX1 domain of RGA5 with a defined interaction surface and interacts, in addition, through independent surfaces with other sites in RGA5. These additional interactions are not sufficient to relieve the repression RGA5 exerts on RGA4. Indeed, AVR-Pia mutants that associate with RGA5, but do not bind RGA5$_{RATX1}$ as well as RGA5 mutants that lack the RATX1 domain, do not permit activation of resistance. We propose that simultaneous interactions of AVR-Pia with different parts of RGA5, including the RATX1 domain, stabilize conformational changes that activate the RGA4/RGA5 complex.

Supplemental Figure Legends

Supplemental Figure 1. Comparison of NMR relaxation of AVR-Pia and AVR-Pia-H3.

Generalized order parameter (S2) obtained from 15N relaxation data at 500 MHz for AVR-Pia (circles) and AVR-Pia-H3 (triangles) determined from Lipari-Szabo formalism with the program DYNAMOF (Barthe et al., 2006). The different colors indicate the use of the “simple” Lipari-Szabo formalism by black triangles for AVR-Pia-H3, and by white circles for AVR-Pia, with possible additional exchange contributions shown by the symbols in grey color.

Supplemental Figure 2. HSQC spectra of AVR-Pia and AVR-Pia-H3 recorded upon titration with $RGA5_{RATX1}$.

HSQC spectra are shown for 50 µM AVR-Pia (A) and AVR-Pia-H3 (B) without $RGA5_{RATX1}$ (black) and various concentration of $RGA5_{RATX1}$ at 12.5 µM (R=0.25 in grey), 25 µM (R=0.5 in green), 50 µM (R=1 in orange) and 100µM (R=2 in pink). Residues belonging to the N-terminal tag are indicated by negative numbering and (*). For AVRPia (A) some of the largest chemical shift differences between the free (R=0) and bound (R=2) form are shown in the zoom panels at the upper left corner. With AVR-Pia-H3 (B), most of the HSQC peaks remain unperturbed and the enlarged plot shows the peaks present at the beginning of the titration for Y27, Y41, T51, N46 (in black or green for R=0 or R=0.5) that are missing at the end of the titration (in pink, R=2).

Supplemental Figure 3. AVR-Pia mutants affected in $RGA5_{RATX1}$-binding are well-structured.

(A) AVR-Pia NMR structure (Guillen et al., 2015) showing the amino acids that were replaced in the AVR-Pia variants carrying mutations in the candidate interaction surface (left) and, for control purposes, in others surfaces (right). (B) 1H 1D-NMR spectra for AVR-Pia point mutants E58A, R43G, R23A, R36A and D29A. Chemical shift assignments of wild type AVR-Pia are shown by vertical bars above the spectra.

Supplemental Figure 4. AVR-Pia mutants not-affected in $RGA5_{RATX1}$-binding trigger HR in *N. benthamiana*

HA:RGA5, RGA4:HA and untagged AVR-Pia variants were transiently expressed in *N. benthamiana* leaves by *A.tumefaciens* infiltration and cell death responses (A and B) and AVR-Pia protein levels were determined (C). (A) The HR response was evaluated after 4 days by visual inspection (left) and a lack of chlorophyll fluorescence (right). (B) The activity of AVR-Pia mutants was determined in comparison to the positive control AVR-Piawt and the negative control AVR-PiaF24S. (C) Proteins were extracted 48 h after infiltration, immunoprecipitation of AVR-Pia variants was performed with α-AVR-Pia antibodies and anti- protein A/G agarose beads (a-protein A/G) due to low abundance of certain AVR-Pia variants and the immunoprecipitate was analyzed by immunoblotting with α-AVR-Pia antibodies.

Supplemental Figure 5. Characterization of transgenic *M. oryzae* isolates carrying *AVR-Pia* mutant constructs.

(A) Culture filtrates of 3 different transgenic *M. oryzae* isolates per construct grown for 5 days in a liquid complete medium was analyzed by immunoblotting with a-AVR-Pia antibodies for expression and secretion of AVR-Pia mutant proteins. (B) For each *AVR-Pia* mutant construct, as well as the controls *AVR-Piawt* and *mRFP*, three independent transgenic *M. oryzae* isolates were sprayed on resistant Kitaake (+*Pia*) and susceptible Maratelli (-*Pia*) rice plants. Symptoms were recorded 7 days after inoculation.

Supplemental Figure 6. RGA5$_{\Delta RATX1}$ represses RGA4-mediated cell death but does not recognize AVR-Pia.

AVR-Pia and RGA4:HA were co-expressed with HA:RGA5$_{\Delta RATX1}$ (A) or HA:RGA5$_{\Delta RATX1}$ and HA:RGA5$_{RATX1}$ (B) to evaluate the importance of RGA5$_{RATX1}$-binding for AVR-Pia- recognition. As controls, the de-repression of RGA4 by AVR-Pia in the presence of HA:RGA5 (C) and the repression of RGA4 -mediated cell death by HA:RGA5$_{\Delta RATX1}$ (D) or HA:RGA5$_{\Delta RATX1}$ and HA:RGA5$_{RATX1}$ (E) are shown. On all leaves, cell death induced by RGA4:HA and repression of cell death upon co-expression of RGA4:HA and HA:RGA5 were recorded as positive and negative controls. Cell death was evaluated after 4 days by visual inspection (left) and lack of chlorophyll fluorescence (right).

Supplemental Figure 7. Comparison of the AVR-Pia and AVR-PikD structures and their complexes with RAX1/HMA domains.

(A) AVR-PikD structure with the Pikp-1HMA interaction surface in pink and labeling of important residues from the N-terminal loop, ß2 and ß3 (based on PDB structure 5A6W, Maqbool *et al*., 2015). (B) Superposition of AVR-Pik and AVR-Pia structures using DALI software. The rmsd over 50 residues is 2.4 Å. (C) AVR-Pia / RATX1 docking simulations using the Rosetta suite. One copy of RATX1 is displayed as a grey cartoon in the middle, and was used as a reference frame for the superposition. The top green model corresponds to the position of AVR-PikD in the AVR-PikD/ Pikp-1HMA complex (5A6W), while the other colored models are AVR-Pia from the docking by RosettaDock. Each colored AVR-Pia model is representative of a cluster of models sharing the same orientation relative to RATX1. Ordered with respect to their relative size, from the biggest to the smallest cluster, they are red, orange, yellow, cyan and blue.

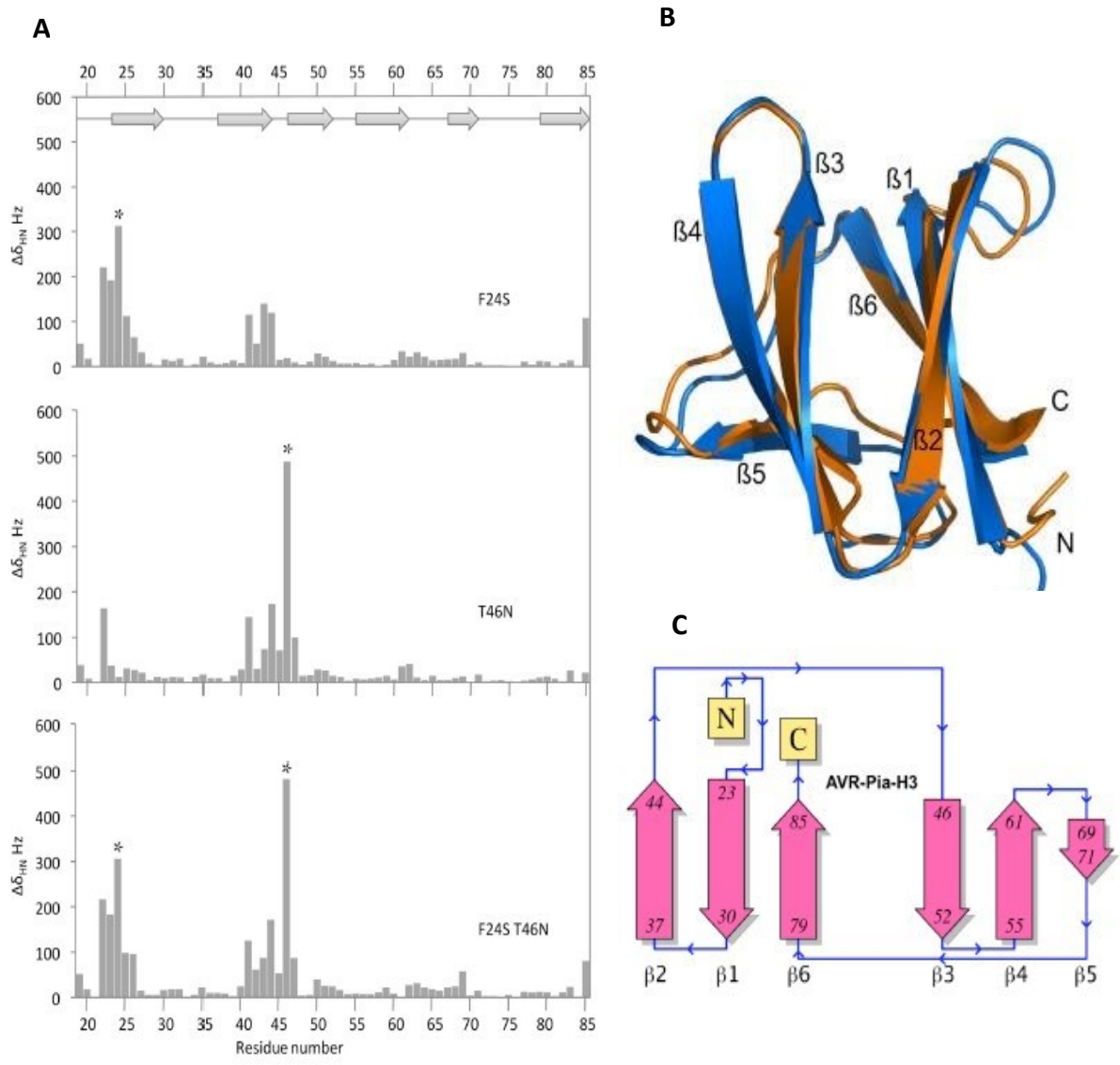
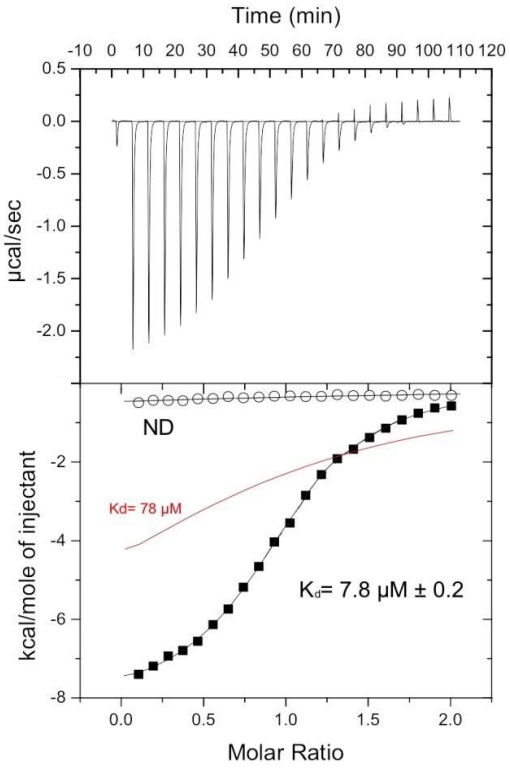


Figure 1. The AVR-Pia-H3 NMR structure is similar to the structure of wild-type AVR-Pi

A



B

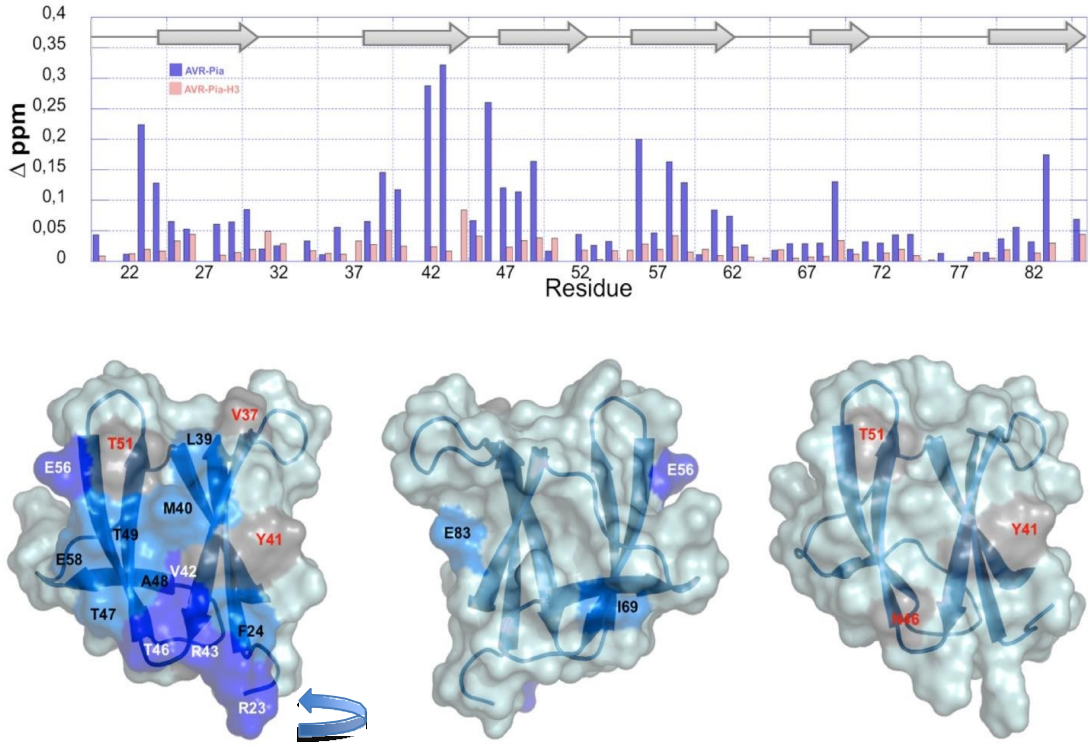


Figure 2. AVR-Pia binds RGA5$_{RATX1}$ with intermediate affinity and a well-defined interaction surface

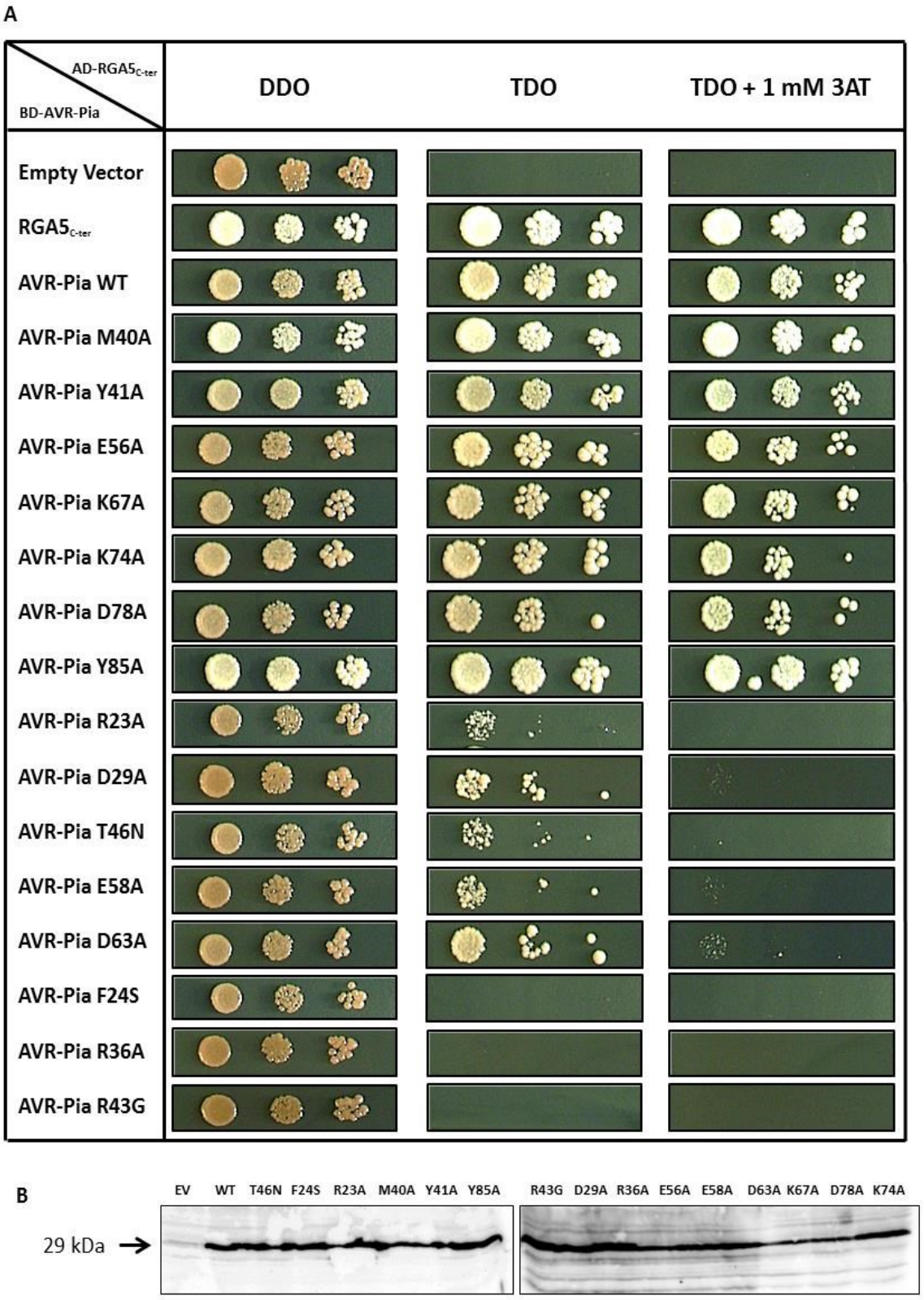


Figure 3. Mutations in the binding surface of AVR-Pia affect binding to $RGA5_{C\text{-}ter}$ in yeast two hybrid assays

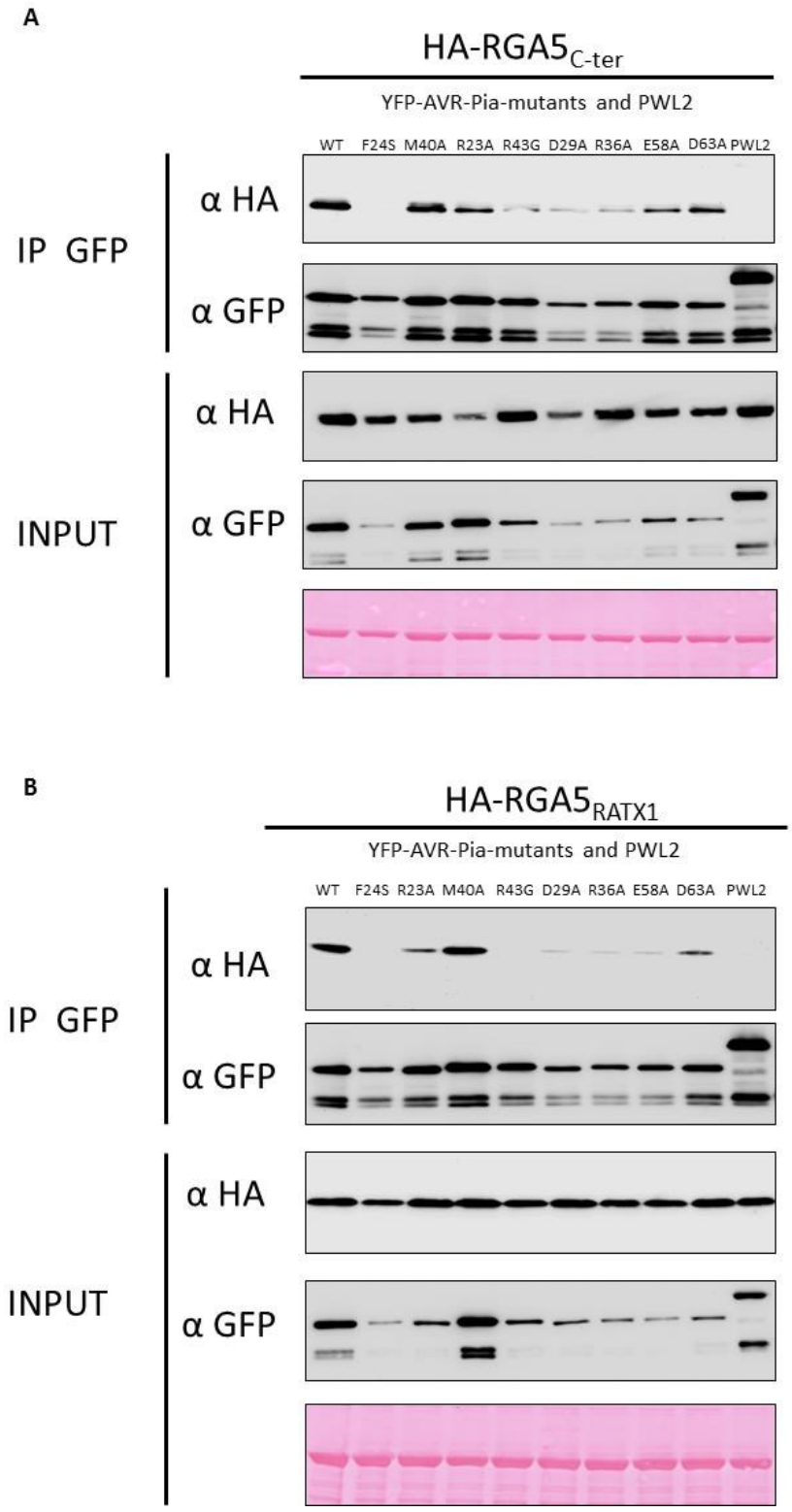
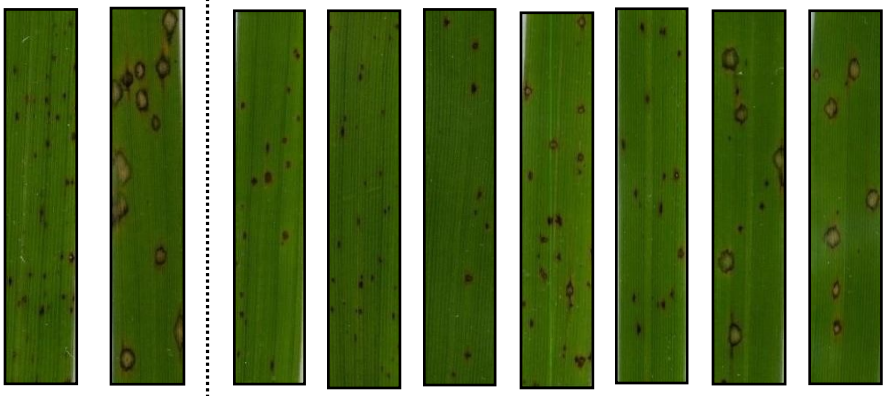


Figure 4. AVR-Pia mutants with reduced RGA5$_{C\text{-}ter}$ binding in yeast are also impaired in binding to RGA5$_{C\text{-}ter}$ and RGA5$_{RATX1}$ *in planta*.

A

	M. oryzae isolates								
	Controls		AVR-Pia mutants						
	AVR-Pia	mRFP	R23A	D29A	R36A	R43G	E58A	F24S	D63A
Kitaake (Pia)									
Avirulence activity	Active	Control	Active	Active	Active	Reduced activity	Active	Inactive	----
α-AVR-Pia (7,5 KDa) Immunoblot									

B

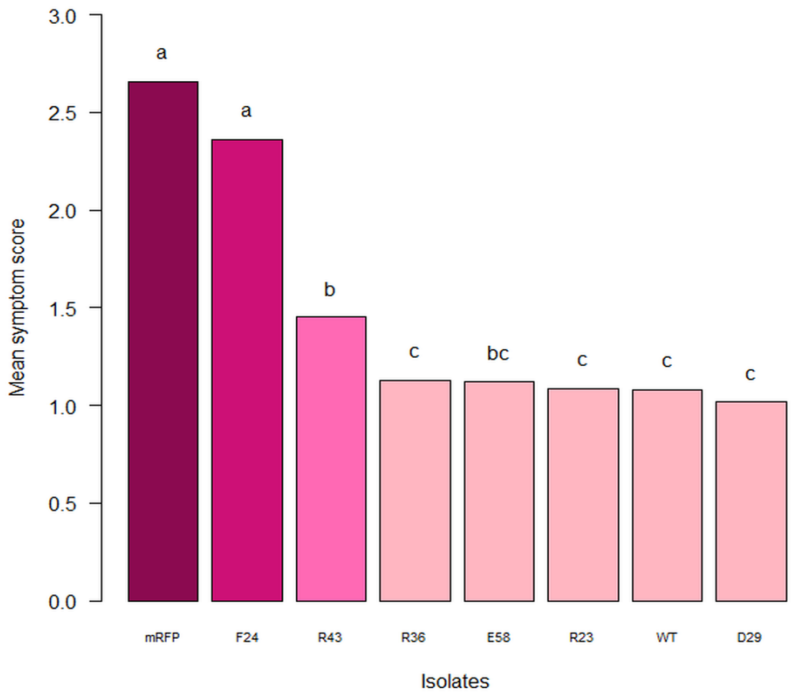


Figure 5. Effector recognition by RGA5 requires binding to the RATX1 domain

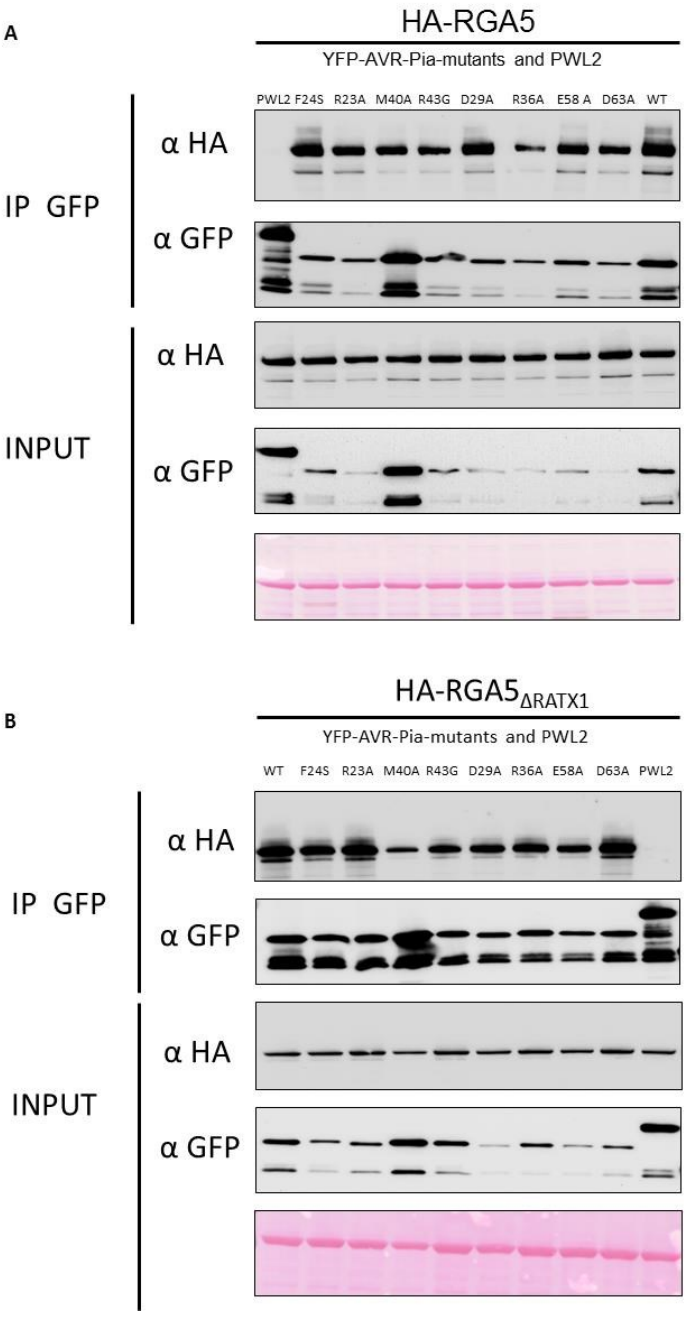


Figure 6. AVR-Pia interacts with RGA5 outside the RATX1 domain

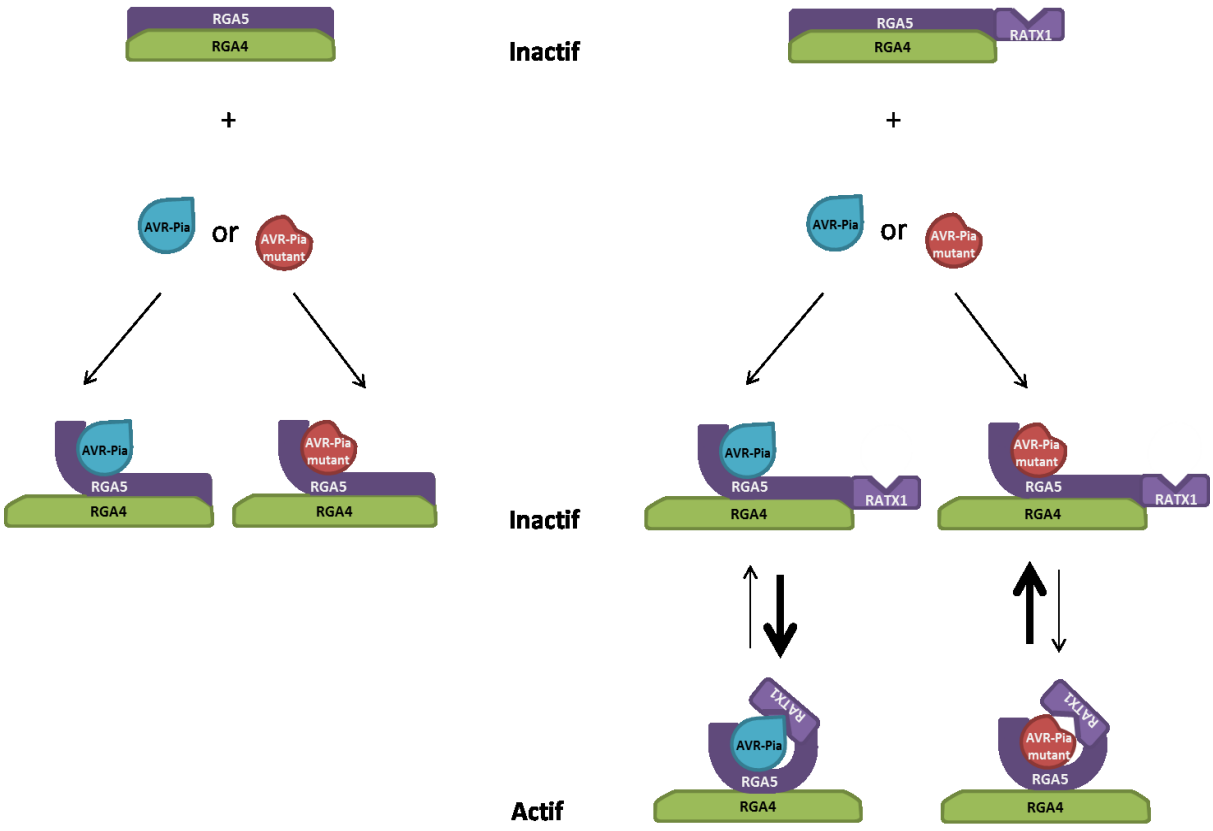
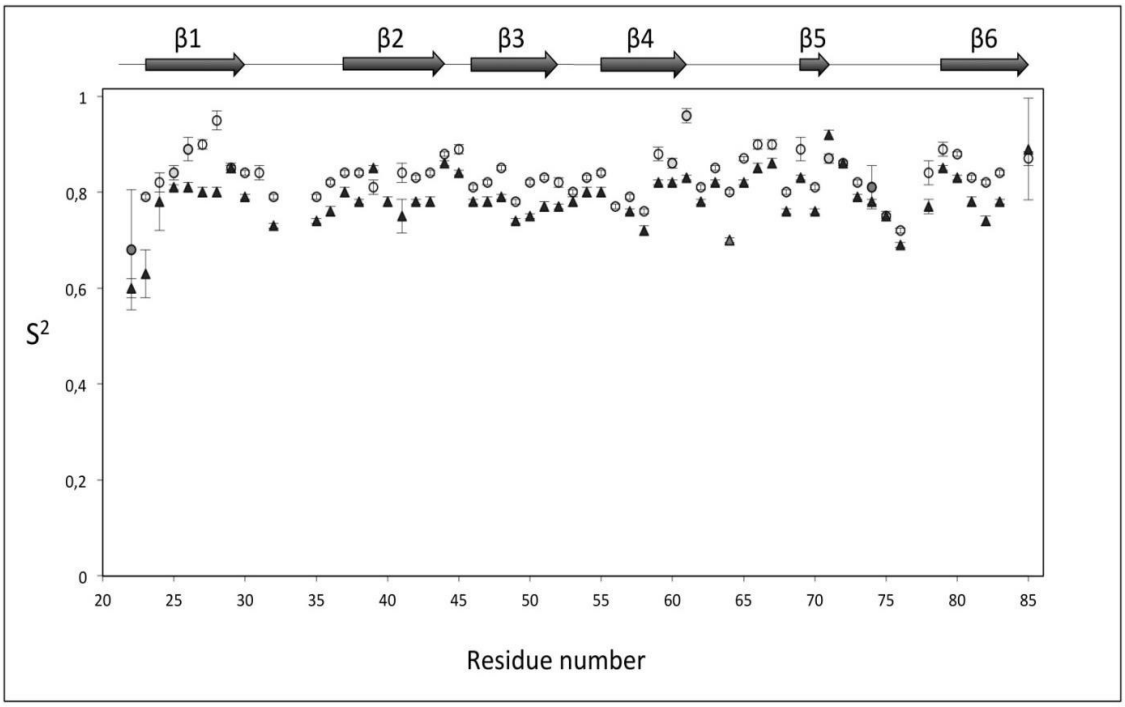
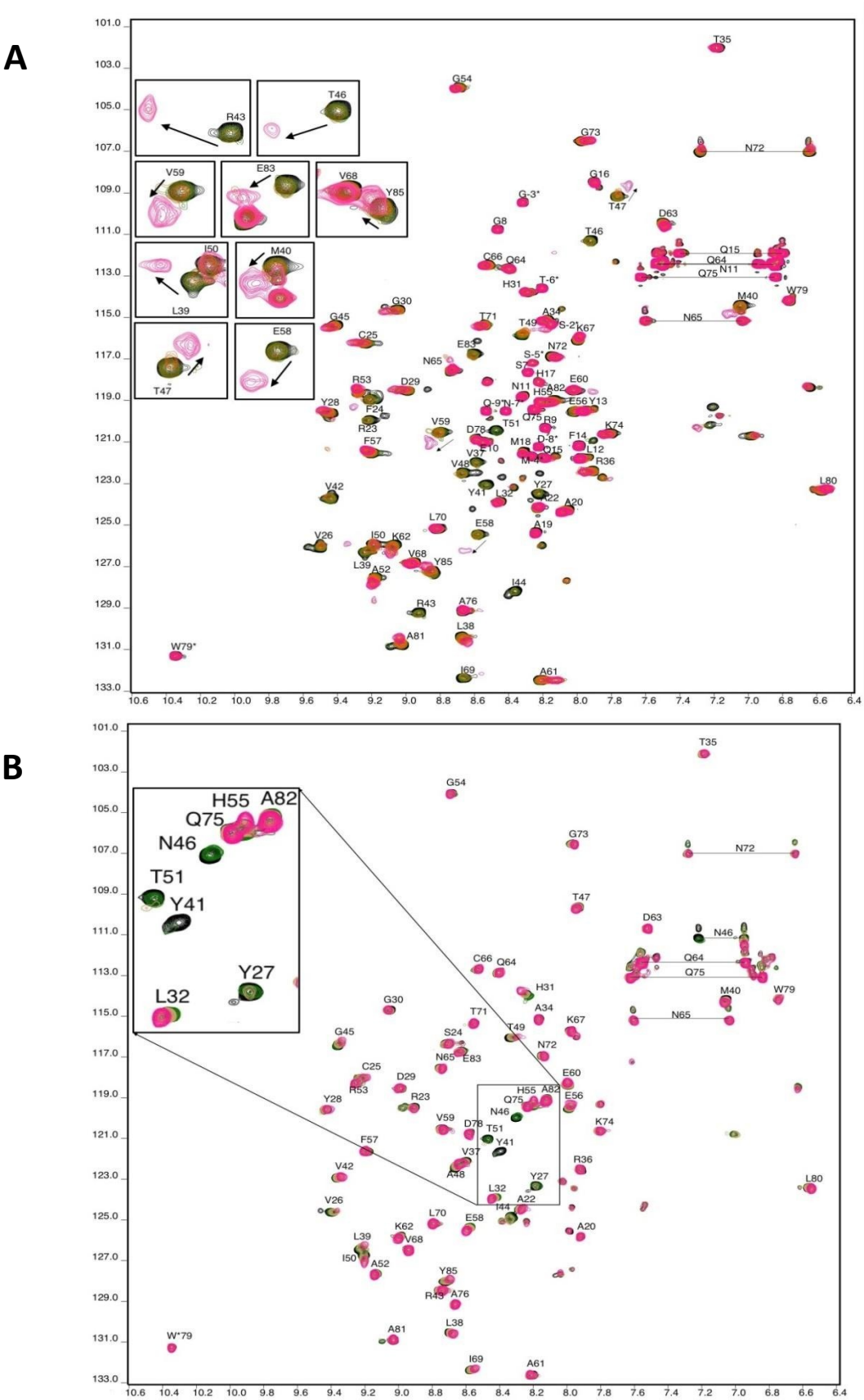


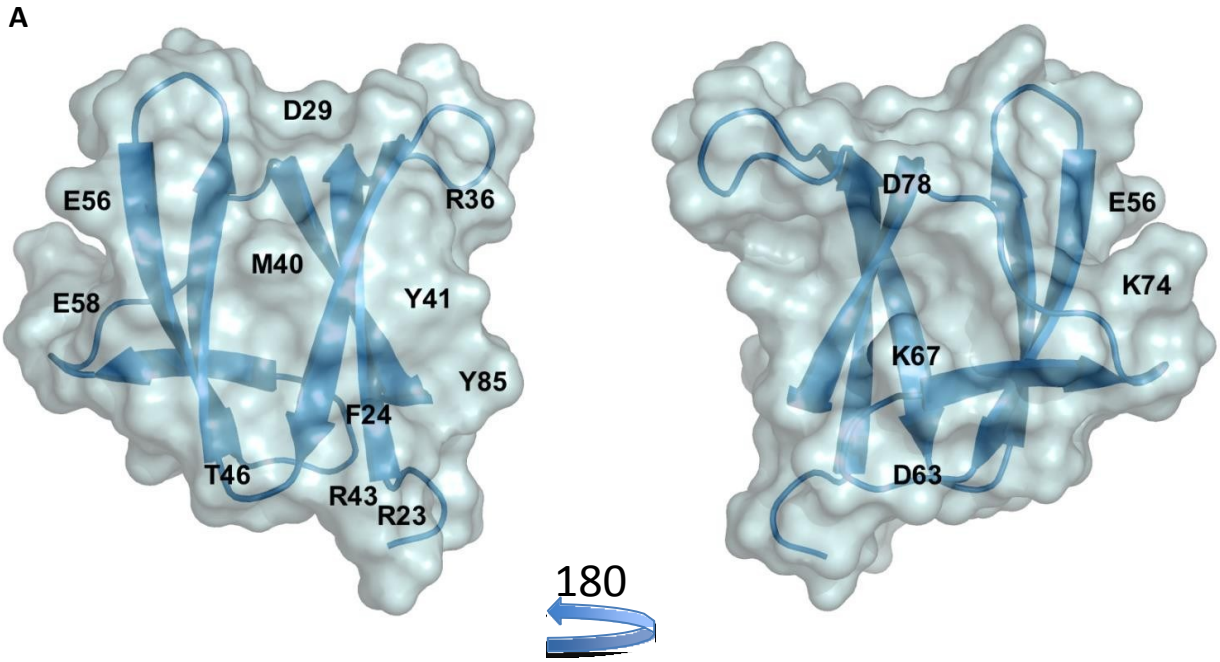
Figure 7. Model of AVR-Pia recognition by the RGA4/RGA5 receptor complex

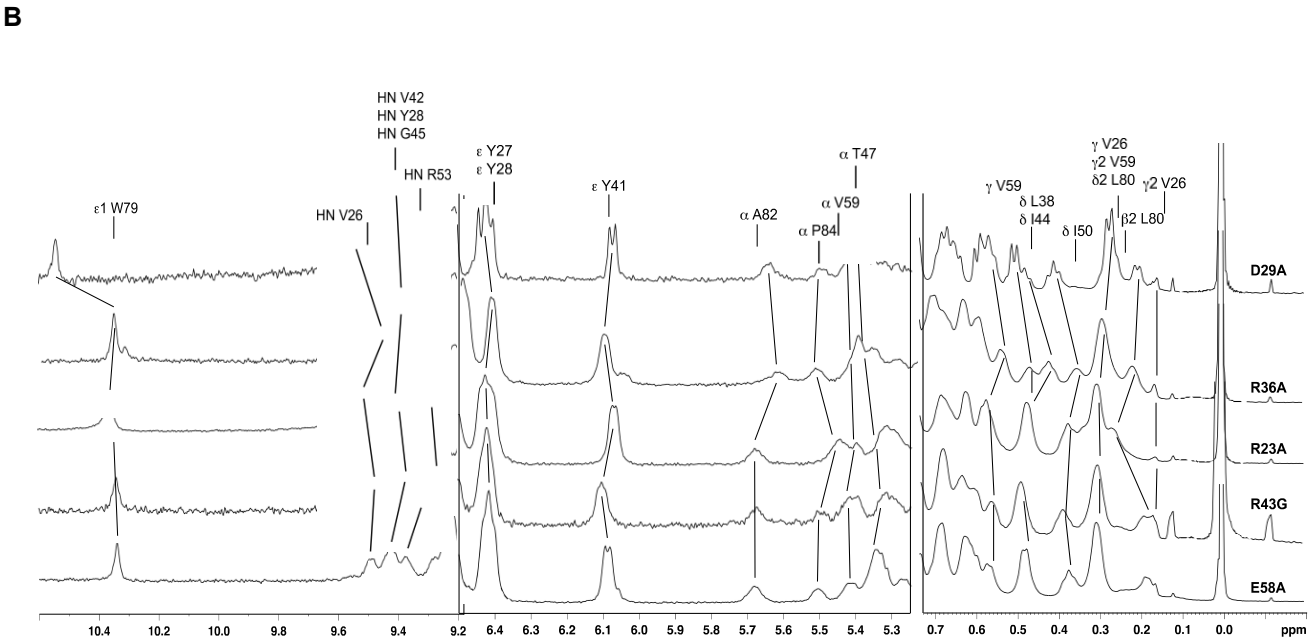


Supplemental Figure 1. Comparison of NMR relaxation of AVR-Pia and AVR-Pia-H3.

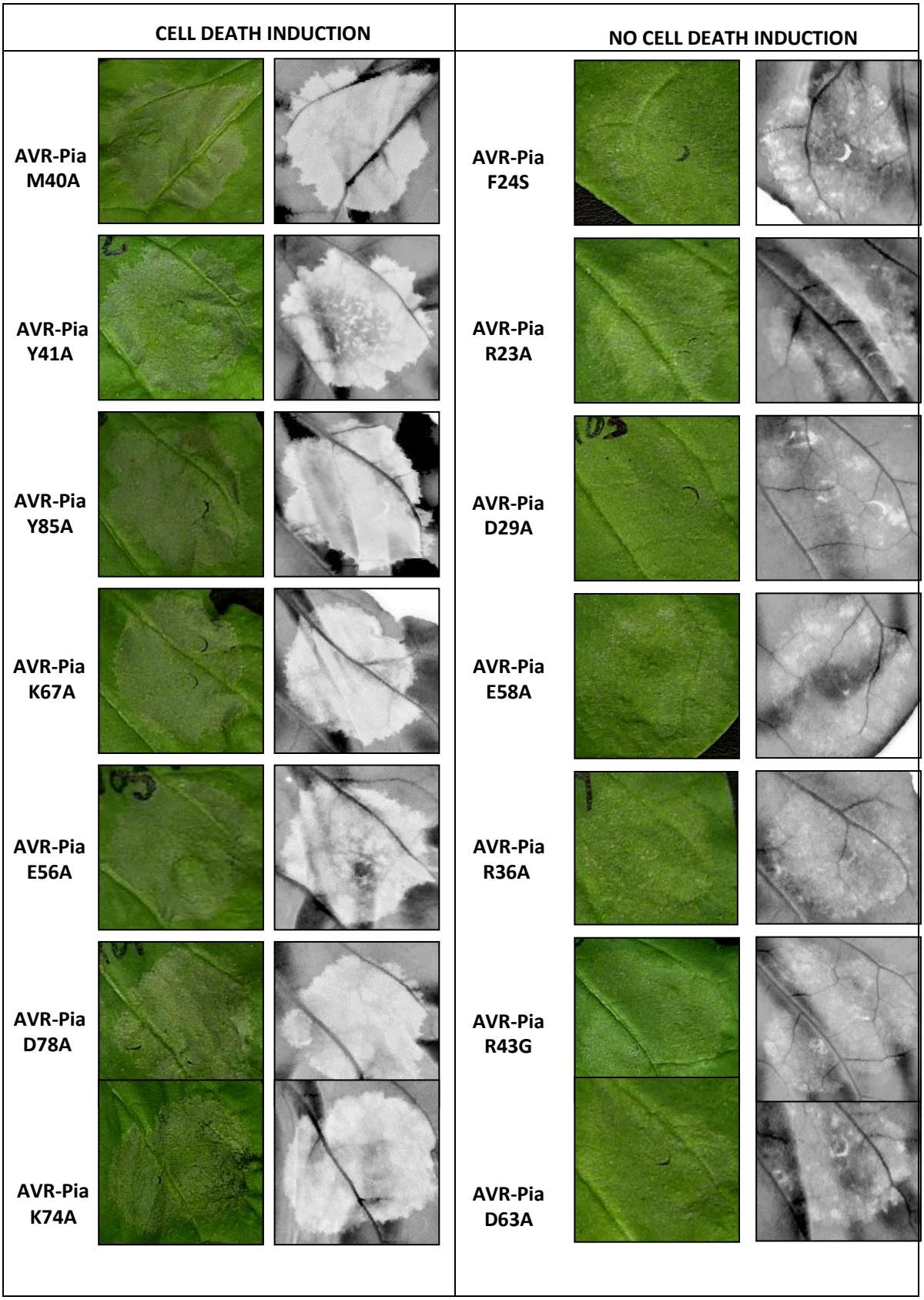


Supplemental Figure 2. HSQC spectra of AVR-Pia and AVR-Pia-H3 recorded upon titration with the RGA5$_{RATX1}$ domain

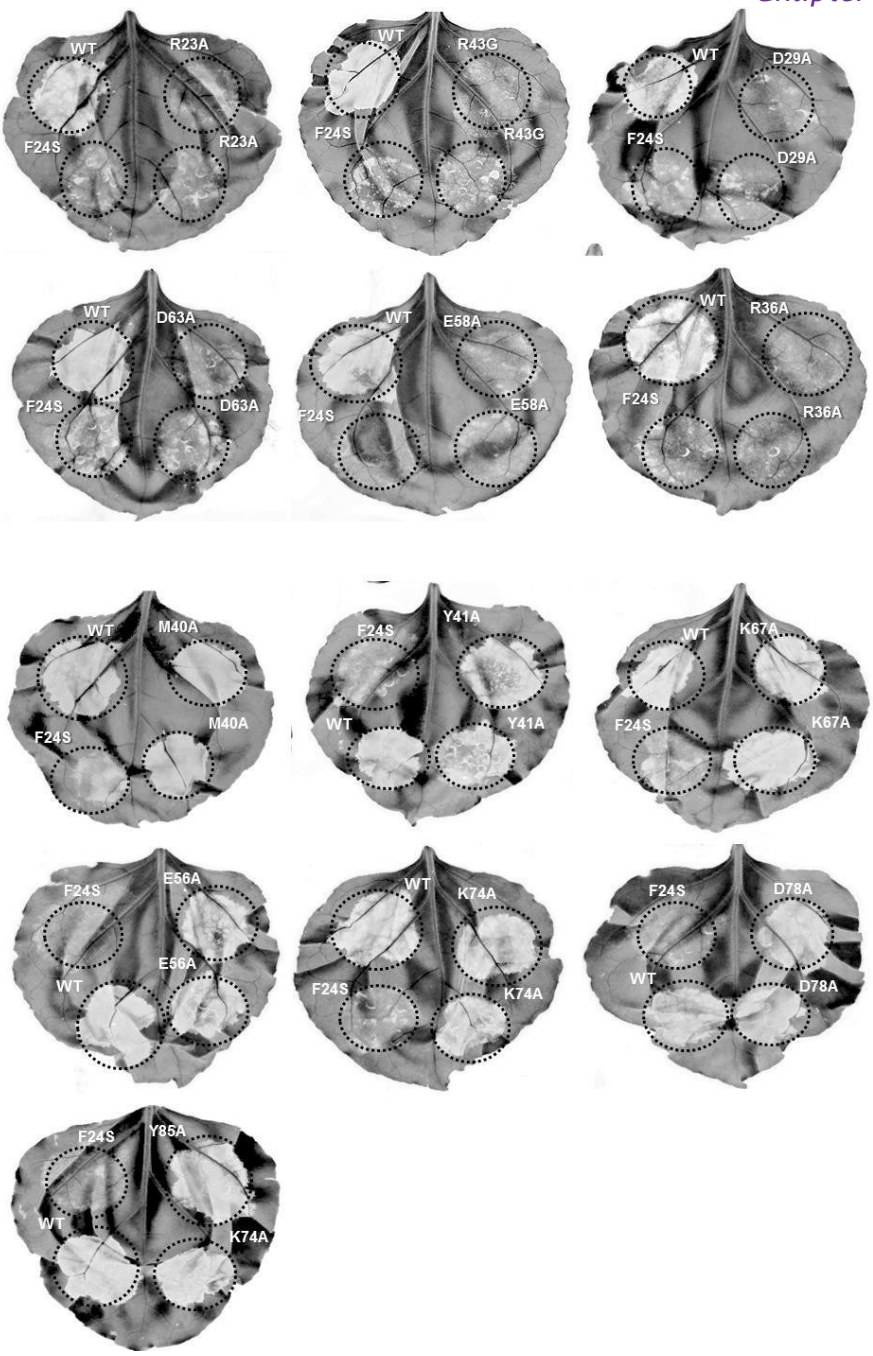




Supplemental Figure 3. AVR-Pia mutants affected in RGA5$_{RATX1}$-binding are well-structured

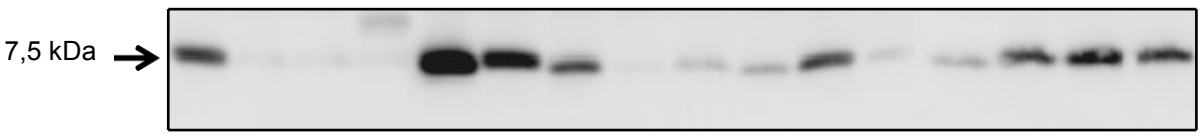


Supplemental Figure 4. AVR-Pia mutants not-affected in RGA5$_{RATX1}$-binding trigger HR in *N. benthamiana*

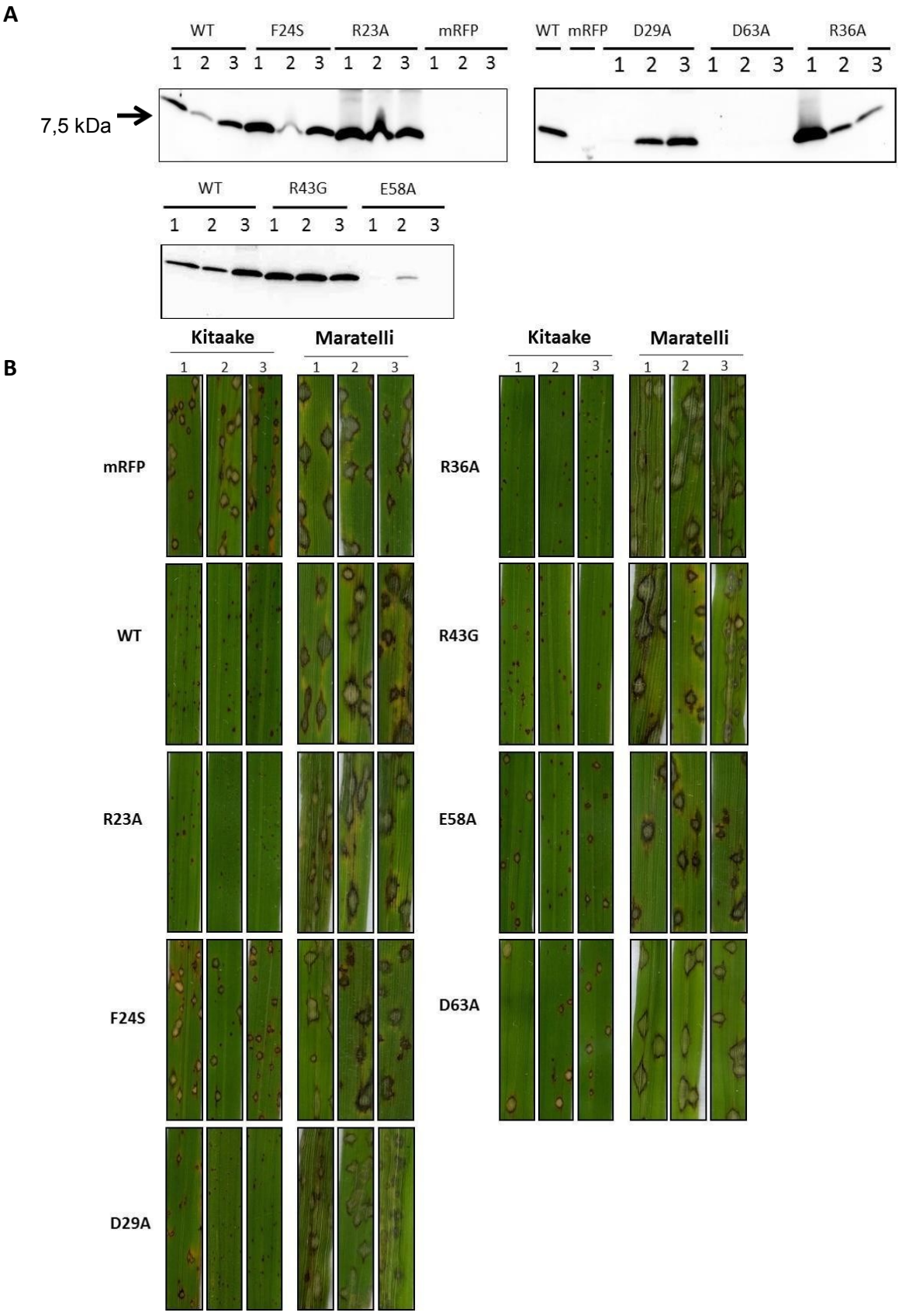


C

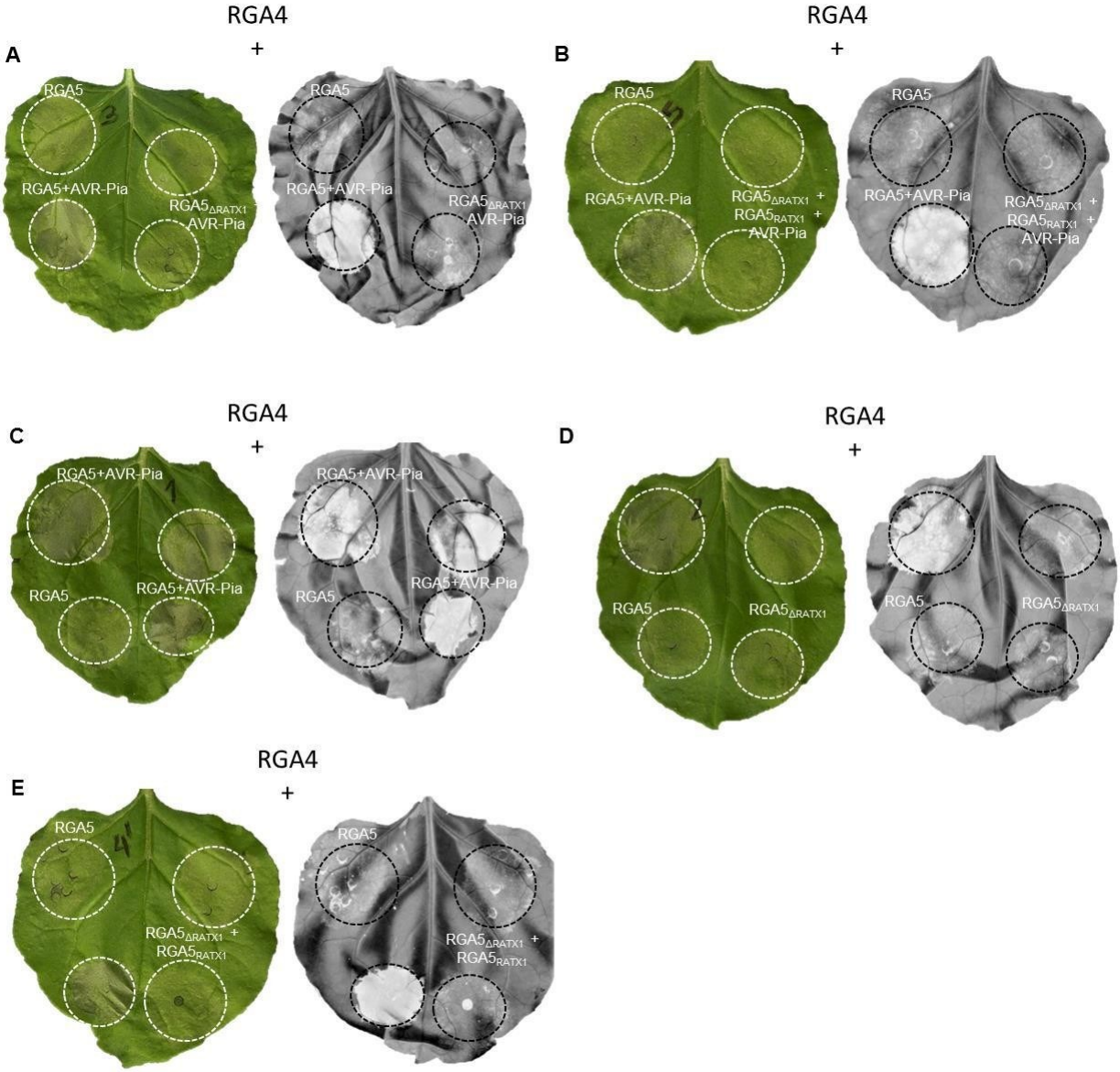
WT F24S F24S T46N R23A M40A Y41A Y85A R43A D29A R36A E56A E58A D63A K67A D78A K74A



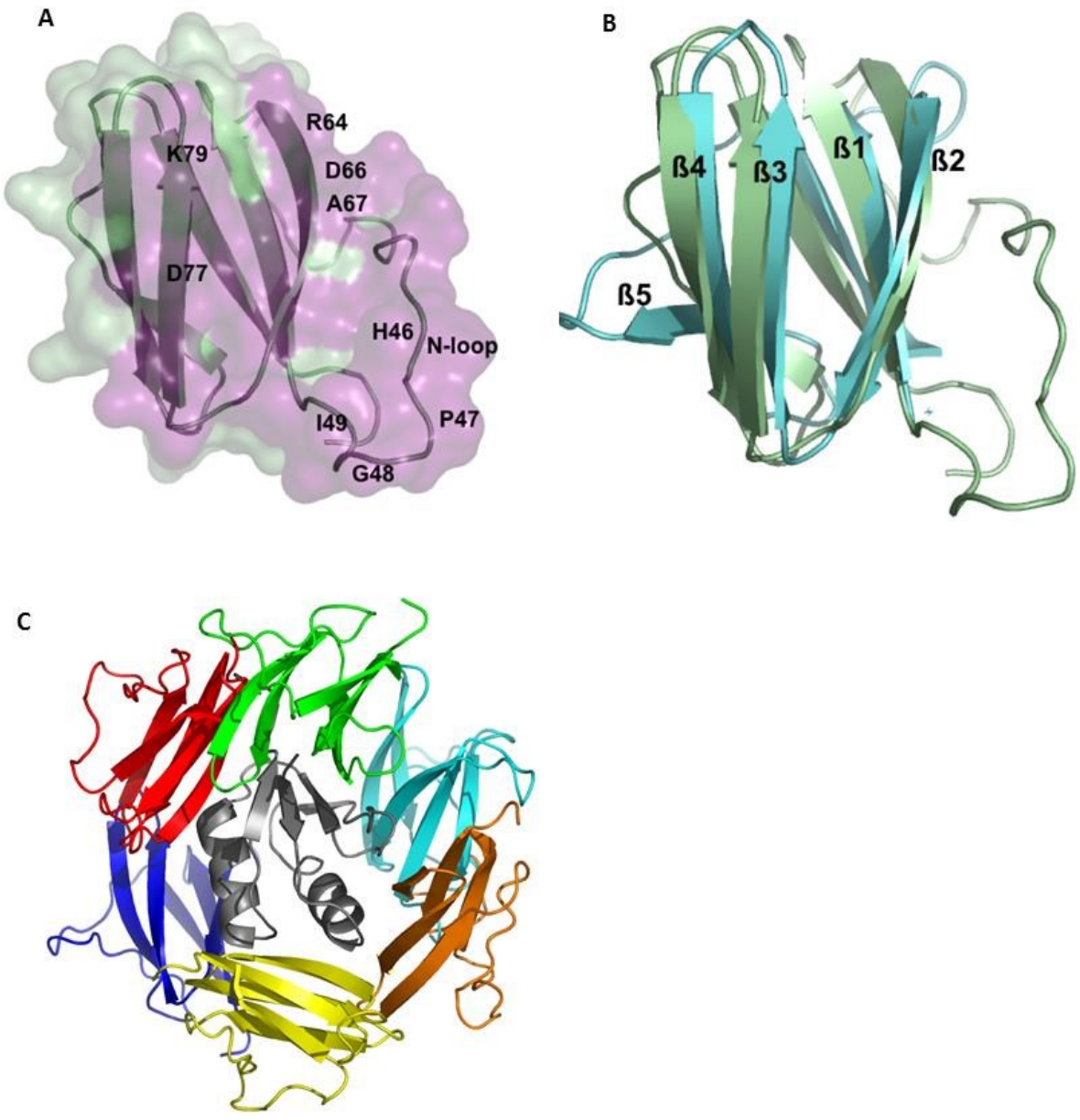
Supplemental Figure 4. AVR-Pia mutants not-affected in RGA5$_{RATX1}$-binding trigger HR in *N. benthamiana*



Supplemental Figure 5. Characterization of transgenic *M. oryzae* isolates carrying *AVR-Pia* mutant constructs.



Supplemental Figure 6. RGA5$_{\Delta RATX1}$ represses RGA4-mediated cell death but does not recognize AVR-Pia.



Supplemental Figure 7. Comparison of the AVR-Pia and AVR-PikD structures and their complexes with RAX1/HMA domains.

Supplemental Table 1. *NMR experiments acquired for structure calculations and chemical shift assignments of AVR-Pia-H3*

		Size			Sweep width (ppm)						
Experiments	nuclei	F3	F2	F1	F3	F2	F1	Mix(ms)	NS	D1(s)	B_o(MHz)
^{15}N-HSQC	^{1}H, ^{15}N	1500	256	-	-	-	-	-	8	1	700
^{15}N-NOESY-HSQC	^{1}H,^{15}N, ^{1}H	1500	64	280	14	40	14	150	8	1	700
NOESY (H_2O)	^{1}H, ^{1}H	1500	512	-	15.95	15.95	-	150	96	1	700
TOCSY (H_2O)	^{1}H, ^{1}H	1500	512	-	15.95	15.95	-	41.2	64	1	700
^{13}C-HSQC(D_2O)	^{1}H, ^{13}C	2048	400		14	160	-	-	240	1	700
^{13}C-TOCSY(D_2O)	^{1}H, ^{13}C	2048	400	-	12.03	100	-	40.25	320	1	700
^{1}H,^{15}N NOE	^{1}H, ^{15}N	1024	128	-	14	35	-	Sat. 3s	64	6	500
R_1	^{1}H, ^{15}N	1024	128	-	14	35	-		16	2.5	500
R_2	^{1}H, ^{15}N	1024	128	-	14	35	-		16	2.5	500

Supplemental Table 2. *Statistics for 20 NMR structures of AVR-Pia-H3*

	AVR-Pia-H3
NOE restraints	1366
Short range (|i-j|≤1)	891
Medium range (1<|i-j|<5)	133
Long range (|i-j|≥5)	342
H-bond restraints	20
Dihedral restraints (a)	100
Average Number of NOE Violations per structure	
> 0.1 Å	23.75
> 0.2 Å	1.0
> 0.3 Å	0.05
> 0.4 Å	0
Dihedral violations	
> 2°	0
Ramachandran plot statistics	
most favourable regions (%)	84.3
additionally allowed regions (%)	15.7
generously allowed regions (%)	0.0
disallowed regions (%)	0.0
Pairwise RMSD (Å) (b)	
Backbone	0.77 ± 0.19
Heavy atoms	1.43 ± 0.16

Structures were calculated using CYANA, refined using CNS, and analysed using PROCHECK.

(a) Residues in regular secondary structures were derived from the chemical shifts using TALOS+ software.

(b) Main chain atoms (N, Cα, C) over the residues 20-85.

Supplemental Table 3. Primer

oDO01	gcttaatggcgccagctgctttttgcgtctattacgac
oDO02	gtcgtaatagacgcaaaaagcagctggcgccattaagc
oDO03	gacacgtgtcctgcttgcttacgttagaatcggca
oDO04	tgccgattctaacgtaagcaagcaggacacgtgtc
oDO05	acgtgtcctgcttatggctgttagaatcggcacta
oDO06	tagtgccgattctaacagccataagcaggacacgt
oDO07	ggcttgctgccgagcctgcttaggacccagctttctt
oDO08	aagaaagctgggtcctaagcaggctcggcagcaagcc
oDO17	attttgcgtctattacgctggccaccttcccgcga
oDO18	tcgcgggaaggtggccagcgtaatagacgcaaaat
oDO19	ccaccttcccgcgacagctgtcctgcttatgtacgt
oDO20	acgtacataagcaggacagctgtcgcgggaaggtgg
oDO21	tacggcccgtgggcacgctttcgaagttgaagcaa
oDO22	ttgcttcaacttcgaaagcgtgcccacgggccgta
oDO23	ccgtgggcacgaattcgctgttgaagcaaaagacca
oDO24	tggtcttttgcttcaacagcgaattcgtgcccacgg
oDO25	tcgaagttgaagcaaaagctcagaattgcaaagttat
oDO26	ataactttgcaattctgagcttttgcttcaacttcga
oDO27	caaaagaccagaattgcgctgttattctcaccaatgg
oDO28	ccattggtgagaataacagcgcaattctggtcttttg
oDO29	atggcaaacaagcaccggcttggcttgctgccgagcct
oDO30	aggctcggcagcaagccaagccggtgcttgtttgccat
oDO31	ttattctcaccaatggcgctcaagcaccggattggct
oDO32	agccaatccggtgcttgagcgccattggtgagaataa
oDO41	atatggctgcgccagctgctttttgcgtctattacgac
oDO42	gtcgtaatagacgcaaaaagcagctggcgcagccatat
oDO43	taagcgctgcgccagctgctttttgcgtctattacgac
oDO44	gtcgtaatagacgcaaaaagcagctggcgcagcgctta
oDO45	gcgctgcgccagctagaagctgcgtctattacgacgg
oDO46	ccgtcgtaatagacgcagcttctagctggcgcagcgc
oGT013	catatggcgccagctagatcttgcgtctattacgacggc
oGT014	gccgtcgtaatagacgcaagatctagctggcgccatatg
oGT015	atgtacgttagaatcggcaatacagcgactattacggcc
oGT016	ggccgtaatagtcgctgtattgccgattctaacgtacat
oGT09	gtcctgcttatgtacgttggaatcggcactacagcga
oGT10	tcgctgtagtgccgattccaacgtacataagcaggac
oCS84	ggggaccactttgtacaagaaagctgggtcCTAGTAAGGCTCGGCAGCAAGC
oTK333	tatcatatggctGCGCCAGCTAGATTTTGCGTCTAT
oTK334	tatggatccCTAGTAAGGCTCGGCAGCAAG
oTK344	AGGACCCAATCTTCAAAATGCATTTTTCGACAATTTTCATCCC
oTK345	AATGTTGAGTGGAATGATGCGGCTAGTAAGGCTCGGCAGCAAG
oTK409	ggggacaagtttgtacaaaaaagcaggcttaATGGCGCCAGCTAGATcTTGC
oTK439	tatcatatggctGCGCCAGCTAGATcTTGCGTCTAT
oTK472	attaCATATGCAGCGTACCAAAATTGTTGTTAAAG
oTK473	attaGGATCCtcaTTTTTTCACGCTTTCGACAACCA

Supplemental table 4. Primers

Use	Plasmid name	Construct	Plasmid backbone	Insert	Cloning method	Primers	PCR template or pENTRY	Reference
Entry clones	pSC060	AVR-Pia	pDONOR207	AVR-Pia_20-85	-	-	-	Césari et al. 2013
	pSC41	RGA4 cDNA	pDONOR207	RGA4	-	-	-	Césari et al. 2013
	pSC42	RGA5 cDNA	pDONOR207	RGA5	-	-	-	Césari et al. 2013
	pSC120	PWL2	pDONOR207	PWL2	-	-	-	Césari et al. 2013
	pSC129	$RGA5_{C\text{-}ter}$	pDONOR207	RGA5_882-1116	-	-	-	Césari et al. 2014
	pSC207	$RGA5_{RATX1}$	pDONOR207	RGA5_997-1072	-	-	-	Césari et al. 2014
	pSC210	$RGA5_{RATX1}$	pDONOR207	RGA5_1-996	-	-	-	Césari et al. 2014
	pDO01	AVR-Pia R23A	pDONOR207	AVR-Pia R23A	Quick change	oDO01/oDO02	pSC060	this study
	pDO02	AVR-Pia M40A	pDONOR207	AVR-Pia M40A	Quick change	oDO03/oDO04	pSC060	this study
	pDO03	AVR-Pia Y41A	pDONOR207	AVR-Pia Y41A	Quick change	oDO05/oDO06	pSC060	this study
	pDO04	AVR-Pia Y85A	pDONOR207	AVR-Pia Y85A	Quick change	oDO07/oDO08	pSC060	this study
	pDO09	AVR-Pia R43G	pDONOR207	AVR-Pia R43G	Quick change	oGT09/oGT10	pSC060	this study
	pDO52	AVR-Pia D29A	pDONOR207	AVR-Pia D29A	Quick change	oDO17/oDO18	pSC060	this study
	pDO53	AVR-Pia R36A	pDONOR207	AVR-Pia R36A	Quick change	oDO19/oDO20	pSC060	this study
	pDO54	AVR-Pia E56A	pDONOR207	AVR-Pia E56A	Quick change	oDO21/oDO22	pSC060	this study
	pDO55	AVR-Pia E58A	pDONOR207	AVR-Pia E58A	Quick change	oDO23/oDO24	pSC060	this study
	pDO56	AVR-Pia D63A	pDONOR207	AVR-Pia D63A	Quick change	oDO25/oDO26	pSC060	this study
	pDO57	AVR-Pia K67A	pDONOR207	AVR-Pia K67A	Quick change	oDO27/oDO28	pSC060	this study
	pDO58	AVR-Pia D78A	pDONOR207	AVR-Pia D78A	Quick change	oDO29/oDO30	pSC060	this study
	pDO59	AVR-Pia K74A	pDONOR207	AVR-Pia K74A	Quick change	oDO31/oDO32	pSC060	this study
	pCV84	AVR-Pia F24S	pDONOR207	AVR-Pia F24S	Gateway BP	oCS84/oTK409	pSC060	this study
Yeast two hybrid	pDO50	AD-empty	pGADT7-GW	no insert	-	-	-	Bernoux et al., 2011
	pDO51	BD-empty	pGBKT7-GW	no insert	-	-	-	Bernoux et al., 2011
	pSC003	**BD**-AVR-Pia	pGBKT7	AVR-Pia_20-85	-	-	-	Césari et al. 2013
	pDO49	**BD**-$RGA5_{C\text{-}ter}$	pGBKT7-GW	RGA5_882-1116	Gateway LR	-	pSC129	this study
	pDO38	AD-$RGA5_{C\text{-}ter}$	pGADT7-GW	RGA5_882-1116	Gateway LR	-	pSC129	this study
	pDO48	BD-AVR-Pia wt	pGBKT7- GW	AVR-Pia wt	Gateway LR	-	pSC060	this study
	pDO39	BD-AVR-Pia R23A	pGBKT7- GW	AVR-Pia R23A	Gateway LR	-	pDO01	this study
	pDO40	BD-AVR-Pia M40A	pGBKT7- GW	AVR-Pia M40A	Gateway LR	-	pDO02	this study
	pDO41	BD-AVR-Pia Y41A	pGBKT7- GW	AVR-Pia Y41A	Gateway LR	-	pDO03	this study
	pDO42	BD-AVR-Pia Y85A	pGBKT7- GW	AVR-Pia Y85A	Gateway LR	-	pDO04	this study
	pDO47	BD-AVR-Pia R43G	pGBKT7- GW	AVR-Pia R43G	Gateway LR	-	pDO09	this study
	pDO69	**BD**-AVR-Pia D29A	pGBKT7- GW	AVR-Pia D29A	Gateway LR	-	pDO52	this study
	pDO70	**BD**-AVR-Pia R36A	pGBKT7- GW	AVR-Pia R36A	Gateway LR	-	pDO53	this study
	pDO71	**BD**-AVR-Pia E56A	pGBKT7- GW	AVR-Pia E56A	Gateway LR	-	pDO54	this study

	pDO72	**BD**-AVR-Pia E58A	pGBKT7- GW	AVR-Pia E58A	Gateway LR	-	pDO55	this study
	pDO73	**BD**-AVR-Pia D63A	pGBKT7- GW	AVR-Pia D63A	Gateway LR	-	pDO56	this study
	pDO74	**BD**-AVR-Pia K67A	pGBKT7- GW	AVR-Pia K67A	Gateway LR	-	pDO57	this study
	pDO75	**BD**-AVR-Pia D78A	pGBKT7- GW	AVR-Pia D78A	Gateway LR	-	pDO58	this study
	pDO76	**BD**-AVR-Pia K74A	pGBKT7- GW	AVR-Pia K74A	Gateway LR	-	pDO59	this study
	pGT22	**BD**-AVR-Pia F24S	pGBKT7	AVR-Pia F24S	Quick change	oGT013/oGT14	pSC003	this study
	pGT23	**BD**-AVR-Pia T46N	pGBKT7	AVR-Pia T46N	Quick change	oGT015/oGT16	pSC003	this study
***N. benthamiana* cell**	pCV129	HA-RGA5	pBIN19 3XHA-GTW	RGA5	-	-	-	César et al. 2014
&	pDO120	**HA**-RGA5$_{RATX1}$	pBIN19 3XHA-GTW	RGA5_997-1072	Gateway LR	-	pSC207	this study
coIP	pSC144	HA-RGA5$_{C\text{-}ter}$	pBIN19 3XHA-GTW	RGA5_882-1116	Gateway LR	-	pSC129	this study
	pDO121	**HA**-RGA5$_{\Delta RATX1}$	pBIN19 3XHA-GTW	RGA5_1-996	Gateway LR	-	pSC210	this study
***N. benthamiana* cell**	pSC61	RGA4-3XHA	pBIN19 GTW-3XHA	RGA4	-	-	-	César et al. 2014
	pSC95	AVR-Pia -stop-**3XHA**	pBIN19 GTW-3XHA	AVR-Pia	-	-	-	César et al. 2014
	pCV91	AVR-Pia F24S-stop-**3X**	pBIN19 GTW-3XHA	AVR-Pia F24S	Gateway LR	-	pCV84	this study
	pDO10	AVR-Pia R23A-stop-3X	pBIN19 GTW-3XHA	AVR-Pia R23A	Gateway LR	-	pDO01	this study
	pDO11	AVR-Pia M40A-stop-**3**	pBIN19 GTW-3XHA	AVR-Pia M40A	Gateway LR	-	pDO02	this study
	pDO12	AVR-Pia Y41A-stop-**3X**	pBIN19 GTW-3XHA	AVR-Pia Y41A	Gateway LR	-	pDO03	this study
	pDO13	AVR-Pia Y85A-stop-**3X**	pBIN19 GTW-3XHA	AVR-Pia Y85A	Gateway LR	-	pDO04	this study
	pDO18	AVR-Pia R43G-stop-**3**	pBIN19 GTW-3XHA	AVR-Pia R43G	Gateway LR	-	pDO09	this study
	pDO103	AVR-Pia D29A-stop-**3**	pBIN19 GTW-3XHA	AVR-Pia D29A	Gateway LR	-	pDO52	this study
	pDO104	AVR-Pia R36A-stop-**3X**	pBIN19 GTW-3XHA	AVR-Pia R36A	Gateway LR	-	pDO53	this study
	pDO105	AVR-Pia E56A-stop-**3X**	pBIN19 GTW-3XHA	AVR-Pia E56A	Gateway LR	-	pDO54	this study
	pDO106	AVR-Pia E58A-stop-**3X**	pBIN19 GTW-3XHA	AVR-Pia E58A	Gateway LR	-	pDO55	this study
	pDO107	AVR-Pia D63A-stop-**3**	pBIN19 GTW-3XHA	AVR-Pia D63A	Gateway LR	-	pDO56	this study
	pDO108	AVR-Pia K67A-stop-**3X**	pBIN19 GTW-3XHA	AVR-Pia K67A	Gateway LR	-	pDO57	this study
	pDO109	AVR-Pia D78A-stop-**3**	pBIN19 GTW-3XHA	AVR-Pia D78A	Gateway LR	-	pDO58	this study
	pDO110	AVR-Pia K74A-stop-**3X**	pBIN19 GTW-3XHA	AVR-Pia K74A	Gateway LR	-	pDO59	this study
***N. benthamiana* col**	pDO119	**YFP**-Pwl2	pBIN19-YFP-GTW	Pwl2	Gateway LR	-	pSC120	this study
	pSC310	YFP	pBIN19-YFP-GTW	YFP	-	-	-	César et al. 2014
	pSC80	**YFP**-AVR-Pia	pBIN19 YFP-GTW	**YFP**-AVR-Pia	Gateway LR	-	pSC060	this study
	pDO19	**YFP**-AVR-Pia R23A	pBIN19 YFP-GTW	AVR-Pia R23A	Gateway LR	-	pDO01	this study
	pDO20	**YFP**-AVR-Pia M40A	pBIN19 YFP-GTW	AVR-Pia M40A	Gateway LR	-	pDO02	this study
	pDO21	**YFP**-AVR-Pia Y41A	pBIN19 YFP-GTW	AVR-Pia Y41A	Gateway LR	-	pDO03	this study
	pDO22	**YFP**-AVR-Pia Y85A	pBIN19 YFP-GTW	AVR-Pia Y85A	Gateway LR	-	pDO04	this study
	pDO27	**YFP**-AVR-Pia R43G	pBIN19 YFP-GTW	AVR-Pia R43G	Gateway LR	-	pDO09	this study
	pDO111	**YFP**-AVR-Pia D29A	pBIN19-YFP-GTW	AVR-Pia D29A	Gateway LR	-	pDO52	this study
	pDO112	**YFP**-AVR-Pia R36A	pBIN19-YFP-GTW	AVR-Pia R36A	Gateway LR	-	pDO53	this study
	pDO113	**YFP**-AVR-Pia E56A	pBIN19-YFP-GTW	AVR-Pia E56A	Gateway LR	-	pDO54	this study
	pDO114	**YFP**-AVR-Pia E58A	pBIN19-YFP-GTW	AVR-Pia E58A	Gateway LR	-	pDO55	this study

	pDO115	**YFP**-AVR-Pia D63A	pBIN19-YFP-GTW	AVR-Pia D63A	Gateway LR	-	pDO56	this study
	pDO116	**YFP**-AVR-Pia K67A	pBIN19-YFP-GTW	AVR-Pia K67A	Gateway LR	-	pDO57	this study
	pDO117	**YFP**-AVR-Pia D78A	pBIN19-YFP-GTW	AVR-Pia D78A	Gateway LR	-	pDO58	this study
	pDO118	**YFP**-AVR-Pia K74A	pBIN19-YFP-GTW	AVR-Pia K74A	Gateway LR	-	pDO59	this study
***E-coli* protein expre**	pCV64	AVR-Pia	pET15b	AVR-Pia	-	-	-	de Guillen et al., 2016
	pCV184	$RGA5_{RATX1}$	pET15b	RGA5_995-1069	restriction NdeI-BamHI	oTK472/oTK473	pTK207	this study
	pCV148	AVR-Pia_H3	pET15b	AVR-Pia_H3	restriction NdeI-BamHI	oTK439/oTK334	pGT5	this study
	pCV147	AVR-Pia F24S	pET15b	AVR-Pia F24S	restriction NdeI-BamHI	oTK439/oTK334	pSC60	this study
	pCV150	AVR-Pia T46N	pET15b	AVR-Pia T46N	restriction NdeI-BamHI	oTK333/oTK334	pGT5	this study
	pDO83	AVR-Pia R23A	pET15b	AVR-Pia R23A	Quick change	oDO41/oDO42	pCV64	this study
	pDO84	AVR-Pia R43G	pET15b	AVR-Pia R43G	Quick change	oGT09/oGT10	pCV64	this study
	pDO85	AVR-Pia D63A	pET15b	AVR-Pia D63A	Quick change	oDO25/oDO26	pCV64	this study
	pDO86	AVR-Pia D29A	pET15b	AVR-Pia D29A	Quick change	oDO17/oDO18	pCV64	this study
	pDO87	AVR-Pia R36A	pET15b	AVR-Pia R36A	Quick change	oDO19/oDO20	pCV64	this study
	pDO88	AVR-Pia E58A	pET15b	AVR-Pia E58A	Quick change	oDO23/oDO24	pCV64	this study
***M. oryzae* transgeni**	pCV76	AVR-Pia	pDL02	AVR-Pia	Yeast gap repair cloning	oTK344/oTK345	pSC60	this study
	pCR17	mRFP	pDL02	mRFP	-	-	-	Ribot et al., 2013
	pDO89	AVR-Pia R23A	pDL02	AVR-Pia R23A	Quick change	oDO43/oDO44	pCV76	this study
	pDO90	AVR-Pia R43G	pDL02	AVR-Pia R43G	Quick change	oGT09/oGT10	pCV76	this study
	pDO91	AVR-Pia D63A	pDL02	AVR-Pia D63A	Quick change	oDO25/oDO26	pCV76	this study
	pDO92	AVR-Pia D29A	pDL02	AVR-Pia D29A	Quick change	oDO17/oDO18	pCV76	this study
	pDO93	AVR-Pia R36A	pDL02	AVR-Pia R36A	Quick change	oDO19/oDO20	pCV76	this study
	pDO94	AVR-Pia E58A	pDL02	AVR-Pia E58A	Quick change	oDO23/oDO24	pCV76	this study
	pDO95	AVR-Pia F24	pDL02	AVR-Pia F24S	Quick change	oDO45/oDO46	pCV76	this study

Supplemental Methods

Production and purification of recombinant proteins for ITC and NMR

The expression and purification of recombinant AVR-Pia mutant proteins were carried out as previously described for the AVR-Pia wild type (de Guillen et al., 2015). Protein expression was performed in *E.coli* BL21 (DE3) in autoinducing minimal media C-750501 (Studier, 2005) supplemented with 15NH4Cl, 13C3-glycerol and 13C6-glucose for NMR experiments. Transformed cells grew at 37°C for 8h then at 30°C for 16h. A standard purification for Histagged protein was used with these following purification buffers (A): 50 mM TrisHCl, pH 7.0, 300 mM NaCl, 1mM DTT, 0.1 mM Benzamidine; (B): A supplemented with 500 mM imidazole) then a size exclusion chromatography was performed in buffer A. Fractions containing the protein were identified by SDS-PAGE, pooled and concentrated in a Centricon®. The RATX1 domain does not contain tryptophan and a bicinchoninic acid (BCA) assay (Pierce®) was used for protein quantification.

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were carried out on a VP-ITC isothermal titration calorimeter (Microcal, Northampton, USA) at 25°C. The protein samples were all buffer exchanged using dialysis at 4°C into the ITC buffer (20 mM potassium-sodium phosphate, pH 5.4 and 150 mM NaCl) to minimize undesirable buffer-related effects. The dialysis buffer was used in all preliminary equilibration and washing steps. Protein concentrations were measured using a NanoDrop 2000spectrometer and BCA Kit assays (Pierce®). The same RATX1 sample was used in all the binding reactions. Titration of RATX1 (80uM) in the cell (2mL) was performed by sequential addition of AVR-Pia (WT and H3 at 1 mM; 22 injections of 10uL). Data were analysed by a single site model using the Origin Software. A 1H-1D-NMR control spectrum was recorded for the AVR-Pia-H3 sample. All protein samples were checked by SDS-PAGE and by gel exclusion chromatography applied to the protein mixtures at the end of the ITC experiments.

NMR samples

The NMR samples for assignments and structure calculations were prepared with 1mM purified protein at 10% D2O and 0.5 mM DSS as a reference. The purification buffer was exchanged with phosphate buffer (20 mM potassiumsodium phosphate, pH 5.4 and 150 mM NaCl), by filtrating with Centricon®. For the D2O experiments, a 15Nlabeled sample was lyophilized and dissolved in D2O.

^{15}N backbone amide NMR Relaxation data

Relaxation data for AVR-Pia-H3 were acquired at 305K on a Bruker Avance 500 MHz spectrometer using R1, R2 and $^{15}N^{1}H$ heteronuclear NOE pulse sequences (TOPSPIN library, v 2.1) with a protein sample at 1 mM and the experimental settings already given in (de Guillen et al., 2015). Relaxation parameters, R1, R2 and NOEs were determined from the analysis module of CCPN (Vranken et al., 2005). Lipari-Szabo analysis was performed using the DYNAMOF software (Barthe et al., 2006).

Structure Calculation

The programs CYANA (Güntert, 2004), 2004) and CNS (Brunger, 2007) were used for automatic NOE assignments and structure calculations. The NH, Ha, ^{15}N, ^{13}Ca and ^{13}Cb chemical shifts were converted into F/Y dihedral angle constraints using TALOS+ (v. 1.2) (Shen et al., 2009). Final structure calculations were performed with CYANA (v. 2.1) using 1366 distance restraints and 100 F/Y dihedral angle constraints. The 20 conformers with lowest target function starting from 200 initial structures, were refined by CNS (v. 1.2) using the refinement in water of RECOORD (Nederveen et al., 2005). These are the structures discussed herein and deposited (PDBs, 5JHJ). The final 20 structures contained no NOE violations greater than 0.4 Å and no dihedral angle constraint violations greater than 2°. Structures were validated using PROCHECK(Laskowski et al., 1993) (Table S2).

AVR-Pia / RATX1 Docking Procedure

Three models of RATX1 have been obtained from the I-tasser server (http://zhanglab.ccmb.med.umich.edu/ITASSER), providing the chain B of the 5A6W PDB structure as an explicit homologous template. The obtained Itasser models were docked onto the AVR-Pia NMR structure (PDB identifier 2MYW) using the protein-protein docking program RosettaDock from the Rosetta modeling suite [2015.12.57698]. Each docking simulation was initialized by randomizing the relative positions and orientations of both partners and was performed with default parameters. The following AVR-Pia residues were restrained to lie closer than 12 Å from the RATX1 residues using Rosetta site constraints: R23, F24, V37, L39, M40, Y41, V42, R43, T46, T47, A48, T49, T51, T56 and E58. These residues correspond to the AVR-Pia central surface displaying the highest chemical shift differences from the NMR titration with RATX1. For each of the 3 RATX1 models, 3000 complexes were built with the same protocol. The 2% best scored complexes (180/9000) were clustered according to the orientation of AVR-Pia relatively to RATX1. All complex interfaces of the resulting clusters were evaluated using Pisa from the CCP4 suite. Some conclusions arising from these simulations could be stressed:

1) In all but one clusters, the central residue on AVR-Pia interface is M40, which is also central in the NMR mapped surface. The last cluster has a slightly displaced interface whose central residue is Y41.
2) The complex clusters display various interface orientations on RATX1 and none of them can be clearly distinguished neither from the number of poses, the Rosetta scores nor from the Pisa ΔGs.
3) None of the best clusters has an interface oriented as in the AVR-Pik / HMA complex. Other docking simulations done using the Zdock server (http://zdock.umassmed.edu) instead of Rosetta also failed to reproduce an interface oriented as seen in the AVR-Pik / HMA complex.

Protein extraction and Western blotting

Yeast proteins were extracted with a post-alkaline extraction method (Kushnirov, 2000), resuspended in Laemmli buffer, boiled 10 min at 90°C and separated in 12% Tris-tricine SDS-PAGE gels (GE Healthcare). After transfer to nitrocellulose membrane (Millipore), protein detection was performed using a-AVR-Pia antibodies (1000x dilution) raised against purified recombinant AVR-Pia protein in rabbits (Eurogentec).

Proteins from fungal culture filtrates were dialyzed against water (cut-off 1kDa, 1 part of medium by 5 parts of water), lyophilized, resuspended in 10 mM Tris-HCL pH 7.5, separated on 16% Tris-tricine NuPAGE gels (GE Healthcare) and, after blotting, analyzed using a-AVR-Pia antibodies.

SUPPLEMENTAL REFERENCES

Barthe, P., Ropars, V., and Roumestand, C. (2006). DYNAMOF: a program for the dynamics analysis of relaxation data obtained at multiple magnetic fields. Comptes Rendus Chim. **9**: 503– 513.

Brunger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. Nat. Protoc. **2**: 2728– 2733.

Güntert, P. (2004). Automated NMR structure calculation with CYANA. Methods Mol. Biol. **278**: 353– 378.

Kushnirov, V. V. (2000). Rapid and reliable protein extraction from yeast. Yeast **16**: 857–860.

Laskowski, R.A., Moss, D.S., and Thornton, J.M. (1993). Main-chain bond lengths and bond angles in protein structures. J. Mol. Biol. **231**: 1049–1067.

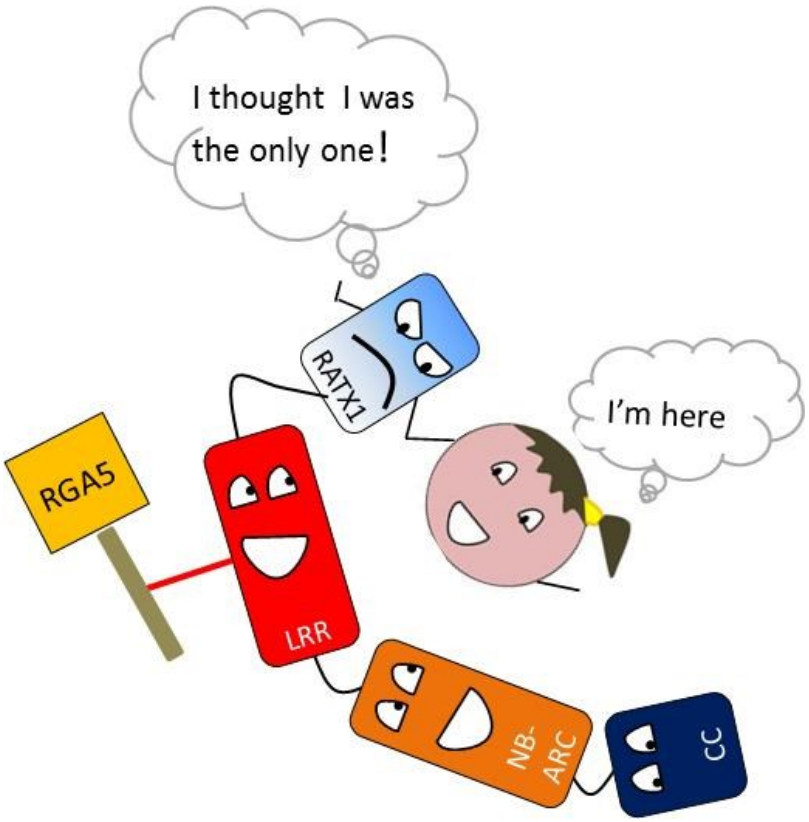
Nederveen, A.J. et al. (2005). RECOORD: A recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. Proteins Struct. Funct. Bioinforma. **59**: 662– 672.

Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J. Biomol. NMR **44**: 213–223.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: Development of a software pipeline. Proteins Struct. Funct. Bioinforma. **59**: 687– 696.

CHAPTER III

Analysis of the interaction of AVR-Pia with RGA5 domains other than RATX1



CONTEXT

In **Chapter II,** we describe the characterization of the AVR-Pia surface that directly binds to the RATX1 domain of the RGA5 immune receptor. We demonstrated that the AVR-Pia-RATX1 interaction is required for AVR-Pia recognition but that reduced binding affinity is well-tolerated. Next we showed that RATX1 is dispensable for the binding of AVR-Pia to RGA5. Indeed, the RATX1 deletion mutant $RGA5_{\Delta RATX1}$ strongly associates with AVR-Pia in *in planta* experiments. Based on these results, we proposed a model to explain a robust effector recognition mediated by integrated decoy domains in combination with additional effector-NLR interactions. Furthermore, Césari et al. 2013 showed that in addition to AVR-Pia, RGA4/RGA5 mediate recognition of AVR1-C039 and that the RATX1 domain is also involved in the binding of this effector protein. Furthermore, a weak interaction between $RGA5_{RATX1}$ and $RGA5_{NB\text{-}ARC}$ but not with $RGA5_{LRR}$ or $RGA5_{CC}$ was reported in yeast two hybrid experiments by Césari et al. 2014.

To evaluate the role of other RGA5 domains in effector recognition, we performed yeast two-hybrid (Y2H) and co-immunoprecipitation (co-IP) assays using different RGA5 domains: $RGA5_{\Delta RATX1}$, $RGA5_{CC}$, $RGA5_{NB}$, $RGA5_{CC\text{-}NB}$, $RGA5_{CC\text{-}NB\text{-}ARC}$, $RGA5_{NB\text{-}ARC}$, $RGA5_{LRR\text{-}C\text{-}ter}$ and $RGA5_{C\text{-}ter}$ (Figure 21) and we analyzed their interaction with the effector protein AVR-Pia and AVR1-CO39.

To investigate the role of RATX1 interactions with others RGA5 domain (Figure 21) in effector recognition and activation of the resistance, we analyzed the association of RATX1 domain with different RGA5 domains before and after AVR-Pia recognition by co-IP assays.

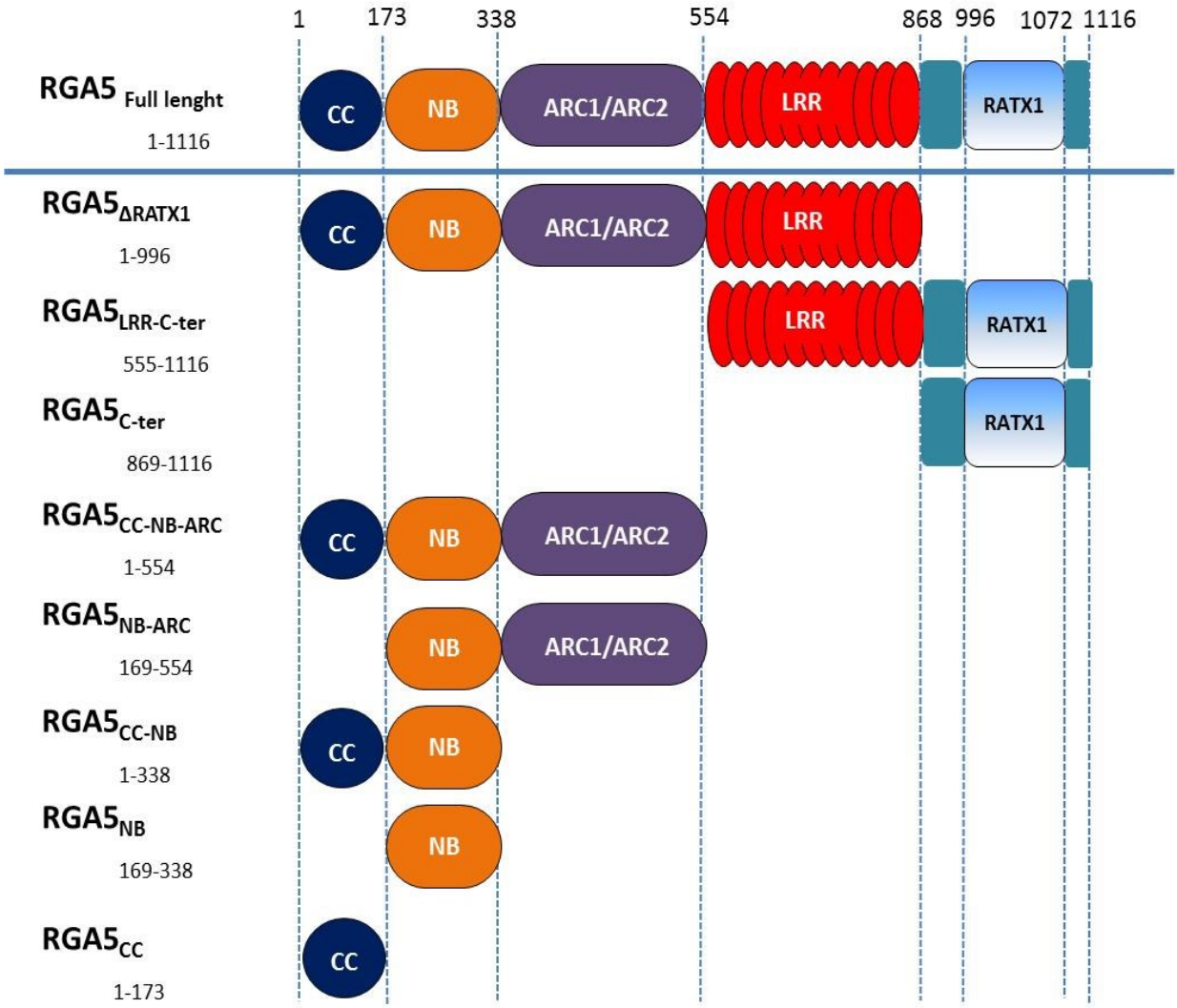


Figure 21. Borders of RGA5 domains analyzed for their interaction with AVR-Pia and AVR1-CO39

RESULTS

AVR-Pia associates with NB-ARC, LRR and RATX1 domains of RGA5

To analyze the interaction of AVR-Pia with other RGA5 domains than the RATX1 domain, BD and AD fusions of $RGA5_{\Delta RATX1}$, $RGA5_{CC}$, $RGA5_{NB}$, $RGA5_{CC\text{-}NB}$, $RGA5_{CC\text{-}NB\text{-}ARC}$, $RGA5_{NB\text{-}ARC}$ and $RGA5_{LRR\text{-}C\text{-}ter}$ were generated (Figure 21). AD and BD fusions of $RGA5_{C\text{-}ter}$ were used as positive controls and, as previously reported, yeasts co-expressing either BD-AVR-Pia and AD-$RGA5_{C\text{-}ter}$ or AD-AVR-Pia and BD-$RGA5_{C\text{-}ter}$ grew on selective medium indicating physical binding between AVR-Pia and $RGA5_{C\text{-}ter}$ (Figure 22). Yeasts co-expressing AD- or BD-AVR-Pia in combination with different AD- or BD-RGA5 domains, other than $RGA5_{C\text{-}ter}$, did not grow on selective medium (Figure 22) suggesting that AVR-Pia does not bind to those RGA5 domains in yeast. The BD- and AD-AVR-Pia fusions were correctly expressed and most of the BD- and AD-RGA5 domains were expressed at similar levels except for BD- and AD-$RGA5_{\Delta RATX1}$, BD- and $AD_{LRR\text{-}Cter}$, BD- $RGA5_{NB}$ and AD-$RGA5_{CC}$ that were not expressed (Figure 23). Therefore, for these constructs, the lack of yeast growth could be explained by a failed protein expression.

To test whether AVR-Pia associates with RGA5 domains *in planta*, we used HA-tagged RGA5 domains and YFP-tagged AVR-Pia. A YFP fusion of PWL2, a *M. oryzae* effector that does not interact with RGA5 was used as a negative control. All proteins were properly expressed and both YFP fusion proteins were efficiently precipitated with anti-GFP agarose beads (Figure 24). All RGA5 domains but $RGA5_{CC}$ were co-precipitated by YFP-AVR-Pia and not by PWL2, suggesting that AVR-Pia associates with $RGA5_{NB\text{-}ARC}$ *in planta*. Since the isolated LRR domain of RGA5 could not be expressed in *N. benthamiana* (data not shown) and the $RGA5_{LRR\text{-}C\text{-}ter}$ construct contains the C-ter domain that binds AVR-Pia we could not conclude whether AVR-Pia also interacts with the RGA5 LRR domain. To address this question, we tested the association with AVR-Pia^{F24S} that does not interact with $RGA5_{Cter}$. For this, we tested the co-IP of HA-tagged RGA5 fragments by YFP-AVR-Pia^{F24S} in comparison with the positive and negative controls YFP-AVR-Pia^{WT} and YFP-PWL2. The three YFP fusion proteins were well precipitated. YFP-AVR-Pia^{F24S} co-precipitated $RGA5_{LRR\text{-}Cter}$ as did YFP-AVR-Pia^{WT} but failed to co-precipitate $RGA5_{C\text{-}ter}$, suggesting that AVR-Pia associates with the LRR domain of RGA5. As observed for YFP-AVR-Pia^{WT}, $RGA5_{CC}$ was not co-precipitated by YFP-AVR-Pia^{F24S}. Strikingly,

YFP-AVR-Pia^{F24S} efficiently co-precipitate RGA5$_{CC\text{-}NB}$ and RGA5$_{CC\text{-}NB\text{-}ARC}$ but not RGA5$_{NB}$ (Figure 25).

Altogether, these results indicate that AVR-Pia associates with the RATX1, NB-ARC and LRR domains of RGA5 in *in planta* assays. For the NB-ARC and LRR domains, these interactions were not detected in yeast two hybrid assays. Therefore, it remains to be elucidated by other approaches whether additional proteins are required for these interactions or whether they rely on direct physical binding.

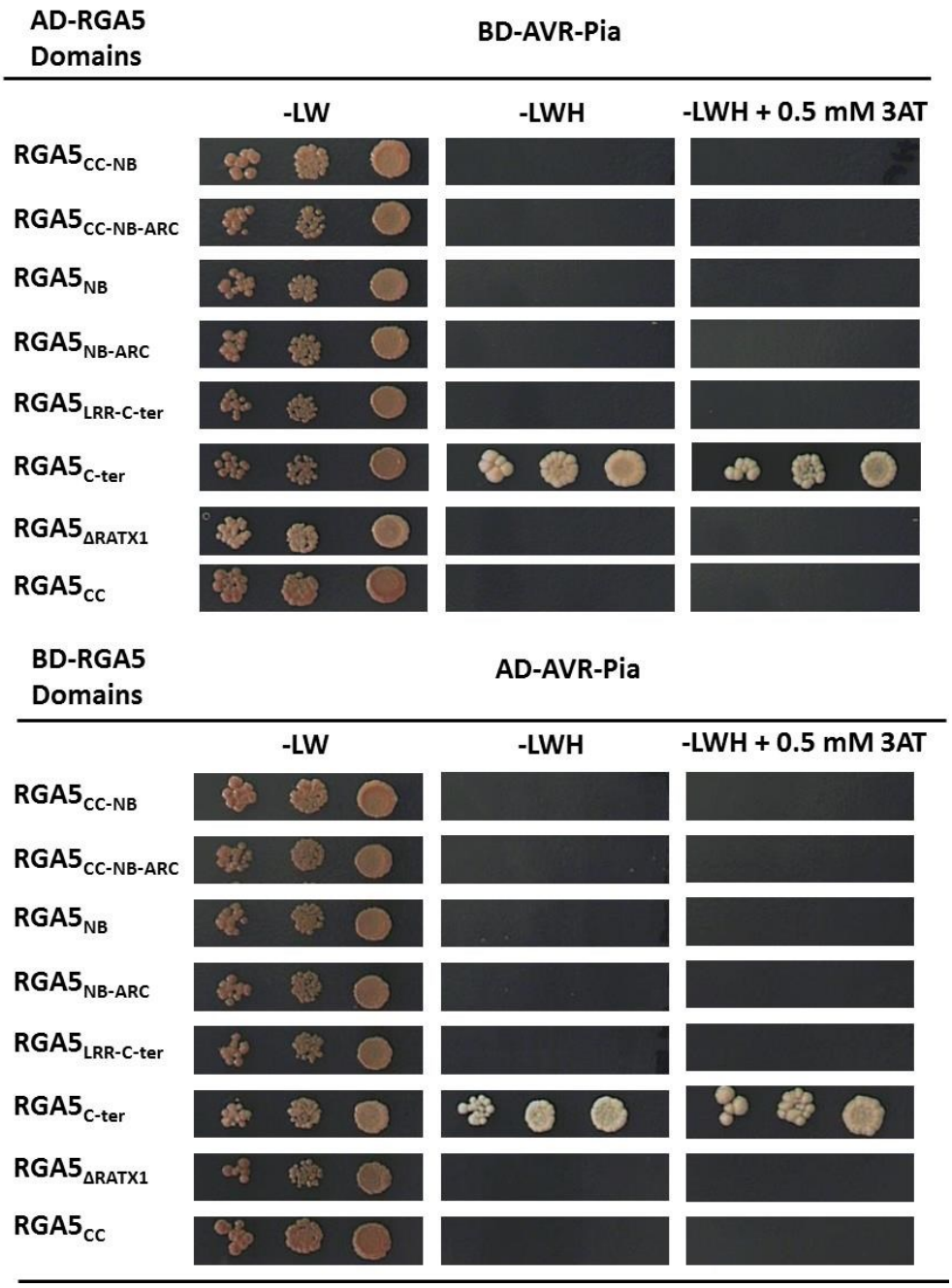


Figure 22. AVR-Pia binds specifically to the RGA5$_{C\text{-}ter}$ domain in yeast. The interaction between AD- and BD- AVR-Pia fusion proteins and different BD and AD-RGA5 fusion domains respectively was tested by a yeast two-hybrid experiment. Three dilutions (1/10, 1/100, 1/1000) of yeast cultures adjusted to an OD of 0.2 were spotted on synthetic double drop out (DDO) medium (-LW) to control proper growth and on synthetic TDO (-LWH) either without or supplemented with 3-amino-1,2,4-triazole (3AT) to test the strength of the interaction. Photos were taken after 4 days of growth.

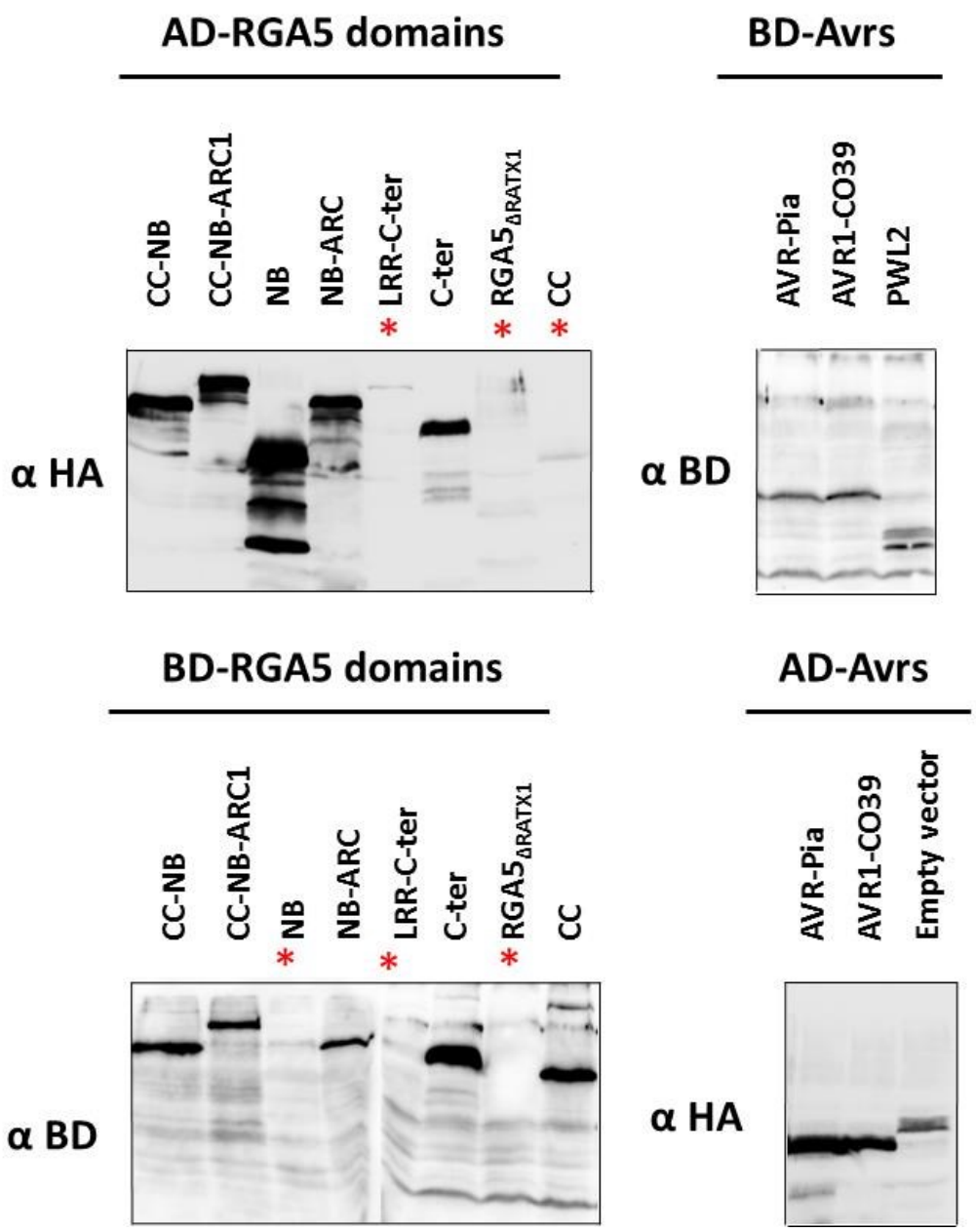


Figure 23. Western blotting of transgenic yeast protein extracts. Equal production of AD and BD-AVR-Pia and AVR1-CO39 fusion proteins was determined by immunoblotting with anti-HA and anti-BD antibodies respectively. Similar quantities of AD and BD-RGA5 fusion domains were specifically detected with anti-HA and anti-BD antibodies, except for the constructs indicated with one asterisk.

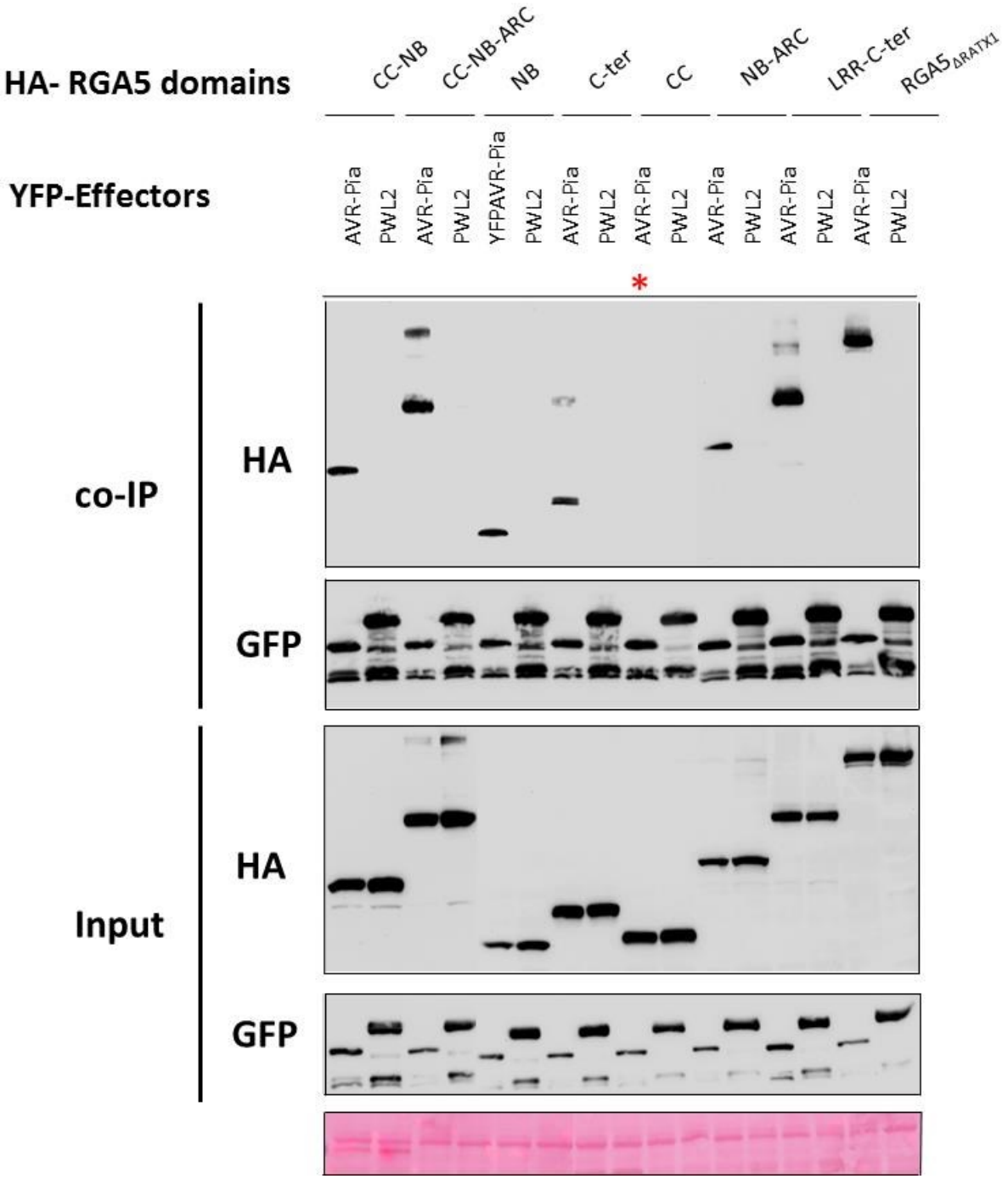


Figure 24. Different RGA5 domains associate with AVR-Pia *in planta*. HA-RGA5 domains were transiently expressed with YFP-AVR-Pia or YFP-PWL2 in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection of immunoprecipitated AVR-Pia and PWL2. Co-precipitated HA-RGA5 domains proteins were detected using α-HA antibody. RGA5 domains not co-precipitated by AVR-Pia are indicated by one asterisk

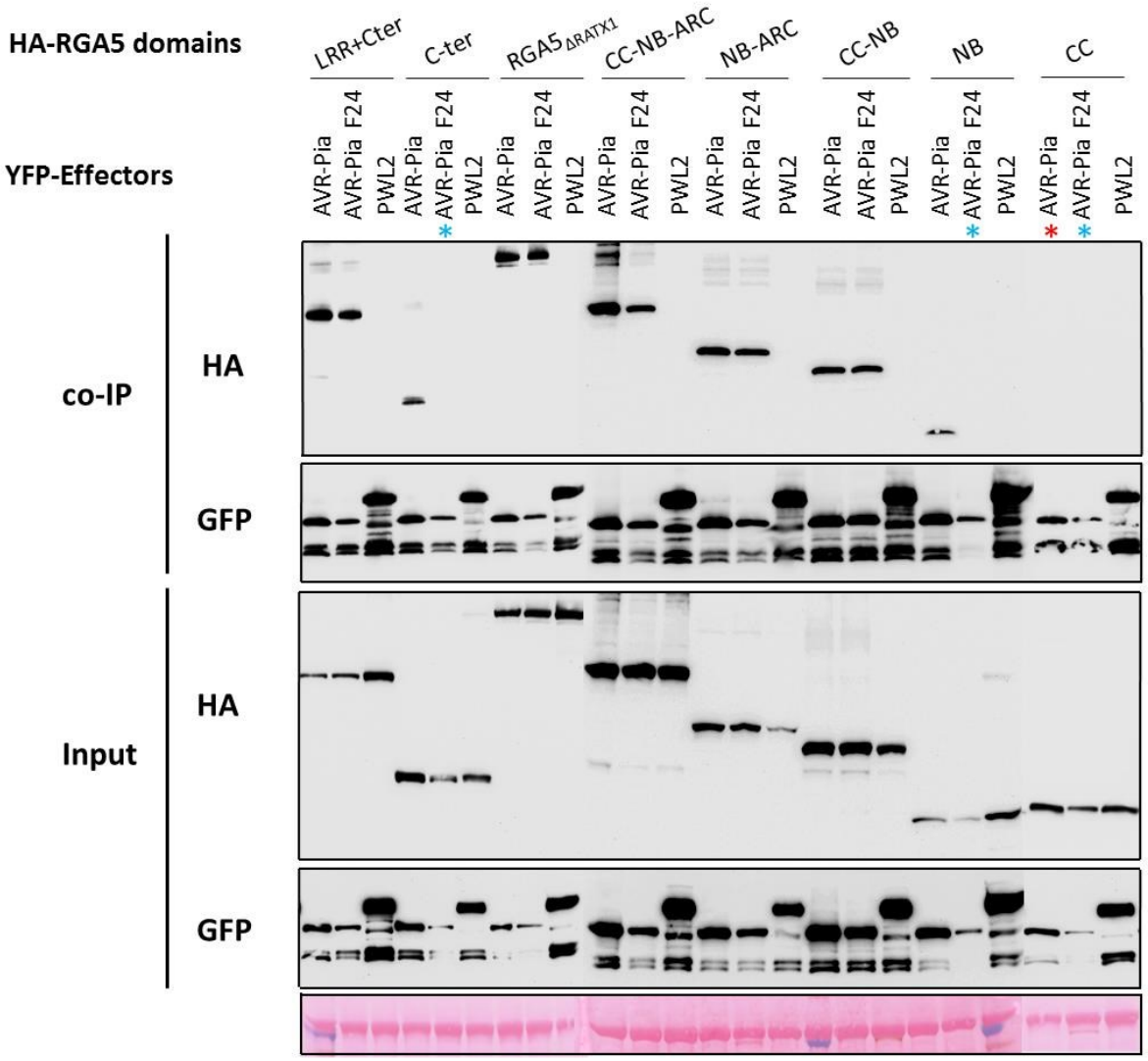


Figure 25. AVR-Pia associates with RGA5$_{LRR}$, RGA5$_{NB-ARC}$ and RGA5$_{C-ter}$ domains *in planta*. HA-RGA5 domains were transiently expressed with YFP-AVR-Pia, YFP-AVR-Pia^{F24S} or YFP-PWL2 in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection of immunoprecipitated AVR-Pia, AVR-Pia^{F24S} and PWL2. Co-precipitated HA-RGA5 domains proteins were detected using α-HA antibody. RGA5 domains not co-precipitated by AVR-Pia or AVR-Pia^{F24S} are indicated by one asterisk red and blue respectively.

The NB-ARC, LRR and RATX1 domains of RGA5 interact with two different *M. oryzae* effector proteins

To analyze whether AVR1-CO39 interacts also with additional RGA5 domains, yeast two hybrid experiments were performed with BD and AD fusions of $RGA5_{\Delta RATX1}$, $RGA5_{CC}$, $RGA5_{NB}$, $RGA5_{CC\text{-}NB}$, $RGA5_{CC\text{-}NB\text{-}ARC}$, $RGA5_{NB\text{-}ARC}$ and $RGA5_{LRR\text{-}C\text{-}ter}$ (Figure 21). AD and BD fusions of $RGA5_{C\text{-}ter}$ were used as positive controls.

Yeast co-expressing BD-AVR1-CO39 with AD-$RGA5_{CC\text{-}NB}$ and with AD-$RGA5_{CC\text{-}NB\text{-}ARC}$ efficiently grew on elective medium and to similar levels as the positive control (Figure 26). In contrast, yeast expressing BD-AVR1-CO39 in combination with AD-$RGA5_{NB}$ showed a very weak growth suggesting that AVR1-CO39 interacts with the NB domain of RGA5 and that the ARC domain is important to stabilize this interaction or establishes additional interactions with the effector (Figure 26). Under higher selection stringency, on –LWH + 3AT medium, only yeasts co-expressing BD-AVR1-CO39 and AD-$RGA5_{CC\text{-}NB\text{-}ARC}$ grew but this growth was weaker than that of yeasts co-expressing BD-AVR1-CO39 and AD-$RGA5_{C\text{-}ter}$ (Figure 26), suggesting that at least in yeast the interaction of AVR1-CO39 with $RGA5_{C\text{-}ter}$ is stronger than with the NB-ARC domain.

Results with AD-$RGA5_{LRR\text{-}Cter}$, AD-$RGA5_{CC}$ or AD-$RGA5_{\Delta RATX1}$ are not informative because there is insufficient protein production for these RGA5 constructs (Figure 23). The lack of growth observed for yeasts co-expressing BD-AVR1-CO39 and AD-$RGA5_{LRR\text{-}Cter}$, AD-$RGA5_{CC}$ or AD-$RGA5_{\Delta RATX1}$ therefore does not indicate lack of interaction. Strikingly, when we co-expressed AVR1-CO39 fused to AD with RGA5 domains fused to BD we only detected interaction between AD-AVR1-CO39 and BD-$RGA5_{C\text{-}ter}$ (Figure 26). A possible reason is that the AD domain interferes with the physical binding of AVR1-CO39 to the NB-ARC domain of RGA5 that was efficiently expressed in yeast (Figure 23).

To investigate the interaction between AVR1-CO39 and RGA5 domains *in planta* we performed co-IP of different HA- tagged RGA5 domains with YFP-tagged AVR1-CO39 or PWL2 with the same co-IP conditions used for AVR-Pia. All proteins were properly expressed and anti-GFP agarose beads efficiently immunoprecipitated YFP-AVR1-CO39 and YFP-PWL2 (Figure 27). Under these conditions, YFP-AVR1-CO39 exhibit strong association with HA-$RGA5_{CC\text{-}NB\text{-}ARC}$ and HA-$RGA5_{NB\text{-}ARC}$. Association with HA-$RGA5_{CC\text{-}NB}$ or HA-$RGA5_{NB}$ is also

observed but was very weak. With HA-RGA5$_{LRR\text{-}C\text{-}ter}$, HA-RGA5$_{CC}$ or HA-RGA5$_{C\text{-}ter}$ no association was observed (Figure 27). For the constructs containing the RATX1 domain, HA-RGA5$_{LRR\text{-}C\text{-}ter}$ and HA-RGA5$_{C\text{-}ter}$ this is surprising since RGA5$_{C\text{-}ter}$-AVR1-CO39 interaction was observed in Y2H and was reported previously (Césari et al., 2014c). However, in these *in planta* association studies between AVR1-CO39 and RGA5$_{C\text{-}ter}$ less stringent condition for co-IP experiments had been used (Cesari et al. 2013). Taken together these results suggest that AVR1-CO39 binds directly to RGA5$_{NB\text{-}ARC}$ in addition to RGA5$_{C\text{-}ter}$.

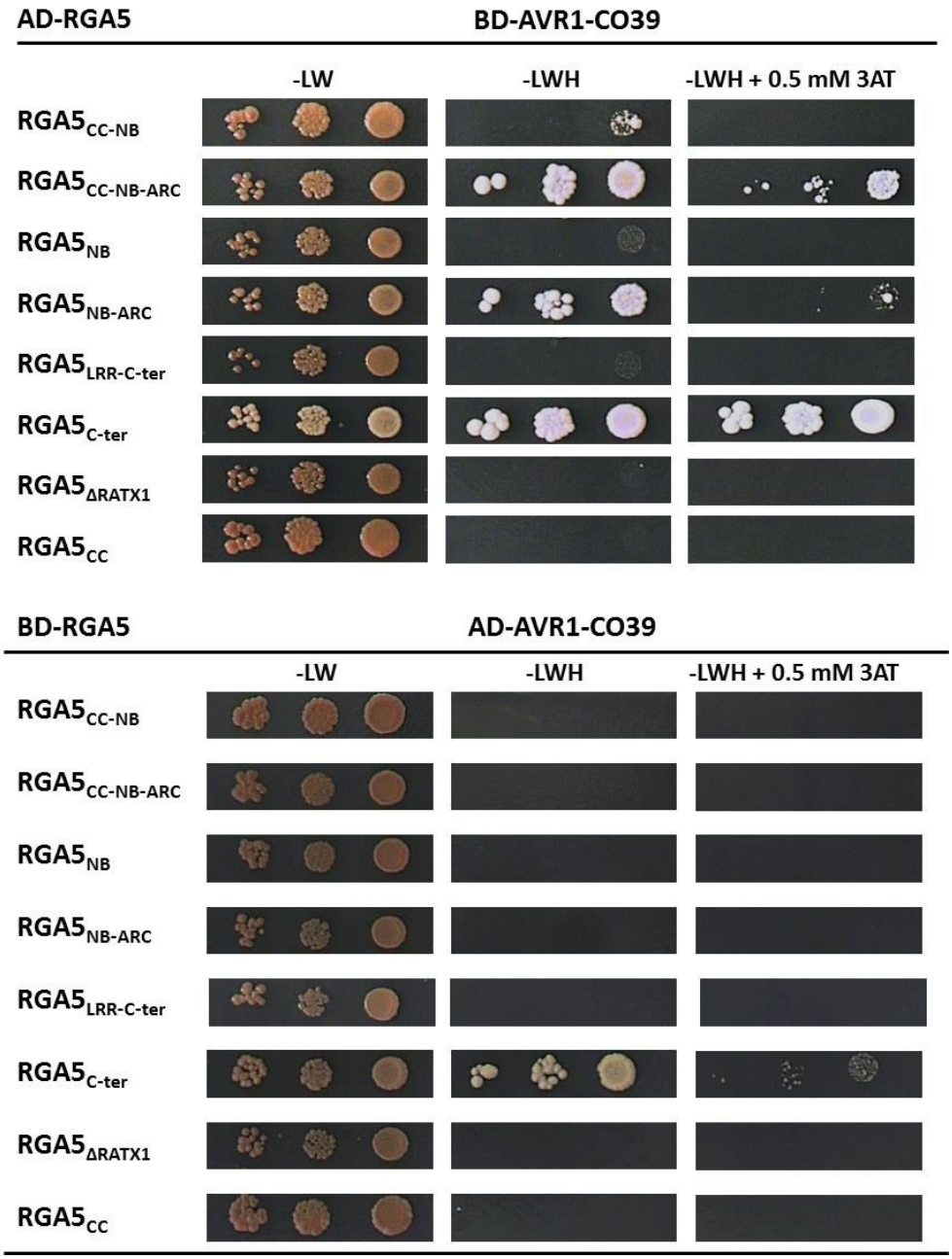


Figure 26. AVR1-CO39 binds to RGA5$_{NB-ARC}$ and RGA5$_{C-ter}$ domains in yeast. The interaction between AD- and BD- AVR-Pia fusion proteins and different BD and AD-RGA5 fusion domains respectively was tested by a yeast two-hybrid experiment. Three dilutions (1/10, 1/100, 1/1000) of yeast cultures adjusted to an OD of 0.2 were spotted on synthetic double drop out (DDO) medium (-LW) to control proper growth and on synthetic TDO (-LWH) either without or supplemented with 3-amino-1,2,4-triazole (3AT) to test the strength of the interaction. Photos were taken after 4 days of growth.

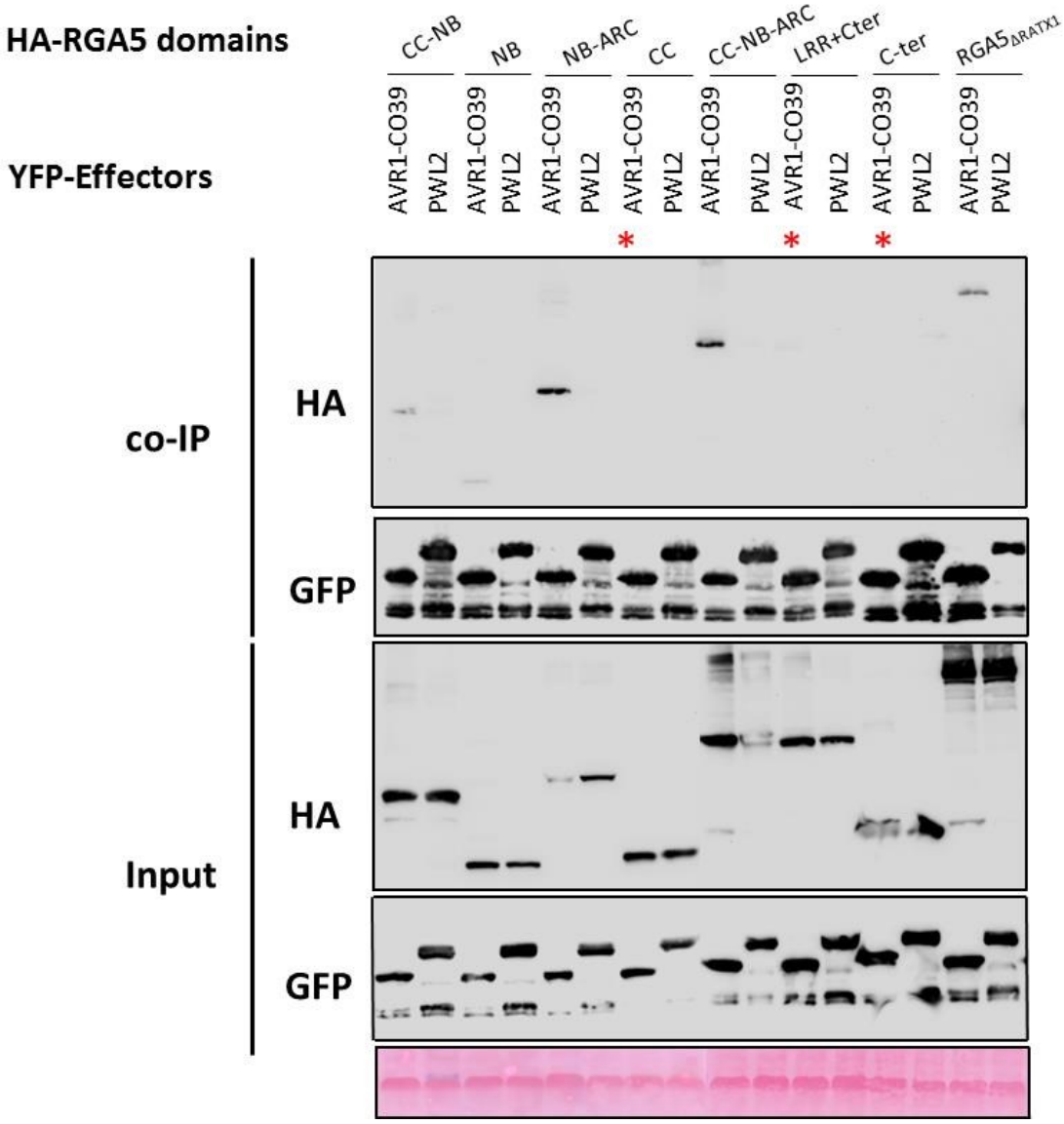


Figure 27. AVR1-CO39 associates with RGA5$_{NB-ARC}$ *in planta* and in yeast. HA-RGA5 domains were transiently expressed with YFP-AVR1-CO39 and YFP-PWL2 in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and

analyzed by immunoblotting with α-GFP for the detection of immunoprecipitated AVR1-CO39 and PWL2. Co-precipitated HA-RGA5 domains were detected using α-HA antibody. RGA5 domains not co-precipitated by AVR1-CO39 are indicated by one asterisk.

The RATX1 domain self-associates and also associates with the NB-ARC domain of RGA5

To analyze the interaction of the RATX1 domain with other domains of RGA5, HA-tagged RATX1 and YFP-tagged RGA5 domains ($RGA5_{C\text{-}term}$, $RGA5_{CC}$, $RGA5_{NB\text{-}ARC}$) were co-expressed in *N. benthamiana*. YFP-RGA5 domains were efficiently immunoprecipitated. The HA-RATX1 domain was efficiently co-precipitated by YFP-$RGA5_{C\text{-}ter}$ and YFP-$RGA5_{NB\text{-}ARC}$ but not YFP-$RGA5_{CC}$, indicating that RATX1 interacts with $RGA5_{NB\text{-}ARC}$ and self-associates which has also previously been found in yeast two hybrid analysis (Césari, et al. 2014) (Figure 28).

To determine whether the interactions between RATX1 and RGA5 domains are modified by AVR-Pia recognition, we tested the co-IP of HA-RATX1 and YFP-tagged RGA5 domains ($RGA5_{C\text{-}term}$, $RGA5_{CC}$, $RGA5_{NB\text{-}ARC}$ or $RGA5_{\Delta RATX1}$) in the presence of AVR-Pia. In the presence of AVR-Pia, the association between HA-$RGA5_{RATX1}$ with YFP-$RGA5_{NB\text{-}ARC}$ or YFP-$RGA5_{C\text{-}term}$ is weaker than in its absence (Figure 29), indicating that AVR-Pia modifies these interactions. However, the co-IP of HA-$RGA5_{RATX1}$ by YFP-$RGA5_{\Delta RATX1}$ was not affected in the presence of AVR-Pia potentially because association is stabilized by the LRR domain. Therefore, the association between RATX1 and LRR should be analyzed in the future.

To get a better view of the strength of the RATX1 self-interaction and the interaction of RATX1 with the NB-ARC domain we also used more stringent conditions for co-immunoprecipitation of HA-RATX1 with YFP-$RGA5_{NB\text{-}ARC}$, YFP-$RGA5_{\Delta RATX1}$, YFP-$RGA5_{CC}$ or YFP-$RGA5_{RATX1}$ (used here instead YFP-$RGA5_{C\text{-}ter}$,) in presence of AVR-Pia. Under these conditions, a strong association was only observed between HA-RATX1 and YFP-RATX1 in the absence of AVR-Pia (Figure 30). In addition, a weak association of HA-RATX1 with YFP-$RGA5_{NB\text{-}ARC}$ and YFP-$RGA5_{\Delta RATX1}$ was detected and in all cases the interactions were lost in the presence of AVR-Pia.

Taken together these results suggest that the RATX1 domain interacts with itself and with the NB-ARC domain of RGA5. Both, RATX1-RATX1 and RATX1-NB-ARC interactions seem to be affected by AVR-Pia.

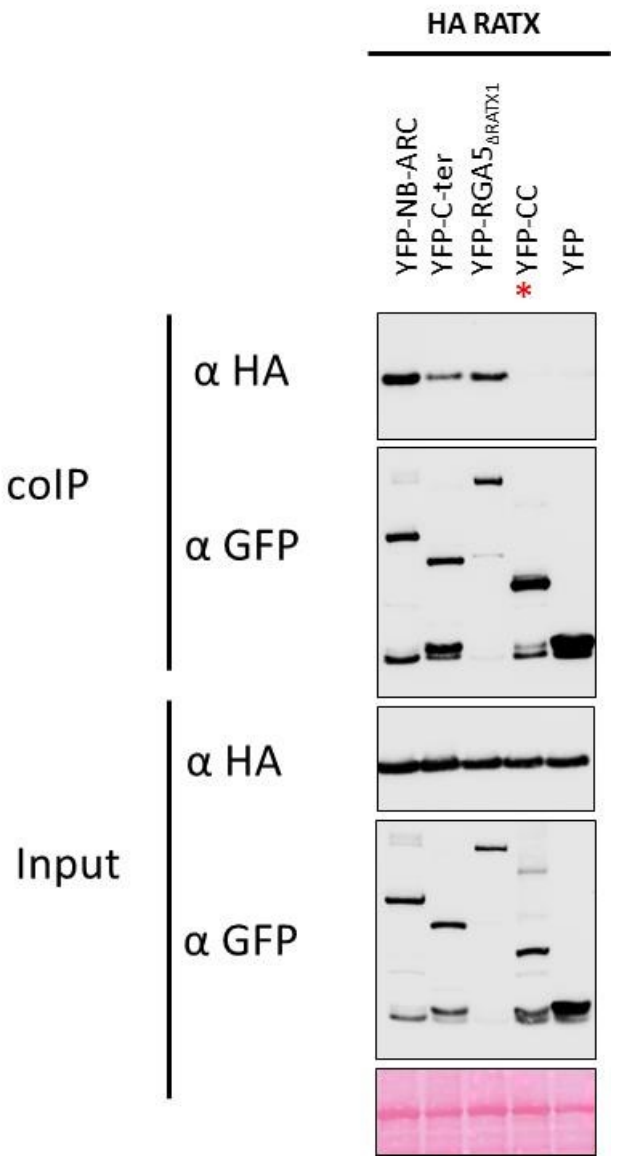


Figure 28. RATX1 associates with RGA5$_{NB-ARC}$ and RGA5$_{C-ter}$ but not with RGA5$_{CC}$ *in planta*. HA-RATX was transiently expressed with YFP-RGA5 domains and YFP in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection of immunoprecipitated RGA5 domains. Co-precipitated HA-RAXT1 was detected using α-HA antibody. RATX1 not co-precipitated by RGA5 domains is indicated by one asterisk.

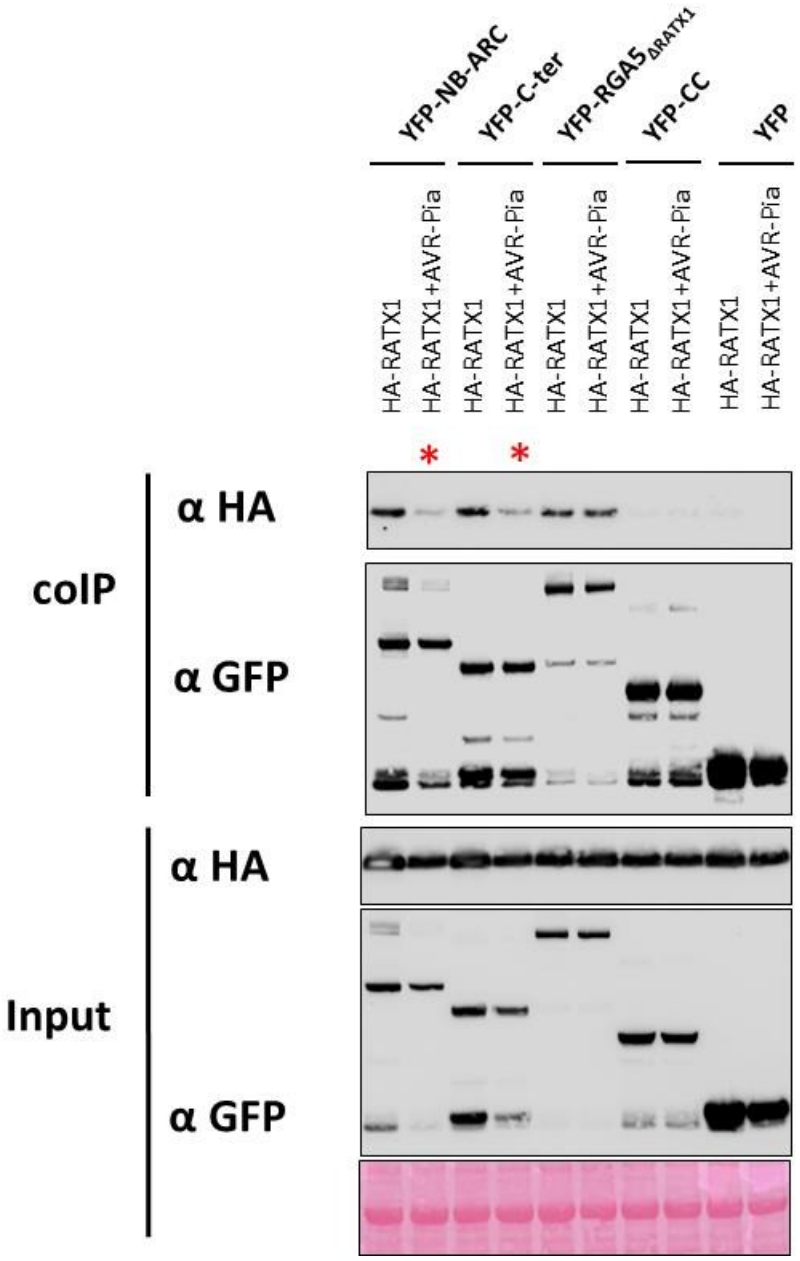


Figure 29. RATX1 association with RGA5$_{NB-ARC}$ and RGA5$_{C-ter}$ is affected by AVR-Pia. HA-RATX was transiently expressed with YFP-RGA5 domains and YFP in presence of AVR-Pia without tag in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection of immunoprecipitated RGA5 domains. Co-precipitated HA-RAXT1 was detected using α-HA antibody. The associations of RATX1 with RGA5 domains affected by AVR-Pia are indicated by one asterisk.

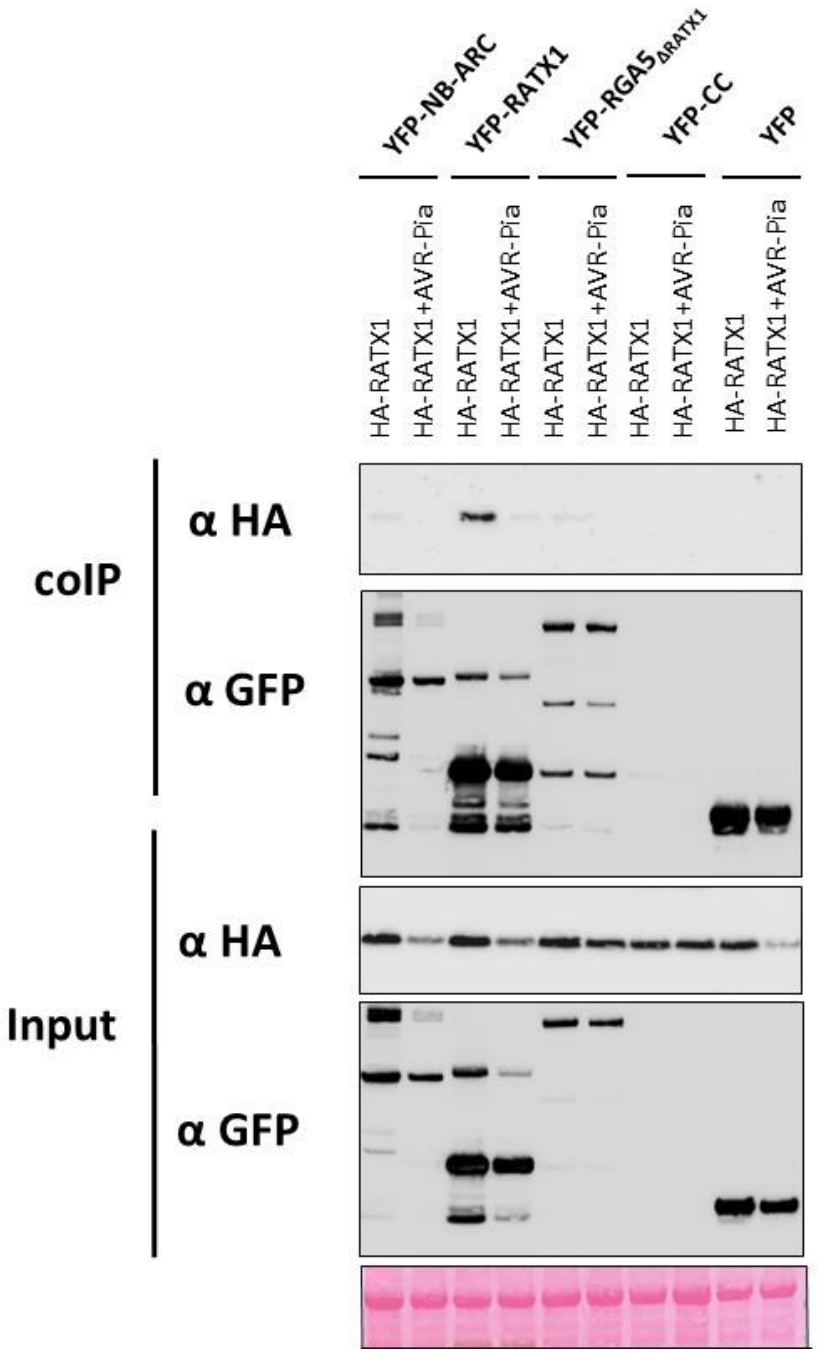


Figure 30. RATX1 strongly self-associates in absence of AVR-Pia. HA-RATX was transiently expressed with YFP-RGA5 domains and YFP in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection of immunoprecipitated RGA5 domains. Co-precipitated HA-RAXT1 was detected using α-HA antibody. *In this experiment the expression of YFP-CC failed.

DISCUSSION

The immune receptor RGA5 interacts with different effector proteins through multiple domains

In this study, we provide evidence that RGA5 interacts with AVR-Pia and AVR1-CO39 through its NB-ARC and RATX1 domains. In addition we showed that AVR-Pia also interact with the LRR domain of RGA5. Since association of AVR-Pia with $RGA5_{NB\text{-}ARC}$ and $RGA5_{LRR}$ was only detected in *in planta* assays and not in yeast, we cannot exclude that additional proteins are required to mediate this interaction and that the interaction is indirect. Interaction between AVR1-CO39 and $RGA5_{NB\text{-}ARC}$ was detected in both *in planta* and in yeast two hybrid assays suggesting that it relies on direct physical binding. Altogether, these results strongly support that both AVR-Pia and AVR1-CO39 bind not only to the RATX1 domain of RGA5 but also to its NB-ARC domain. In chapter II we propose a cooperative model for AVR-Pia-RGA5 interaction in which effector recognition depends on simultaneous binding to the integrated domain and other NLR domains. Here, we provide additional support for this model and show that at least one second domain, $RGA5_{NB\text{-}ARC,}$ binds or indirectly associates with the two structure-related effectors AVR-Pia and AVR1-CO39. This opens now the way to a more precise characterization of this interaction including the identification of the interaction surfaces and the individual residues that mediate this interaction. Subsequently, these residues can be mutated to verify if the effector-NB-ARC interactions are required for effector recognition and to determine their contribution to the recognition specificity. This would give also some clues about the question to what extent synchronous binding to the integrated decoy and the NB-ARC domain contributes to the extended recognition specificity of RGA5 that recognizes two different effectors.

The NACHT domain from animal immune receptors such us NAIP (Neuronal Apoptosis Inhibitor Protein) or CIITA (MHC class 2 transcription factor) is closely related to the NB-ARC domain (Leipe et al., 2004; van Ooijen et al., 2008). Interestingly, this domain confers ligand recognition specificity to NAIP proteins and e.g. NAIP2 directly interacts through its nucleotide-binding domain with PrgJ a type III secretion system (T3SS) component of the bacteria *Salmonella* (Tenthorey et al., 2014). Therefore it is tempting to speculate that like $NAIP_{NACHT}$, the $RGA5_{NB\text{-}ARC}$ domain has evolved the ability to interact with pathogen-derived

ligands which would be an interesting case of convergent evolution.

$RGA5_{RATX1}$ self-association and $RGA5_{RATX1}$- $RGA5_{NB-ARC}$ interaction is perturbed by AVR-Pia

The interaction of $RGA5_{RATX1}$ with $RGA5_{NB-ARC}$ and $RGA5_{RATX1}$ in the presence of AVR-Pia was characterized *in planta* by co-immunoprecipitation assays and showed that the interaction or $RGA5_{RATX1}$ with $RGA5_{NB-ARC}$ as well as its self-association is weakened by AVR-Pia. Interestingly, AVR-Pia associates with both domains suggesting that AVR-Pia interacts with the RATX1 surfaces mediating self-association. In addition, they suggest that AVR-Pia and RATX1 may interact with similar $RGA5_{NB-ARC}$ surfaces or that the AVR-Pia-binding surface in the RATX1 domain is also involved in $RGA5_{NB-ARC}$-binding. Since the RATX1 domain of RGA5 is not required for the repression of RGA4 it can be speculated that it is not the removal of the $RGA5_{RATX1}$-$RGA5_{NB-ARC}$ association by AVR-Pia that triggers activation of resistance. Rather, the simultaneous interaction of AVR-Pia with the RATX1 and the NB-ARC domain could force RGA5 in a conformation that no longer allows RGA4 repression and thereby triggers resistance. The role $RGA5_{RATX1}$ self-association in effector recognition requires further analyses since the stoichiometry of the RGA4-RGA5 complex is unknown and it is unclear whether it has biological relevance. It would be particularly interesting to better characterize the composition of the RGA4/RGA5 complex and to investigate whether $RGA5_{RATX1}$ mutants impaired in the formation of homo-complexes are affected in AVR-Pia or AVR1-CO39 recognition.

MATERIAL AND METHODS

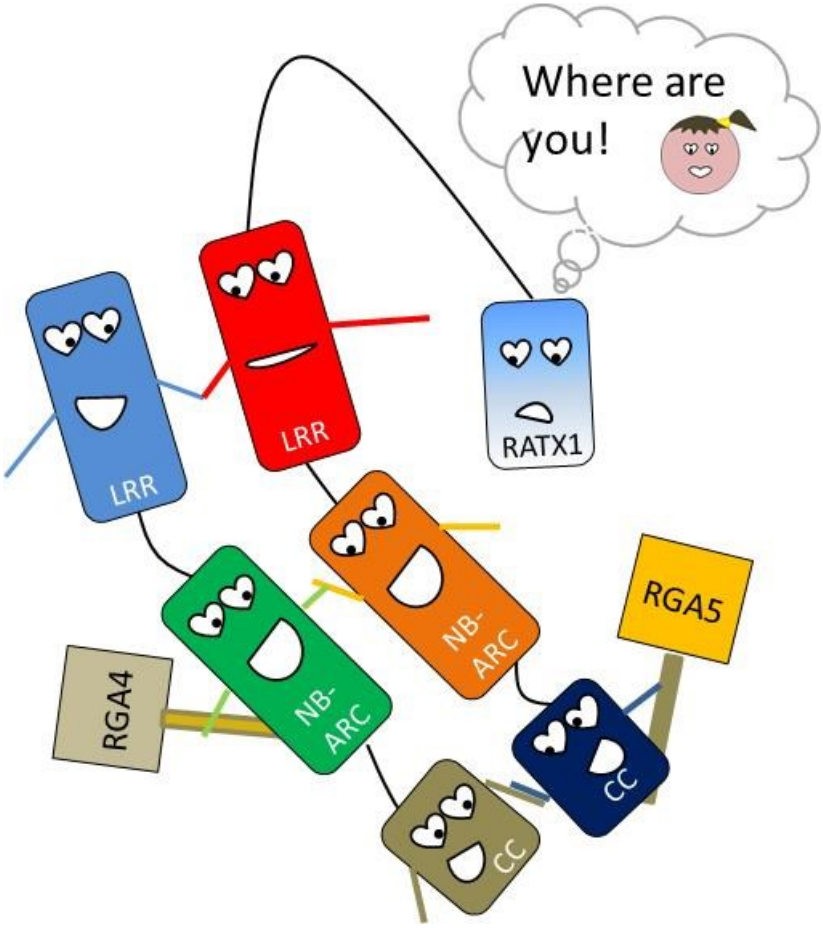
The material and methods of this chapter are detailed in the annex.

SUMMARY POINTS (CHAPTER II and III)

- ✓ AVR-Pia binds the RGA5 RATX1 domain through multiple residues located in a specific surface exposed region of the effector protein.
- ✓ *M. oryzae* effector proteins with homologous structures can bind similar domains through different protein surfaces.
- ✓ The RATX1-interaction surface of AVR-Pia undergoes mutations to escape recognition by the RGA4/RGA5 complex.
- ✓ The RATX1 domain is crucial to mediate recognition of AVR-Pia by the RGA5/RGA4 complexes.
- ✓ AVR-Pia binds the RGA5 RATX1 domain with intermediate affinity and reduced binding is well tolerated in recognition.
- ✓ The RATX1 domain is dispensable for the association of AVR-Pia with RGA4/RGA5 complexes.
- ✓ AVR-Pia associates with RGA5 domains other than RATX1.
- ✓ Simultaneous binding of AVR-Pia to different RGA5 domains could explain the high resilience of RGA4/RGA5-mediated resistance to mutations in the effector.

CHAPTER IV

Functional analyses of RGA4-RGA5 interaction



INTRODUCTION

One of the most amazing features of NLRs is their ability to confer resistance to a large number of pathogens from different kingdoms such as insects, oomycetes, bacteria, viruses and nematodes. To respond to these highly diversified groups of pathogens, NLRs possess flexible structures and varied modes of action that allow the recognition of a maximum number of effector proteins, one of the most powerful molecular arms of pathogens to cause infection.

Cooperativity between NLRs appears to be one way to maximize pathogen detection by using a limited set of resistance proteins (Williams et al., 2014). The NLR proteins RRS1 and RPS4 e.g. act together to mediate resistance to three different pathogens; *Pseudomonas syringae*, *Ralstonia solanacearum* and *Colletotrichum higginsianum.* Functional analyses have shown that RRS1 and RPS4 proteins interact in part via their TIR domains to form heterocomplexes that are essential for defense activation. Further, it has been shown that effector proteins PopP2 from *R. solanacearum* and AvrRPS4 from *P. syringae* exclusively interact with RRS1 and not with RPS4 showing that RRS1 acts as the effector sensor (Williams et al., 2014). P-loop motif of many NLRs is required for nucleotide binding and cell death triggering. Mutations in the P-loop motif in RPS4 and not in RRS1 impaired recognition of PopP2 and AvrRPS4 and cell death induction. Furthermore overexpression of RPS4 induced cell death showing that this NLR act as a cell death executor (Williams et al., 2014). Together these results present an example of cooperation between NLRs, in which one protein act as an “effector plataform” that facilities the perception to several pathogens by using a second protein that mediates the defense signaling.

The NLRs RGA4 and RGA5 from rice are another couple of resistance proteins working together to mediate pathogen recognition and the molecular bases of their mode of function started to be elucidate by Césari, et al., in 2014. It has been shown that RGA5 mediate the recognition to the sequence unrelated effector proteins AVR-Pia and AVR1-CO39 and the acts an effector sensor whereas RGA4 is a cell death executor as was also observed in RRS1 and RPS4 respectively. In addition, it was demonstrated that RGA5 acts as a RGA4 repressor and therefore in the presence of the AVR-Pia, RGA5 binds the effector protein relieving RGA4 which in turn trigger cell death.

In this chapter, we will provide details about the physical interaction between RGA5 and RGA4 domains and the nature of complexes before and after effector binding.

CONTEXT

Previously it was shown that RGA4 and RGA5 functionally interact to mediate effector recognition (Césari, et al. 2014). Furthermore, it was demonstrated that RGA5 and RGA4 form homo and hetero-complexes and that they physically interact through their CC domains. However it has not yet been elucidate whether domains in RGA4 and RGA5 others than CC also interact. In the absence of effector proteins, RGA5 interacts with RGA4 and repress its activity. Effector recognition by RGA5 releases RGA4 repression but it is unknown whether the active form of RGA4 still interacts with RGA5 or whether the interaction between RGA5 and the effector protein disrupts the RGA5-RGA4 hetero-complex.

To better understand the link between physical RGA4-RGA5 interaction and the activation of resistance after effector perception we evaluated the interaction of different RGA4 and RGA5 domains (Figure 31) as well as the interaction of RGA4 and RGA5 in the presence of AVR-Pia.

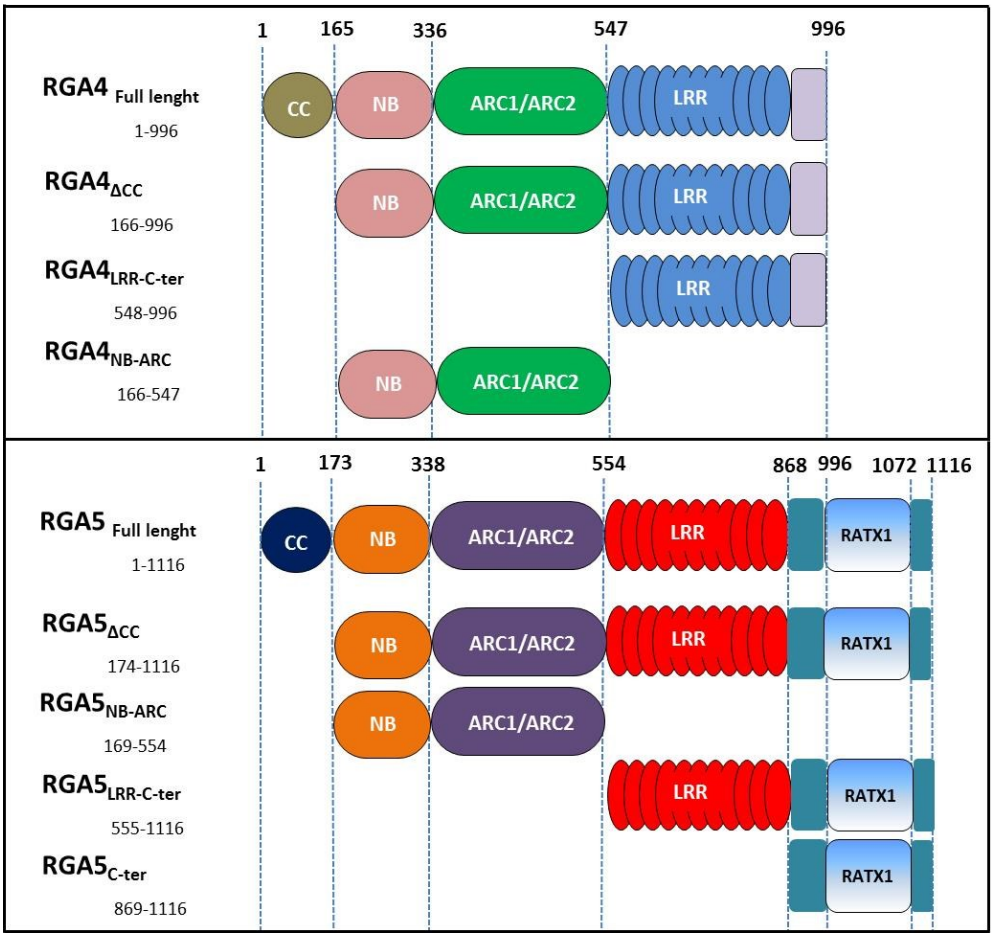


Figure 31. RGA4 and RGA5 domains used to analyze their interaction *in planta*.

RESULTS

RGA5 and RGA4 interact through multiple domains

Interaction between the CC domains of RGA5 and RGA4 was previously reported by Césari et al. (2014). To investigate whether other domains also interact, we generated mutants for RGA4 and RGA5 where the CC domains is deleted and fused them to HA and GFP or YFP tags: $RGA4_{\Delta CC}$-HA, $RGA4_{\Delta CC}$-GFP, HA-$RGA5_{\Delta CC}$ and YFP-$RGA5_{\Delta CC}$. Then, we tested their interaction by co-IP assays in comparison with the co-IP of the full length proteins RGA4-HA, RGA4-GFP, HA-RGA5 and YFP-RGA5. The Co-IP of $RGA4_{\Delta CC}$-HA and HA-$RGA5_{\Delta CC}$ by YFP was used as a negative control.

All proteins were properly expressed and proteins tagged with GFP or YFP were efficiently immunoprecipitated by anti-GFP agarose beads (Figure 32). Both $RGA4_{\Delta CC}$-HA and HA-$RGA5_{\Delta CC}$ were co-immunoprecipitated by YFP-$RGA5_{\Delta CC}$ and $RGA4_{\Delta CC}$-GFP respectively (Figure 2). The interaction of HA-$RGA5_{\Delta cc}$ with $RGA4_{\Delta CC}$-GFP was weaker than that of HA-RGA5 with $RGA4_{\Delta CC}$-GFP probably due to a reduced quantity of $RGA4_{\Delta CC}$-GFP in protein extracts (input Figure 32). Taken together, these results indicate that the CC domains of RGA4 and RGA5 are dispensable for association and that additional domains in these NLRs seem to interact to mediate hetero-complex formation.

The $RGA5_{C\text{-}ter}$ does not interact with the NB-ARC and LRR-C-ter domains of RGA4

To investigate what domains of RGA5 are involved in the interaction with RGA4, we analyzed the interaction of HA-$RGA5_{NB\text{-}ARC}$, HA-$RGA5_{LRR\text{-}C\text{-}ter}$ and HA-$RGA5_{C\text{-}ter}$ with $RGA4_{NB\text{-}ARC}$-GFP and $RGA4_{LRR\text{-}C\text{-}ter}$-GFP by co-IP assays in *N. benthamiana*. The interactions between YFP and RGA5 variants were used as negative controls.

All proteins were properly expressed and the proteins fused to YFP or GFP were efficiently immune-precipitated (Figure 33). HA-$RGA5_{NB\text{-}ARC}$ and HA-$RGA5_{LRR\text{-}C\text{-}ter}$ were well co-precipitated by $RGA4_{NB\text{-}ARC}$-GFP and $RGA4_{LRR\text{-}C\text{-}ter}$-GFP respectively, suggesting that the NB-ARC and LRR domains of RGA4 and RGA5 are involved in hetero-complex formation, in addition to the previously reported interaction between their CC (Césari, et al in 2014).

On the contrary, HA-RGA5$_{C\text{-}ter}$ was not co-precipitated by either RGA4$_{NB\text{-}ARC}$-GFP or RGA4$_{LRR\text{-}C\text{-}ter}$-GFP. Taken together these results suggest that all RGA5 domains but the C-ter participates in the interaction with RGA4. The interaction between RGA5$_{C\text{-}ter}$ and RGA4$_{CC}$ will have to be tested in the future.

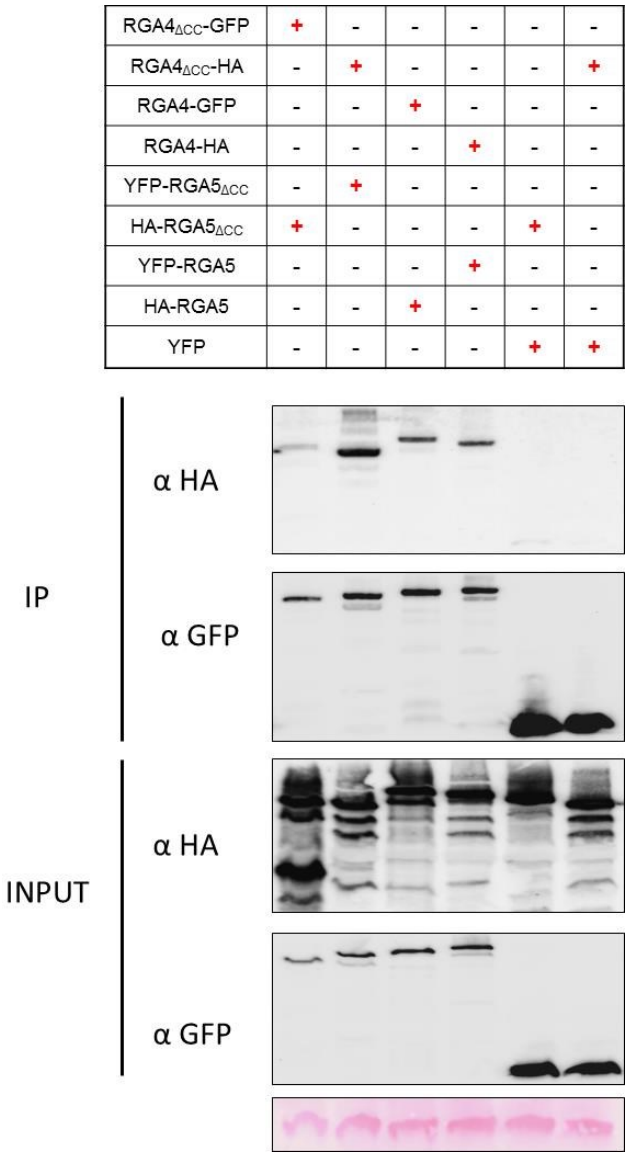


Figure 32. The CC domain of RGA4 and RGA5 is dispensable for their interaction. Full length and delta-CC of both RGA5 and RGA4 proteins tagged with HA, YFP or GFP were transiently expressed in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection

of immunoprecipitated RGA5 and RGA4 as well as YFP. Co-precipitation of RGA4-HA, $RGA4_{\Delta CC}$-HA or HA-RGA5 and HA-$RGA5_{\Delta CC}$ was analyzed using α-HA antibody.

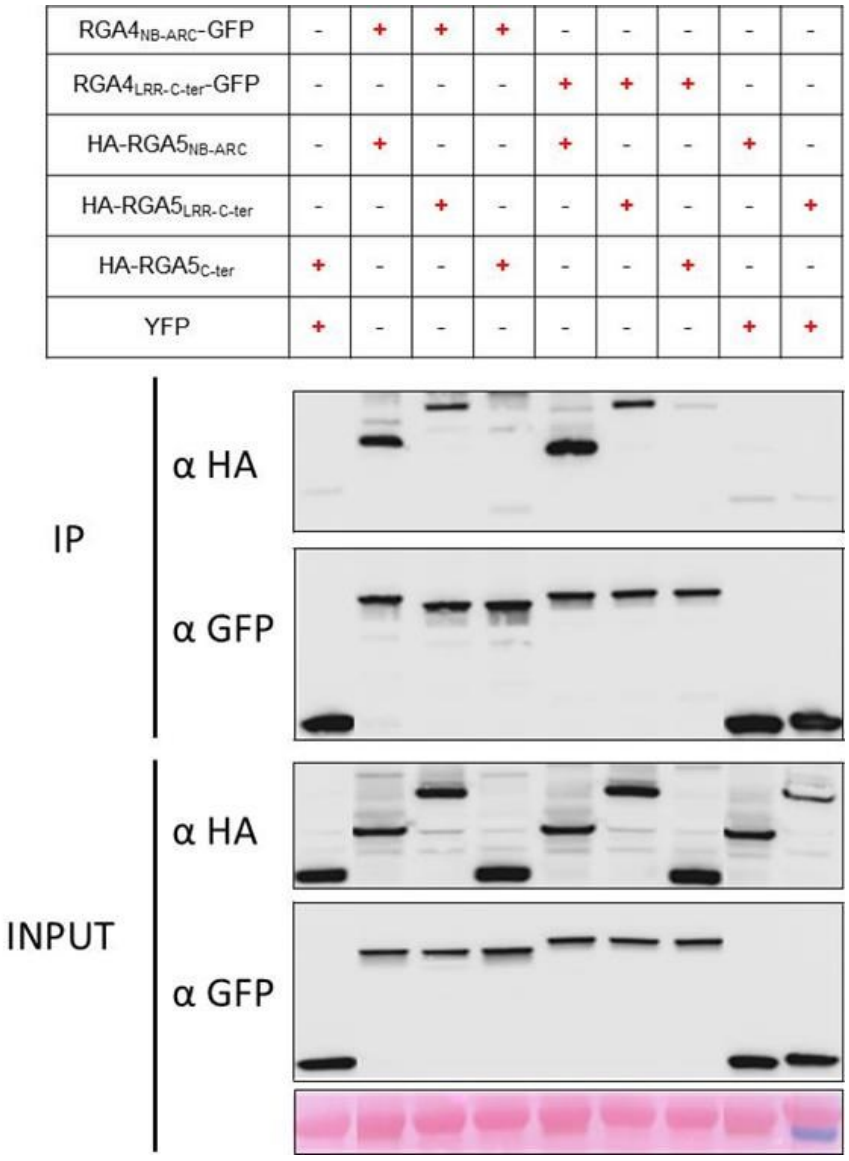


Figure 33. **RGA4 and RGA5 interact through their NB-ARC and LRR domains**. $RGA4_{NB-ARC}$-GFP and $RGA4_{LRR-Cter}$ were transiently expressed with HA-RGA5 variants in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection of immuneprecipitated $RGA4_{NB-ARC}$-GFP, $RGA4_{LRR-Cter}$ and YFP. Co-precipitated RGA5 variants were detected using α-HA antibody.

The interaction between RGA4 and RGA5 appears not to be disrupted by AVR-Pia

To investigate whether the interaction between RGA4 and RGA5 is disrupted by AVR-Pia, we co-expressed RGA5 and RGA4 with and without AVR-Pia. For this, we used the tagged proteins RGA4-GFP, YFP-RGA5, HA-RGA5, HA-AVR-Pia and YFP and we analyzed the complex formation by co-IP experiments.

All tagged proteins were well expressed and proteins fused to either YFP or GFP were in all cases efficiently immuneprecipitated. HA-RGA5 was well co-precipitated by RGA4-GFP in the presence and absence of HA-AVR-Pia (Figure 34) indicating that AVR-Pia does not abolish RGA4-RGA5 interaction. In addition, HA-AVR-Pia was not co-precipitated by RGA4-GFP suggesting that no stable trimeric RGA4-RGA5-AVR-Pia complex is formed. However, HA-AVR-Pia was co-precipitated by YFP-RGA5 (Figure 34).

We then used a non-tagged AVR-Pia protein to confirm that the interaction between YFP-RGA5 and RGA4-HA is not affected by this effector. RGA4-HA was co-precipitated by RGA5 in presence of AVR-Pia which supports the idea that AVR-Pia does not abolish the interaction between RGA5 and RGA4 (Figure 35). However, whether AVR-Pia reduces the amount of co-precipitated RGA4-HA cannot be determined from this experiment since the immunoprecipitation of RGA5 failed for the YFP-RGA5-RGA4-HA sample where AVR-Pia is absent.

Taken together these results suggest that RGA4 and RGA5 interaction is not abolished by AVR-Pia, however whether RGA5 and RGA4 form a stable three-protein-complex with AVR- remains to be elucidated.

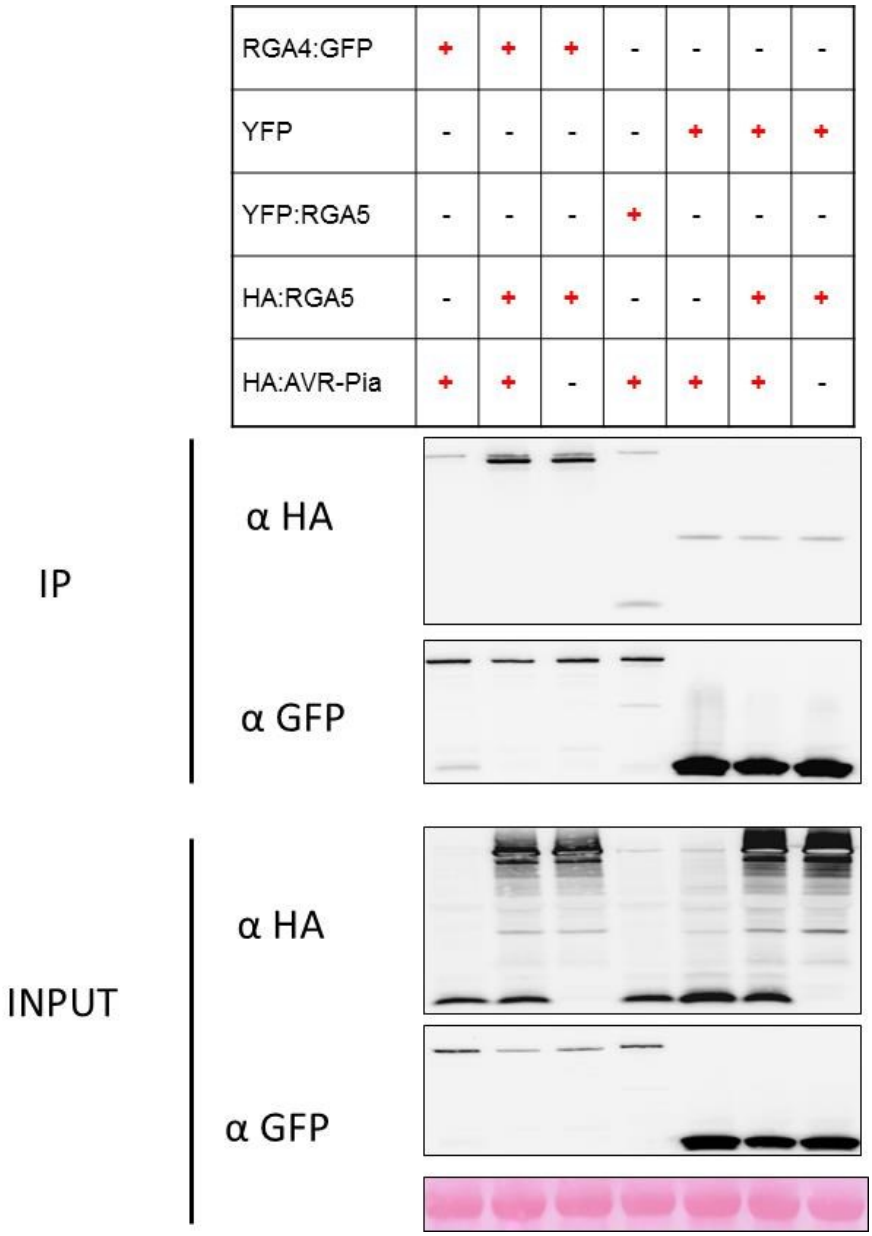


Figure 34. RGA4 and RGA5 interact in presence of HA- tagged AVR-Pia protein. RGA4-GFP was transiently expressed with HA-RGA5 and HA-AVR-Pia in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection of immunoprecipitated RGA4-GFP and YFP. Co-precipitation of HA-AVR-Pia and HA-RGA5 was verified by using α-HA antibody.

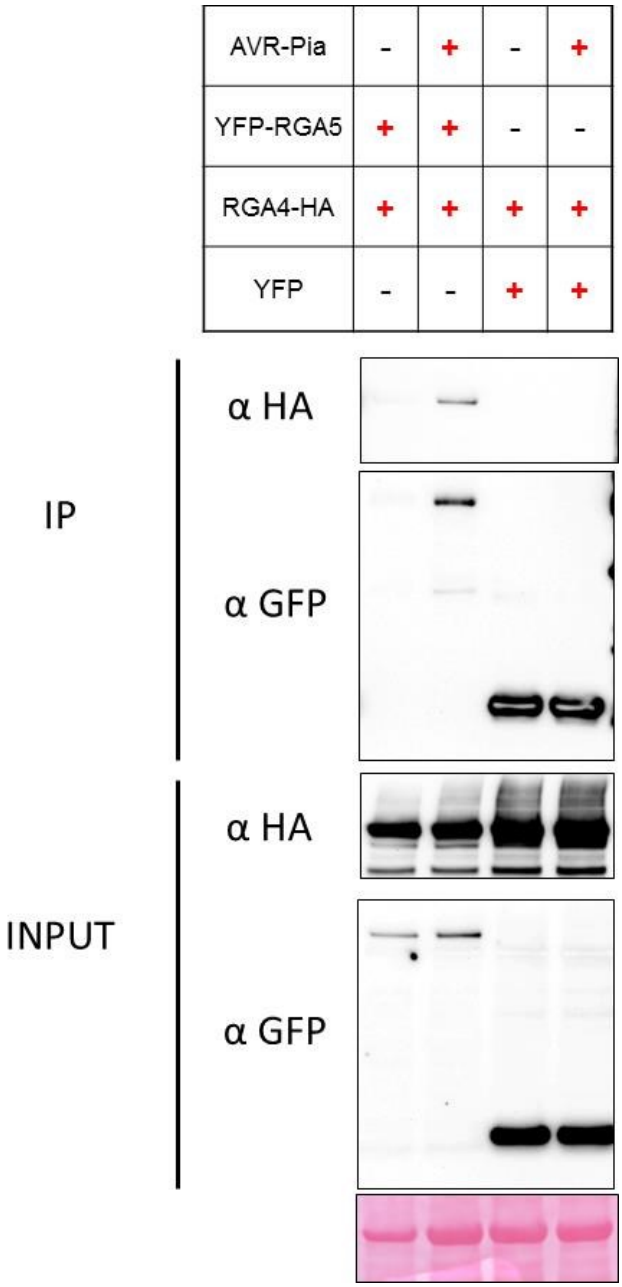


Figure 35. RGA4 and RGA5 interact in presence of the un-tagged AVR-Pia protein. YFP-RGA5 was transiently expressed with RGA4-HA and AVR-Pia in *N. benthamiana*. Protein extracts were analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Immunoprecipitation (IP) was conducted with anti-GFP beads (IP GFP) and analyzed by immunoblotting with α-GFP for the detection of immunoprecipitated YFP-RGA5 and YFP. Co-precipitation of RGA4-HA was verified by using α-HA antibody.

DISCUSSION

RGA5 and RGA4 present a robust interaction mediated by at least three different domains

In this work we investigated the molecular details of the interaction of RGA5 with RGA4 and we showed that $RGA5_{NB\text{-}ARC}$ and $RGA5_{LRR}$ domains interact with their corresponding domains $RGA4_{NB\text{-}ARC}$ and $RGA4_{LRR}$ domains. In addition we showed that $RGA5_{RATX1}$ does not interact with neither $RGA4_{NB\text{-}ARC}$ nor with $RGA4_{LRR}$. These data led us to the conclusion that RGA4 and RGA5 establish tight and multiple interactions mediated by three different domains: the CC domain, as shown by Césari et al 2014 and the NB-ARC and LRR as showed here. Only the $RGA5_{RATX1}$ domain appears not to participate directly in RGA4/RGA5 complex formation. However, since we have not yet characterized the interaction between $RGA5_{C\text{-}ter}$ and $RGA4_{CC}$ we cannot reject the possibility that $RGA5_{RATX1}$ somehow interacts with this domain.

$RGA5_{RATX1}$ is dispensable for repression of RGA4 by RGA5 and AVR-Pia association but is required to mediate AVR-Pia recognition. On the other hand, $RGA5_{RATX1}$ interacts with neither $RGA4_{NB\text{-}ARC}$ nor $RGA4_{LRR}$. Taken together these suggests that binding of AVR-Pia to the RATX1 domain and association to the NB-ARC and eventually the LRR domain induce the release of RGA4 repression, rather than the interaction of $RGA5_{RATX1}$ with $RGA4_{CC}$ or other RGA4 or RGA5 domains. Further confirmation and investigation of the relevance of AVR-Pia association with $RGA5_{NB\text{-}ARC}$ and $RGA5_{LRR}$ domains will therefore be of great importance to better understand the role of inter and intramolecular interactions of RGA4/RGA5 complex in effector recognition and activation of the resistance.

An RGA4/RGA5 complex seems to mediate effector depend cell death induction

Here we provide further evidence that the robust physical interaction between RGA4 and RGA5 is not abolished in the presence of AVR-Pia. Previously, Césari et al 2014 showed that - RGA5 represses RGA4-triggered cell death and that AVR-Pia relieves this repression. We show that YFP-RGA5 co-precipitates RGA4-HA in the presence of AVR-Pia, suggesting that the RGA4/RGA5 complex is not disrupted upon AVR-Pia recognition but that rather conformational changes and modifications of intermolecular interactions between RGA4 and RGA5 occur. In this way, the surfaces in RGA4 required to induce cell death may become

available to interact with other uncharacterized proteins or to form a signaling competent complex with RGA5.

In RRS1/RPS4 e.g., effector recognition by RRS1 leads to conformational changes in the RRS1/RPS4 complex that result in homotypic interactions of the RPS4 TIR domain required for cell death but not in disruption of the complex. In animals NAIP/NLRC4 NLR pairs form a wheel-shaped oligomer called the inflammasome after signal recognition and during defense signaling that contains one NAIP-ligand dimers and up to 11 NRC4 molecules (Tenthorey et al., 2014; Bentham et al., 2016). Formation of the oligomer is initiated by formation of a NAIP-ligand complex. Then, the NAIP-ligand dimer binds to NLRC4, which triggers association of additional NLRC4 molecules and formation of the inflammasome that transduces the immune signal. However in this case, NAIP and NLRC4 interact only after ligand perception suggesting that the mechanism to activate signaling markedly different between plant and animal NLR pairs. Further analyses are therefore required to better understand the nature of NLR heteropair complexes in plants during effector recognition and defense signalling.

Material and methods

The material and methods of this chapter are detailed in the annex.

SUMMARY POINTS (CHAPTER IV)

- ✓ CC domains are dispensable in the intermolecular interaction between RGA5-RGA4.
- ✓ CC, NB and LRR domains of RGA5 are involved in the intermolecular interaction with their corresponding domains in RGA4.
- ✓ RGA4 and RGA5 form hetero-complexes in presence or absence of AVR-Pia.

GENERAL DISCUSSION AND PERSPECTIVES

During their co-evolution, plants and pathogens employ a large repertoire of molecular arms to interact with each other. Pathogens deploy effector proteins that act inside or outside host cells to manipulate plant cellular processes, to promote pathogen growth and cause disease while plants rely on the continuous surveillance of the cell by using immune receptors to avoid disease. A major class of plant immune receptors are the NLR proteins that recognize cytoplasmic pathogen effectors and this recognition results in the activation of defense responses protecting plants from the disease.

In this work, we studied the molecular factors in both effector and NLR proteins that determine recognition specificity and the molecular bases of the activation of the NLR proteins after effector perception. In addition, we investigated the diversity of fungal effector proteins.

Fugal effectors that are highly variable in sequence may be related by structural homologies and common evolutionary origin

The huge effector complements of fungi are extremely diverse and mainly lineage or species specific suggesting that the effectors of different fungal linages evolve rapidly and independently from one another. This can, e.g., be nicely observed in the powdery mildew fungi where genotypes that infect barley and closely-related genotypes that attack wheat share about 500 candidate effectors while they have only few in common with more distantly related powdery mildew species that attack dicotyledonous plants (Pedersen et al., 2012; Spanu et al., 2010). As a consequence, the massive identification of fungal candidate effector proteins in the last decade has unraveled thousands of unique proteins of unknown functions and origin.

In this work, we determined by NMR the 3-dimensional structure of the sequence-unrelated *M. oryzae* effector proteins AVR-Pia and AVR1-CO39. This, in combination with bioinformatic searches and *in planta* expression analyses, led us to the identification of a family of effector proteins that we named MAX-effectors (**M**agnaporthe **A**vrs and To**x**B like). MAX-effectors are the first family of effectors with conserved structure and unrelated sequence identified in fungi. Their discovery is of great importance for effector biology since it suggests that

many of the huge number of apparently unrelated effectors in fungi may actually be structurally and phylogenetically related and gathered in structurally defined families.

Highly similar 3-dimensional structures have been also found in effector proteins with limited sequence homology in oomycetes (Boutemy et al. 2011; Win et al. 2012). The crystal structures of the effector proteins AVR3a11 and PexRD2 from *Phytophthora capsici* and *Phytophthora infestans* revealed that despite extremely weak sequence identity, below 20%, these two effectors displayed a conserved α-helical fold termed the "WY-domain" (Boutemy et al., 2011). Bioinformatic analyses suggest that the core α-helical fold actually occurs in 44% and 26% of annotated *Phytophthora* and *H. arabidopsis* RXLR effectors respectively either as a single domain or in tandem repeats (Boutemy et al., 2011).

The identification of structurally related effector families such as MAX-effectors and WY-RXLR-effector that are highly diversified provides a molecular framework to better understand the way effectors evolved and tolerate sequence hyper-variability. For example, we observed that the β-sandwich fold of the MAX-effectors allows insertion or deletion of amino acids in the loops or the exchange of surface exposed amino acids. A flexible structure like the MAX-effector β-sandwich fold therefore facilitates the generation of multiple effectors with different shapes and surface properties, molecular activities and host targets that nevertheless maintain a stable overall structure.

The identification of fungal candidate effectors relies mainly on bioinformatic sequence analyses and is not very precise because it is usually based on relatively broad criteria. Here we showed that a combination of structure-informed pattern-based searches led to the identification of huge numbers of MAX candidate effectors mainly in *M. oryzae* and *M. grisea.* In addition we showed that a large majority of MAX candidate effectors in *M. oryzae* are expressed specifically during early infection indicating that this group of effectors may be important during the biotrophic colonization of the plant host. This analysis demonstrates that structure determination and structure-informed bioinformatics analysis are extremely powerful approaches in the analysis of fungal effectors.

Interestingly, three of the best characterized MAX-effector proteins AVR-Pia, AVR1-CO39 and AVR-Pik in *M. oryzae* interact with a HMA/RATX1 domain integrated in an NLR protein;

Pik-1 in the case of AVR-Pik and RGA5 in the case of AVR-Pia and AVR1-CO39. This suggests that in addition to sharing a conserved structure, some MAX-effectors may also interact with highly similar domains in host proteins, despite the big differences in their surface properties. The structure of the AVR-Pik-HMA complex has been determined (Maqbool et al. 2015) and it is now a priority to generate also the structure of AVR-Pia and AVR1-CO39 in complex with the $RGA5_{RATX1}$ domain and compare whether the interaction between these avirulence proteins and HMA/RATX1 involves homologous β-strands or loops or whether the residues that determine the specificity and strength of the interaction possess similar properties.

61 HMA/RATX1-like proteins are present in the rice genome according to Panther database (www.pantherdb.org/panther/family.do?clsAccession=PTHR22814) (Cesari et al., 2013). In a mid-term, it would be very interesting to analyze e.g. by yeast two hybrid assays whether other MAX- effectors interact with some of these HMA/RATX1-like proteins in rice. These studies would provide an overview about the conservation of MAX-effector functions and targets. However, from present data it is already clear that MAX effectors target also other host proteins than HMA/RATX1 proteins, since AvrPiz-t interacts with E3 ubiquitin ligases. To get a broader picture of MAX effector functions, it will in the future be important to determine the host targets of a greater number of MAX effectors by approaches without *a priori* such as pull down or yeast two hybrid screens.

The two MAX effector proteins AVR-Pia and AVR-Pik have similar RATX1-binding surfaces

In this thesis, we deciphered in detail the interaction between AVR-Pia and the RATX1 domain of RGA5. We identified the RATX1-binding surface in AVR-Pia and found that this surface resembles to the surface of AVR-Pik that binds the HMA domain of Pik-1. The HMA and the RATX1-domains have similar sequences (55% identity) and according to homology-based structure modeling have very similar structures. However, they seem to have significant differences in the HMA/RATX1-binding surfaces of AVR-Pia and AVR-Pik. Surface properties such as charge distribution and hydrophobicity are different between both MAX effectors and, in particular, an N-terminal extension of AVR-Pik absent from AVR-Pia is crucial for interaction with $Pik\text{-}1_{HMA}$. One interesting question now is how AVR1-CO39, a sequence-unrelated but structurally similar effector to AVR-Pia and AVR-Pik, interacts with

the RATX1 domain? It is e.g. possible that the binding surface in AVR1-CO39 involves like in AVR-Pia and AVR-Pik the β-strands 2 and 3 and loop 2. However, these surfaces have very different properties in AVR-Pia and AVR1-CO39 which makes it difficult to understand how they could bind the same target sequence. In addition, AVR1-CO39 has a C-terminal extension that could be involved in RATX1-binding in a similar way that the N-terminal extension of AVR-Pik that binds the HMA domain. Hence, the question with which surface AVR1-CO39 binds the RATX1 is clearly open.

Another important question is: which are the surfaces of the RATX1 that interact with AVR1-CO39 or AVR-Pia. One possibility is that both AVR1-CO39 and AVR-Pia interact with the same binding surface in the RATX1/HMA structure. However, if this is the case it will be particularly interesting to learn how this is accomplished since both effectors differ strongly in surface properties and shapes. Alternatively, they could interact with different surfaces in the RATX1 domain that may in addition be different from the homologues surface in Pik-1_{HMA} bound by AVR-Pik. Docking models for the AVR-Pia/RATX1 complex support this hypothesis since no model of this complex resembles the structure of the Avr-Pik/Pik_{HMA} complex (chapter II) .

Finally, it would be interesting to identify the virulence targets of AVR1-CO39, AVR-Pia and AVR-Pik, and determine whether they are in all cases small HMA proteins and to investigate by which molecular mechanism these three effectors act on their corresponding virulence targets. Work in the group of R. Terauchi (IBRC, Iwate, Japan) indicates that AVR-Pik targets a clade of small HMA proteins similar to the Pik_{HMA} and $RGA5_{RATX1}$ domains and establishes with them the same type of interaction they set with Pik_{HMA}. These interactions stabilize the small HMA proteins which contribute to increased susceptibility and reduction of ROS production (personal communication). Whether AVR-Pia and AVR1-CO39 target the same clade of sHMAs and whether they target them by the same type of mechanism remains to be determined.

Additional NLR domains cooperate with integrated domains in effector binding and recognition

In this work, we found that AVR-Pia mutants that lost the capacity to bind the $RGA5_{RATX1}$ domain are able to fully interact with $RGA5_{\Delta RATX1.}$, This indicates that other domains of RGA5 than $RGA5_{RATX1}$ interact with AVR-Pia and this probably through AVR-Pia surfaces that are different from the RATX1-binding surface. Indeed further analyses suggested that these additional interactions involve the $RGA5_{NB\text{-}ARC}$ and $RGA5_{LRR}$ domains and that similar interactions occur with AVR1-CO39. The recognition of different effector surfaces by independent binding sites in NLRs, including interactions with IDs, is expected to increase specificity and robustness in effector recognition since reduction in the affinity to one binding side has minor impact when a second binding site is present. Such coordinated binding to several binding sites, including the integrated decoy could constitutes an efficient strategy to avoid the loss of recognition by rapid mutations in the effector and be one of the key advantages of NLR-IDs over a situation where NLRs and integrated decoys are separated.

To validate this model it will be important to demonstrate in the future that the interaction between the effector and the NLR domains other than the integrated domain are required for effector recognition. For this, we will further characterize the interaction of AVR-Pia and AVR1-CO39 with the NB-ARC and LRR domains. Then we are planning to identify AVR-Pia mutants that do no longer associate with either or both of these domains and to analyze the impact of the loss of this association in effector binding and avirulence activity.

Integrated domains are crucial for effector recognition and promise to be useful guides for the identification of plant effector targets.

In the present study we found that binding of AVR-Pia to the RATX1 domain is required for effector recognition. This is an important support for the integrated decoy model that states that **NLR integrated domains (NLR-ID**) serve as effector detector modules. Similarly compelling evidence for this model has for the moment only been provided by the investigation of AVR-Pik/Pik-1, PopP2/RRS1 or AvrRPS4/RRS1. In AVR-Pik/Pik-1 and AvrRPS4/RRS1 effector-binding to the decoy domain have been identified as in AVR-Pia/RGA5 as key events for NLR activation. In PopP2/RRS1, the NLR is rather activated by an effector-mediated acetylation of the integrated decoy domain. This highlights that

integrated domains, very likely as classical guardees, decoys or co-factors, can function in effector recognition by different manners: they can be involved in complex formation or acquire post-translational modifications that are sensed by the associated NLR.

The systematic analysis of plant NLRs has shown that the integration of uncommon domains is frequent; it occurs at a rate of approximately 5%, and is widespread since NLR-IDs have been found in all land plants from mosses to angiosperms (Césari, et al. 2014; Kroj et al. 2016; Sarris et al. 2016). In addition, the diversity of integrated domains is very high. For instance, Sarris et al. 2016 identified about 265 different integrated domains in 750 NLR proteins carrying integrated domains (Sarris et al., 2016). The most highly represented domains are protein kinases, DNA-binding and protein-protein interaction domains (Sarris et al., 2016). Interestingly, many of the integrated domains fused to plant NLRs have been found to interact with effectors in targeted studies or effector interactome screens (Sarris et al., 2016). This suggests that effector recognition is a general feature of integrated domains similar to what has been demonstrated experimentally for RGA5, Pik-1 and RRS1. Furthermore, many of the integrated domains have not yet been associated with plant diseases or immunity and represent an outstanding guide for the identification of novel effector targets and immune components (Kroj et al., 2016; Sarris et al., 2016; Ellis, 2016; Malik and Van der Hoorn, 2016)

Perspectives for plant immune receptor design exploiting the modular structure of NLR-IDs

The uncovered role of integrated domains in pathogen perception as well as the modular structure of NLR-ID has raised the possibility that the recognition specificity of this group of NLR receptors could be modified by simply exchanging the integrated domain. For instance it could be conceivable to plug cellular hubs of immune signaling that are targeted by multiple effectors into NLR-IDs to generate versatile large spectrum immune receptors. However, our study suggests that the design of NLR-IDs with completely novel specificities will not be as simple as that. Indeed, we show that the RATX1 domain as well as AVR-Pia interact with additional RGA5 domains. Indeed, we think that these interactions are required for RGA5 function, and, in particular, for proper RGA4 de-repression in the presence of effect Therefore, we believe that these additional interactions will have to be taken into consideration when integrating novel decoy domains. The work on AVR-Pik/Pik-1 suggests

that NLR-IDs with extended recognition specificities that recognize additional alleles of the same effector may be relatively easy to generate by changing single residues in the integrated domains.

Another first step can be the creation of RGA5 variants where residues that are crucial for AVR-Pik recognition have been introduced into the RATX1 domain to create an RGA5 variant that recognizes AVR-Pik. However, according to our model such variants would require probably additional modifications outside the RATX1 domain to interact correctly with the effector and be completely functional. However, a much more detailed understanding of effector-NLR-ID interactions and intramolecular interaction in NLR-IDs will be required before novel specificities can be designed by integration of completely new decoy domains

NLR pairs interact through multiple domains

Another open and challenging question is how RGA4 and RGA5 and more generally other paired NLRs interact to mediate effector recognition. In the TNL pair RRS1/RPS4, the important role of TIR domain homo and heterotypic interactions in activation and repression has been documented (Williams et al., 2014). In RGA4/RGA5, previously shown CC-interactions could play similar roles (Césari, et al. 2014). In the present study, we provide evidence that additional domains of RGA4 and RGA5 contribute to their hetero-association. Indeed, the NB-ARC and LRR domains of RGA4 seem to interact with the NB-ARC and LRR domains in RGA5 respectively. To evaluate the importance of these interactions for the function of the hetero-complex, it will be important to further characterize these interactions and to create specific mutations within RGA4 and/or RGA5 CC, NB-ARC or LRR domains that prevent their hetero-association and analyze whether these mutations also affect effector recognition and defense signalling.

The RGA4/RGA5 hetero-complex seems not to be disrupted by the presence of the effector

It was previously demonstrated that RGA4 and RGA5 interact in the absence of recognized AVRs (Césari, et al. 2014). In this study, we provide preliminary results suggesting that the presence of the effector does not disrupt the RGA4/RGA5 hetero-complex. However, additional experiments are required to confirm this hypothesis and it remains to be determined whether the functional complex is indeed a tripartite complex constituted by

AVR-Pia/RGA4/RGA5.

One perspective is to perform additional co-IP experiments where AVR-Pia is immunoprecipitated and the co-precipitation of RGA4 and RGA5 is evaluated. In addition, it should be important to confirm whether RGA4 co-precipitates AVR-Pia in presence of RGA5 in *in planta* experiments. Alternatively, BiFC analyses and Fret-Flim experiments can monitor the *in planta* interaction of RGA4 and RGA5 in the absence and the presence of AVR-Pia.

Defense signalling mediated by plant NLR hetero-complex is poorly understood. In the case of RGA4/RGA5 it was shown that RGA4 is able to induce cell death either in absence of RGA5 or after the recognition of AVR-Pia through RGA5 (Césari, et al. 2014). However, how RGA4 mediates defense signaling is still an open question and the downstream signaling partners are not known. In animals it has been shown that a functional hetero-complex of NLRs functioning in pairs actually involves a multiprotein complex called inflammasome. The last, has a wheel-like structure that creates one oligomerization surface required to defense signaling (Hu et al., 2015). For example, in mouse and humans, the NLRs called NAIPs (Neuronal Apoptosis Inhibitor Protein) interact with NLRC4 (NLR family CARD domain-containing protein 4) to mediate resistance to bacterial pathogens. In mice, NAIP2 directly recognizes the rod protein PrgJ, a principal component of the type III secretion systems in bacteria. This recognition induces NAIP2 interaction with NLRC4 that result in the oligomerization of NLRC4 in complex with NAIP and PrgJ. The structure of the PrgJ-NAIP2-NLRC4 complex has been determined by cryogenic electron-microscopy and have revealed that this multiprotein complex is constitute by one molecule of PrgJ, one molecule of NAIP2 and 10-12 molecules of the NLRC4 (Hu et al., 2015). To investigate whether plant hetero-complex can form similar structures, it would be interesting to elucidate the stoichiometry of the active RGA4/RGA5 complex.

CONCLUSION

In this work, we identified the first family of structure-related fungal effector proteins and established new lights for a better understanding of effector evolution and diversification in fungi. Furthermore, we characterized the AVR-Pia surface that is recognized by the integrated RATX1 decoy domain of RGA5. In this way, we expanded the knowledge about

how effector proteins, in particular MAX-effectors, bind integrated domains. Next we showed that RATX1 domain is required for effector recognition and that the NB-ARC and LRR domains of RGA5 also associate with effector proteins. Based on these results we proposed a cooperative model in which the recognition of the effector proteins and the activation of defense signalling depend on the association/binding of multiples NLR domains with the effector protein. We believe that this mode of recognition constitute a good strategy to maximize the recognition of rapidly evolving effector proteins and to avoid the segregation between the NLR and the integrated decoy. Finally we reveled that RGA5 and RGA4 interact through their CC, NB and LRR domains and that some of these interaction seem not to be disrupted by the effector since we observed hetero-complex formation in the presence of AVR-Pia. However, additional experiments are required to confirm these preliminary results. Collectively, these results provide important insights about the recognition of effector proteins mediated by NLR integrated decoys. Nevertheless, we need to put much more effort in the study of the molecular bases of this recognition before the engineered of NLRs with integrated decoys could be considered.

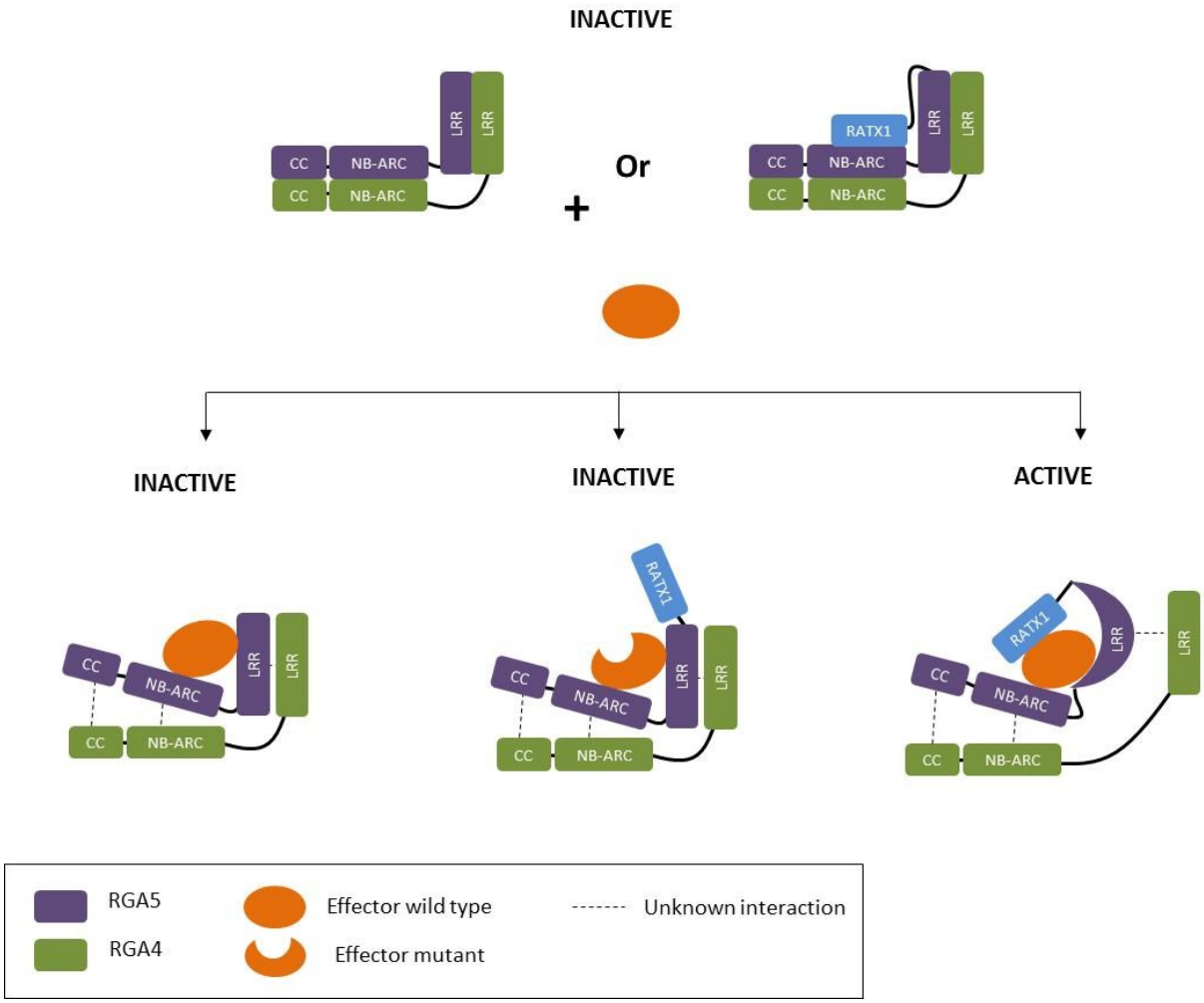
Model of effector recognition and resistance activation by the RGA4/RGA5receptor complex

Based on the work presented in **chapter II**, we proposed a model to describe the recognition of AVR-Pia by RGA5 and RGA4 on the molecular level. The basis of the model is that the AVR-Pia-$RGA5_{RATX1}$ interaction is absolutely required for effector recognition. In addition, it takes into consideration that the strength of AVR-Pia-$RGA5_{RATX1}$ binding can be rather low and that AVR-Pia associates with additional RGA5 domains, outside the RATX1 domain. We hypothesized that those interactions are important to occur simultaneously with binding of AVR-Pia to the RATX1 domain to increase overall binding strength and affinity of the effector to RGA5 .

In **chapter III** we describe further dissection of the AVR-Pia/RGA5 interaction and present evidence that the additional RGA5 domains mediating AVR-Pia-binding are the NB-ARC and the LRR domains. In addition, we show that also AVR1-CO39 interacts with other domains of RGA5 than the RATX1 domain and that again the NB-ARC domain is involved. When analyzing for intramolecular interactions of the RATX1 domain, we found that RATX1 forms homo-dimers and associates in addition with $RGA5_{NB-ARC}$. Interestingly, both interactions are perturbed by AVR-Pia.

Since RGA5 mainly acts through RGA4 the molecular bases of the interaction between RGA4 and RGA5 were also investigated in **chapter IV**. This provided evidence that in addition to previously known interactions between the CC domains, the NB-ARC and the LRR domains are engaged in heterotypic interactions. Interestingly, the RATX1 domain seems not to interact with RGA4, neither with $RGA4_{NB-ARC}$ nor with $RGA4_{LRR-C-ter}$ suggesting that the C-terminus of $RGA5_{C-ter}$ is not directly involved in RGA4 and RGA5 complex formation.

Altogether these results led to a refined model for effector recognition by RGA5 and RGA4.



Model for effector recognition by the RGA4/RGA5 receptor complex

PROJET DE THESE ET DISCUSSION GENERALE

INTRODUCTION

Dans l'agriculture moderne, le développement de variétés exprimant une résistance à la fois durable et efficace contre une large gamme d'agents pathogènes est essentiel à la productivité et la durabilité des systèmes de production. L'une des principales stratégies variétales pour la lutte contre les maladies consiste au déploiement de gènes de résistance (gènes *R*). Ces gènes sont spécifiquement activés par des protéines appelées « effecteurs », codées par des gènes dits « de virulence » chez les agents pathogènes. Cependant, l'efficacité des gènes *R* est souvent de courte durée du fait de la grande diversité génétique dont disposent les populations d'agents pathogènes, et de leur capacité à évoluer de manière extrêmement rapide. Ainsi, le déploiement seul de gènes *R* en monocultures mène bien souvent à la sélection des variants (d'agents pathogènes) dont l'effecteur concerné a été muté ou désactivé, induisant alors une réaction compatible avec la plante pour le déclenchement de la maladie.

Il est donc essentiel de développer des techniques de sélection génétique et des stratégies de déploiement innovantes, afin d'approvisionner les systèmes agricoles avec des gènes de résistance efficaces et durables. Une approche prometteuse consiste à générer de manière synthétique des gènes de résistance spécifiques à de nouveaux effecteurs, avec un large spectre d'action (c'est-à-dire capables de cibler plusieurs agents pathogènes), et ce à un rythme au moins aussi rapide que celui auquel évoluent les agents pathogènes. Pour cela, il est nécessaire d'explorer la riche diversité des gènes de résistance des plantes, et d'approfondir la caractérisation fonctionnelle de leur mode d'action afin de comprendre comment les produits de ces gènes perçoivent l'attaque d'un agent pathogène et déclenchent les réactions immunitaires de la plante.

OBJECTIF GENERAL ET QUESTIONS DE RECHERCHE

Objectif général

Le principal objectif de cette thèse est de contribuer à une meilleure compréhension des déterminants moléculaires de l'immunité des plantes. Plus précisément, nous avons cherché à élucider les bases moléculaires de la reconnaissance des effecteurs de champignons phytopathogènes par les récepteurs de type NLR des plantes, et le lien entre cette reconnaissance et la réaction immunitaire. Pour cela, nous avons caractérisé la structure des effecteurs protéiques AVR1-CO39 et AVR-Pia du champignon *Magnaporthe oryzae*, et étudié la diversité et l'évolution des effecteurs fongiques définis par leur structure chez les champignons ascomycètes. Nous avons également étudié les déterminants moléculaires permettant la reconnaissance de AVR-Pia par les récepteurs immunitaires RGA4/RGA5. Enfin, nous avons analysé la composition et la formation du complexe de récepteurs RGA4/RGA5.

Puisque le pathosystème Riz – *M. oryzae* est un système modèle pour les interactions plantes – champignons, les résultats générés dans le cadre de cette thèse viennent plus généralement alimenter nos connaissances sur les maladies végétales causées par les champignons.

Questions de recherche, approches et résultats principaux

Le projet de thèse s'est penché sur les questions suivantes.

1. **Est-ce que la structure tridimensionnelle des effecteurs est un facteur clé de leur fonction, leur diversité et leur évolution ?**

Pour répondre à cette question nous avons tout d'abord déterminé la structure tridimensionnelle des effecteurs AVR1-CO39 et AVR-Pia par résonnance magnétique nucléaire (RMN). Puisque ces deux effecteurs ont des structures similaires, qu'ils partagent, en outre, avec deux autres effecteurs protéiques : respectivement AvrPiz-t et ToxB chez *M. oryzae* et *Pyrenophora tritici-repentis*, nous avons identifié une famille d'effecteurs de séquences différentes mais qui possèdent une structure conservée. Cette famille a été appelée MAX-effectors (***M****agnaporthe* **A**vrs and To**x**B like). Leur analyse bioinformatique a montré que les effecteurs MAX sont distribués dans de nombreux taxons phylogénétiques,

très répandus et très diversifiés, et ce particulièrement au sein de l'ordre Magnaporthales (**Chapitre I**).

2. Comment AVR-Pia se lie au domaine RATX1 de RGA5?

Précédemment, notre groupe de recherche a démontré que l'effecteur AVR-Pia de *M. oryzae* interagit directement avec le domaine RATX1, à l'extrémité C-terminale du récepteur RGA5. Il a également été démontré que AVR-Pia-H3, un allèle naturel d'AVR-Pia, avait perdu cette capacité d'interaction avec le domaine RATX1, et ainsi n'activait pas la résistance des plantes de riz possédant le complexe de récepteurs RGA4/RGA5.

Dans la présente étude, nous avons approfondi la caractérisation de l'interaction AVR-Pia-$RGA5_{RATX1}$. Nous avons pour cela comparé les structures RMN de AVR-Pia et AVR-Pia-H3, et identifié une surface d'interaction candidate par titrage RMN. Nous avons alors caractérisé la formation *in vitro* du complexe AVR-Pia-$RGA5_{RATX1}$ par titration calorimétrique isotherme (isothermal titration calorimetry, ITC), et validé la surface d'interaction candidate ainsi que la présence de deux résidus clés par des techniques de double hybride (yeast two hybrid) et de co-immunoprécipitation de complexes protéiques (**Chapitre II**).

3. Quel est le rôle de la liaison AVR-Pia-$RGA5_{RATX1}$ dans la reconnaissance d'AVR-Pia ?

Pour déterminer le rôle de la liaison AVR-Pia-$RGA5_{RATX1}$ dans la reconnaissance et la fonction d'avirulence de AVR-Pia, des effecteurs AVR-Pia mutants ont été analysé dans des isolats transgéniques de *M. oryzae* et par leur expression transitoire chez *Nicotiana benthamiana* lors de l'induction de la mort cellulaire programmée. Ces expériences ont montré que la liaison AVR-Pia-$RGA5_{RATX1}$ est nécessaire pour le déclenchement de la résistance, mais que les mutants pour lesquels la liaison AVR-Pia-$RGA5_{RATX1}$ n'est que partielle sont également reconnus (**Chapitre II**).

4. Quel sont les rôles des autres domaines de RGA5 dans la reconnaissance de AVR-Pia ?

Afin de mieux comprendre la reconnaissance d'AVR-Pia par RGA5, nous avons caractérisé individuellement les interactions de AVR-Pia avec les différents domaines de RGA5 (CC, NB-ARC, LRR et RATX1) par des techniques de double hybride et co-immunoprécipitation de complexes protéiques. Ceci a révélé l'existence d'une interaction entre AVR-Pia et les

domaines NB-ARC et LRR de RGA5, en plus de celle précédemment décrite avec RATX1 (**Chapitre III**).

5. Comment se forme le complexe de récepteurs RGA4/RGA5 ?

RGA4 et RGA5 interagissent physiquement et fonctionnellement pour reconnaître AVR-Pia. La formation d'homo- et d'hétéro-complexes de leur domaines CC a été rapportée (Césari et al. 2014). De plus, il a été montré que RGA4 active la mort cellulaire programmée, alors que RGA5 agit comme un inhibiteur de RGA4 ainsi comme un récepteur de AVR-Pia.

Dans cette étude, pour mieux caractériser les interactions entre RGA4 et RGA5, nous avons évalué l'interaction physique entre ces deux récepteurs de type NLR en présence de l'effecteur sauvage AVR-Pia$_{WT}$. Ensuite, nous avons testé la formation de l'hétéro-complexe RGA4$_{\Delta CC}$ et RGA5$_{\Delta CC}$, et son association avec les domaines individuels de RGA4 et RGA5 (NB, CC, NB-LRR) par des techniques co-immunoprécipitation. Ces expériences préliminaires ont mis en évidence une interaction entre les trois diffèrent domaines de RGA4 et RGA5 et il a été observé que la association du complexe RGA4/RGA5 ne est pas perturbé par l'effecteur AVR-Pia (**Chapitre IV**).

DISCUSSION GENERALE ET PERSPECTIVES

Dans ces travaux de thèse, j'ai tout abord étudié la diversité des effecteurs fongiques chez les Ascomycètes. J'ai ensuite analysé les déterminants moléculaires de la spécificité de reconnaissance d'effecteurs fongiques par des récepteurs intracellulaires du système immunitaire des plantes. Enfin, j'ai approfondi l'étude de l'activation des récepteurs immunitaires intracellulaires des plantes qui agissent en paires pour induire la résistance.

Des effecteurs fongiques de séquences très variables peuvent présenter des homologies structurales et posséder une origine évolutive commune.

Les champignons possèdent un répertoire d'effecteurs très large et diversifié. De plus, ces effecteurs sont spécifiques d'une espèce ou d'une lignée, ce qui suggère que les effecteurs appartenant à différents lignages fongiques évoluent rapidement et indépendamment les uns des autres. Par exemple, dans le cas de l'oïdium, les génotypes infectant l'orge sont

étroitement liés à ceux infectant le blé et partagent avec ces derniers environ 500 effecteurs candidats, alors qu'ils ont très peu d'effecteurs en commun avec les génotypes d'oïdium qui infectent les plantes dicotylédones (Pedersen et al., 2012; Spanu et al., 2010). En conséquence, l'identification récente et massive de nouveaux effecteurs fongiques a mis en évidence des milliers de protéines dont les fonctions et les origines sont largement méconnues.

Dans ce travail, nous avons déterminé par RMN la structure tridimensionnelle des effecteurs AVR-Pia et AVR1-CO39 chez *Magnaporthe oryzae*, le champignon pathogène responsable de la pyriculariose du riz. Cette analyse, combinée à des recherches bioinformatiques et des analyses d'expression chez la plante, nous a conduit à l'identification d'une famille de protéines effectrices que nous avons nommée « effecteurs MAX » (**M**agnaporthe **A**vrs and **T**oxB like). Chez les champignons, les effecteurs MAX constituent la première famille identifiée d'effecteurs comportant une structure conservée malgré des séquences en acides aminés très différentes. Cette découverte est d'une grande importance pour l'étude de la biologie des effecteurs car elle suggère qu'un grand nombre d'effecteurs, apparemment sans aucun lien entre eux, peuvent en réalité être liés au niveau structural, être rassemblés en familles et partager une origine phylogénétique commune. En effet, chez les oomycètes, une famille d'effecteurs possédant des structures similaires a également été trouvée (Boutemy et al 2011; Win et al 2012). C'est notamment le cas des effecteurs AVR3a11 et PexRD2 de *Phytophthora capsici* et *Phytophthora infestans*, respectivement. En dépit d'une identité de séquence inférieure à 20%, ces deux effecteurs possèdent dans leurs structures tridimensionnelles une hélice α conservée appelée "WY-domain" (Boutemy et al., 2011). Des analyses bioinformatiques complémentaires ont suggéré que ce WY-domain était présent dans les 44% des effecteurs décrits chez *P. infestans* (Boutemy et al., 2011).

L'identification de familles d'effecteurs structurellement liés, telles que celles des effecteurs MAX et WY-RXLR, fournit un cadre moléculaire permettant de mieux comprendre l'évolution des effecteurs vers une meilleure tolérance de l'hyper-variabilité des séquences en acides aminés. Nous avons par exemple observé que les structures en feuillets β-sandwich des effecteurs MAX permettaient l'insertion ou la délétion d'acides aminés dans les boucles, ou la substitution d'acides aminés en surface. Une structure souple des effecteurs MAX, comme celle en feuillet β-sandwich, facilite donc la génération d'effecteurs avec des surfaces de

formes et de propriétés différentes, tout en préservant leur activité moléculaire, leurs cibles et une structure globale stable.

Jusqu'à présent, l'identification d'effecteurs fongiques reposait principalement sur l'analyse bioinformatique de séquences et pouvait manquer de précision car généralement ces analyses sont basées sur des critères relativement larges. Ici, nous avons montré que la recherche d'effecteurs candidats basée sur leurs structures peut conduire à une identification très précise d'un très grand nombre d'effecteurs. En outre, nous avons montré chez *M. oryzae* que la grande majorité des effecteurs candidats MAX sont exprimés spécifiquement au cours de l'infection précoce, suggérant que ce groupe d'effecteurs a un rôle important dans la colonisation de l'hôte pendant les premières étapes de l'infection. L'analyse fonctionnelle des effecteurs candidats MAX, par des approches visant notamment à identifier leurs partenaires protéiques dans les cellules végétales, ou à déterminer leur localisation, permettront de mieux comprendre le ou les modes d'action de cette famille d'effecteurs.

De manière intéressante, AVR-Pia, AVR1-CO39 et AVR-Pik, les effecteurs MAX les mieux caractérisés, interagissent avec un domaine appelé RATX1/HMA, intégré dans les récepteurs immunitaires Pik-1 pour AVR-Pik, et dans RGA5 dans le cas d'AVR-Pia et AVR1-CO39. Ceci suggère qu'en plus de posséder une structure conservée, certains effecteurs MAX peuvent également interagir avec des domaines très similaires qui se trouvent dans les protéines de l'hôte, malgré les grandes différences dans les propriétés de surface de ces effecteurs. La structure du complexe AVR-Pik-HMA a été déterminée récemment par Maqbool et al. (2015). Il serait donc également intéressant de générer la structure du complexe formé par AVR-Pia ou AVR1-CO39 avec le domaine RATX1 afin d'évaluer si ces différentes interactions AVR-RATX1/HMA font intervenir des sites d'interaction similaires, ou si la spécificité et la force de l'interaction sont déterminées par des interfaces ou des résidus différents.

Les effecteurs MAX AVR-Pia et AVR-Pik ont des surfaces de liaison au domaine RATX1/HMA très similaires

Dans cette thèse, nous avons disséqué en détails la surface d'AVR-Pia qui se lie au domaine RATX1, intégré dans le récepteur immunitaire RGA5. Nous avons également souligné que cette surface est similaire à la surface de liaison d'AVR-Pik avec le domaine HMA, intégré

dans le récepteur immunitaire Pik-1. Curieusement, bien que 34 des 74 résidus qui constituent les domaines HMA et RATX1 diffèrent, ces domaines sont très similaires. De plus, les propriétés de la surface, telles que la distribution des charges ou l'hydrophobie, sont différentes entre les deux effecteurs MAX. En particulier, une extension N-terminale d'AVR-Pik, cruciale pour l'interaction avec Pik-1_{HMA}, est absente chez AVR-Pia. Il serait intéressant de découvrir comment AVR1-CO39, un effecteur dont la séquence en acides aminés est très différente de celles d'AVR-Pia et d'AVR-Pik, mais qui partage une structure similaire, peut se lier au domaine RATX1. Il est possible qu'à la différence d'AVR-Pia et d'AVR-Pik, la surface de liaison d'AVR1-CO39 implique les feuilles β 2 et 3 ainsi que la boucle 2. Néanmoins, il est également possible que l'extension C-terminale d'AVR1-CO39, qui ressemble à l'extension N-terminale d'AVR-Pik, soit impliquée dans l'interaction avec RATX1.

D'autre part, nous avons relativement peu d'indications concernant la surface de RATX1 qui interagit avec AVR1-CO39 ou AVR-Pia. Une possibilité est que ces deux effecteurs interagissent avec la même surface de liaison de RATX1. Toutefois, si tel est le cas, il sera particulièrement intéressant d'en expliquer les raisons, puisque les deux effecteurs présentent des propriétés de surface très différentes. Alternativement, ils pourraient interagir avec des surfaces différentes du domaine RATX1, surfaces qui pourraient également différer de la surface de son homologue HMA (qui se lie à AVR-Pik). La modélisation structurale du complexe AVR-Pia/RATX1 soutient cette dernière hypothèse car aucun modèle de ce complexe ne ressemble à la structure du complexe AVR-Pik/Pik_{HMA}.

Dans le futur, il sera intéressant d'identifier les cibles de la virulence d'AVR1-CO39, AVR-Pia et AVR-Pik, et déterminer si elles sont systématiquement de petites protéines HMA/RATX1. Il sera aussi nécessaire d'étudier, au niveau moléculaire, comment ces trois effecteurs agissent sur leurs cibles de virulence. Le travail réalisé par le groupe de recherche de R. Terauchi (IBRC, Iwate, Japon) indique qu'AVR-Pik cible un clade de petites protéines HMA (sHMA) semblables aux domaines intégrés Pik-HMA ou RGA5-RATX1. Sur la base de ces résultats, ils ont pu déterminer qu'AVR-Pik établit le même type d'interaction avec ses cibles et avec le domaine HMA de Pik. Ces interactions ont également montré que les petites protéines HMA contribuent à l'augmentation de la sensibilité du riz à *M. oryzae* et à la réduction de la production d'espèces réactives de l'oxygène (R. Terauchi, communication

personnelle). Il reste toutefois à déterminer si AVR-Pia et AVR1-CO39 ciblent le même clade de sHMAs, et si elles interagissent par le même type de mécanisme.

Les domaines intégrés des NLRs sont impliqués dans la reconnaissance des effecteurs et peuvent être utilisés pour l'identification de cibles d'effecteurs chez la plante

Dans cette étude, nous avons constaté que la liaison d'AVR-Pia au domaine RATX1 est nécessaire pour la reconnaissance de l'effecteur. Ceci conforte le modèle de leurre intégré, qui prédit que les domaines intégrés aux NLR (**NLR-ID**) servent de « leurres » pour piéger et détecter les effecteurs. En effet, jusqu'à ce jour, la seule preuve convaincante appuyant ce modèle avait été fournie par les études de recherche sur les interactions AVR-Pik/Pik-1, PopP2/RRS1 ou AvrRPS4/RRS1. Dans le cas de AVR-Pik/Pik-1 et AvrRPS4/RRS1, il a été démontré que la liaison de l'effecteur au domaine leurre est nécessaire, comme dans le cas de AVR-Pia/RGA5, pour l'activation du NLR. Dans le cas de PopP2/RRS1, le NLR est plutôt activé par une acétylation du domaine leurre intégré par l'effecteur. Cela met en évidence que les domaines intégrés, tout comme les protéines de garde classiques, les protéines leurres ou les cofacteurs, peuvent permettre la reconnaissance d'effecteurs par différentes manières, par exemple en étant impliqués dans la formation du complexe ou en acquérant des modifications post-traductionnelles détectées par le NLR associé.

L'analyse systématique des NLR végétaux a montré que l'intégration des domaines « non conventionnels » se produit à un taux d'environ 5%. De plus des NLR-ID ont été trouvés dans toutes les plantes terrestres, des mousses aux angiospermes (Cesari, et al 2014; Kroj et al 2016; Sarris et al 2016). En outre, il a été observé que la diversité des domaines intégrés est très élevée. En effet, environ 265 domaines intégrés différents ont été identifiés dans des protéines NLR (Sarris et al. 2016). Les domaines les plus représentés sont des protéines kinases, des domaines de liaison à l'ADN ou des domaines d'interaction protéine-protéine (Sarris et al., 2016). De manière intéressante, plusieurs études sur l'interaction protéine-protéine montrent qu'un grand nombre de domaines intégrés pourraient agir comme cibles des effecteurs. Ceci suggère que la reconnaissance des effecteurs est une caractéristique générale des domaines intégrés. Ceci a justement été démontré expérimentalement pour RGA5, Pik-1 et RRS1. Par ailleurs, un grand nombre de domaines intégrés n'ont pas encore été associés à des cibles impliquées dans l'immunité de l'hôte, et constituent donc une liste

exceptionnelle de candidats pour l'identification de nouveaux composants immunitaires (Cesari et al., 2014 ; Kroj et al 2016; Sarris et al 2016).

Le modèle coopératif pour la reconnaissance des effecteurs par NLR-ID

Dans le cadre de ce travail, nous avons découvert que les mutants d'AVR-Pia qui perdent la capacité à se lier au domaine $RGA5_{RATX1}$ sont pourtant capables d'interagir avec les protéines $RGA5_{\Delta RATX1}$. Ceci indique que différentes surfaces d'AVR-Pia sont impliquées dans l'interaction avec RGA5 via d'autres domaines que RATX1. Plus précisément, nous avons montré qu'AVR-Pia et AVR1-CO39 interagissent avec les domaines $RGA5_{NB\text{-}ARC}$ et $RGA5_{LRR}$. L'interaction de l'effecteur reconnu avec de multiples sites de liaison indépendants au sein de la protéine NLR peut constituer un mécanisme moléculaire de résistance efficace pour éviter une perte de reconnaissance qui pourrait être engendrée par des mutations rapides de l'effecteur. Ainsi, au travers d'interactions avec le domaine de leurre intégré et avec d'autres domaines de la protéine NLR, l'effecteur peut être reconnu de manière très robuste et spécifique. Ces multiples interactions pourraient constituer l'un des principaux avantages des protéines NLR-ID.

Pour valider ce modèle, il serait important de démontrer que l'interaction entre l'effecteur et les domaines de la protéine NLR, autres que le leurre intégré, est nécessaire pour la reconnaissance de l'effecteur. Pour cela, il sera nécessaire d'identifier des mutants d'AVR-Pia qui perdent l'association avec $RGA5_{\Delta RATX1}$ mais conservent l'interaction avec $RGA5_{RATX1}$, puis d'évaluer si la perte de cette interaction a un impact sur la reconnaissance de l'effecteur.

Les NLR qui agissent en duo pour la reconnaissance des effecteurs peuvent interagir au travers de différents domaines

L'une des questions abordées lors de cette thèse était de comprendre comment RGA4/RGA5 et plus généralement d'autres NLR qui forment des hétéro-complexes interagissent pour conférer la reconnaissance de l'effecteur. Dans l'hétéro-complexe RRS1/RPS4, un rôle important des interactions homo et hétérotypiques entre les domaines TIRs dans l'activation et la répression du complexe, a été documenté (Williams et al. 2014). Dans le cas de RGA4/RGA5, des interactions entre leurs domaines CC ont déjà été décrites (Cesari, et al.

2014) ; il est possible qu'elles jouent un rôle similaire à celles des domaines TIR du complexe RRS1/RPS4.

Dans la présente étude, nous mettons en évidence que des domaines supplémentaires de RGA4 et RGA5 contribuent à leur hétéro-association. En effet, les domaines NB-ARC et LRR de RGA4 interagissent avec les domaines correspondants dans RGA5. Pour évaluer l'importance de ces interactions dans la fonction de l'hétéro-complexe, il serait intéressant d'identifier des mutations spécifiques dans RGA4 ou RGA5 au sein des domaines CC, NB-ARC ou LRR qui empêchent leur hétéro-association et d'analyser si ces mutations affectent également la reconnaissance de l'effecteur tout comme la signalisation de la défense.

L'hétéro-complexe RGA5/RGA4 ne semble pas être perturbé par la présence de l'effecteur AVR-Pia

Il a été précédemment démontré que RGA4 et RGA5 forment un hétéro-complexe en l'absence de AVR-Pia (Cesari, et al., 2014). Dans cette étude, nous avons obtenu des résultats préliminaires suggérant que la présence de l'effecteur ne modifiait pas la formation de l'hétéro-complexe. Cependant, des expériences supplémentaires sont nécessaires pour tester cette hypothèse. Des expériences possibles consisteraient à des analyses BiFC dans lesquelles l'interaction de RGA4 et RGA5 serait évaluée en l'absence ou en présence de la protéine d'avirulence. En outre, il reste à vérifier si le complexe fonctionnel est bien un complexe tripartite constitué par AVR-Pia/RGA4/RGA5. Dans ce but, on pourrait tester si AVR-Pia co-précipite RGA4 et RGA5 dans des expériences *in planta*.

La signalisation de défense induite par des hétéro-complexes NLR est mal comprise

Dans le cas du couple RGA4/RGA5 il a été montré que RGA4 est capable d'induire la mort cellulaire, soit en l'absence de RGA5, soit en présence d'AVR-Pia (Césari, et al. 2014). Cependant, la manière dont RGA4 induit la signalisation de défense reste une question ouverte. Chez les animaux, il a été démontré que les protéines NLR fonctionnant par paires peuvent former un complexe tripartite avec le ligand du pathogène. Ce triple complexe est ensuite capable d'induire le recrutement de protéines supplémentaires, ce qui donne lieu à la formation d'un complexe multi-protéique appelé inflammasome (Hu et al., 2015). Cet inflammasome est alors composé par une molécule ligand, une protéine NLR qui agit comme

récepteur, et 10 à 12 molécules NLR qui agit comme adaptateur. Ce complexe multi protéique prend la forme d'une roue et crée une surface d'oligomérisation nécessaire à la l'activation de l'immunité innée (Hu et al. 2015). Afin de déterminer si les NLR hétéro-complexes chez la plante peuvent former des structures similaires, il serait donc intéressant d'élucider la stœchiométrie du complexe RGA4/RGA5 actif.

ANNEX 1

Materials & methods chapter III and IV

Plant material and constructs

N. benthamiana plants were grown in a growth chamber at 22°C with a 16-h light period. Plasmids were generated by Gateway cloning (ThermoFisher, Waltham, USA), restriction/ligation, Gateway entry clones were generated using the pDONR207 plasmid (ThermoFisher, Waltham, USA) (Table 1). Gateway destination vectors were modified pBIN19 plasmids for expression of tagged proteins in *N. benthamiana* (Césari et al. 2013; Césari et al. 2014) or modified pGAD-T7 or pGBK-T7 plasmids (Clontech, Mountain View, USA) for yeast two hybrid experiments (Bernoux et al., 2011) (Table 1).

Coimmunoprecipitation and Yeast two hybrid interaction assays

Protein-protein interaction analyses by co-immunoprecipitation were performed with protein extracts from *N.benthamina* leaf discs harvested 2 days after *Agrobacterium* infiltration (Césari et al., 2013). For the interaction of AVR-Pia and AVR1-C039 with RGA5 variants as well as the interactions between RGA5 and RGA4 domains, 5 leaf disks per sample were homogenized in extraction buffer (50mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 10 mM DTT, 1 mM PMSF, 1.0% IGEPAL CA-630 [NP-40], 0.05% sodium deoxycholate, and 0.1% SDS, supplemented with a complete protease inhibitor cocktail (Roche) and polyvinylpolypyrrolidone 0.5%). And co-IP was performed with 8 uL of agarose GFP_trap_A suspension (Chromotek) and four washes with the modified protein extraction buffer. For the interactions between $RGA5_{RATX1}$ domain and RGA5 variants we used the same extraction buffer described above but we also used a modified extraction buffer with 0.5% IGEPAL CA-630 [NP-40].

Bound proteins were eluted by boiling for 10 min at 70°C in 50 uL of Nupage sample buffer, separated by polyacrylamide gel electrophoresis using NuPAGE 4-12% gels (Invitrogen, Carlsbad, CA, USA), transferred to nitrocellulose membrane (Millipore), and analyzed by immunoblotting. For immunodetection of proteins, rat anti-HA-horseradish peroxidase (clone 3F10; Roche) or mouse anti-GFP (Roche) and goat anti-mouse-horseradishperoxidase (Sigma-Aldrich) were used in combination with the Immobilon western kit (Millipore).

Binding domain (BD) fusions of AVR-Pia, AVR1-CO39 and PWL2 as well RGA5 variants in

pGBKT7-53 and activation domain (AD) fusions of AVR-Pia, AVR1-CO39 and empty vector as well as RGA5 variants in pGADT7, were transformed in Gold and Y187 yeast strain respectively. Interactions assays were performed according to the Matchmaker Gold yeast two-hybrid system protocol (Clontech).

	Entry clone	**Destination vectors yeast**		**Destination vectors in *N. bentamiana***			
Insert	**pDONOR207**	**pGBKT7-GW**	**pGADT7-GW**	**pBIN YFP-GTW**	**pBIN19 3XHA-GTW**	**PTK205 GWT-GFP**	**pBIN19 GTW-3XHA**
AVR-Pia*	pSC060	pDO143	pDO155	pSC80	pSC070	-	-
AVR-Pia F24S	pCV84	-	-	pCV128	-	-	-
AVR1-CO39	pSC059	pDO145	pDO157	pSC079	-	-	-
PWL2	pSC120	pDO146	pDO158	pDO119	-	-	-
RGA5	pSC42	-	-	pSC078	pCV129	-	-
$RGA5_{C\text{-}ter}$	pSC129	pDO38	pDO49	pSC273	pSC144	-	-
$RGA5_{RATX1}$	pSC207	-	-	pSC274	pDO120	-	-
$RGA5_{\Delta RATX1}$	pSC210	pDO141	pDO153	pSC276	pDO121	-	-
$RGA5_{CC\text{-}NB}$	pSC193	pDO135	pDO147	pSC267	pDO123	-	-
$RGA5_{CC\text{-}NB\text{-}ARC}$	pSC195	pDO136	pDO148	pSC268	pDO124	-	-
$RGA5_{NB}$	pSC197	pDO137	pDO149	pSC269	pDO125	-	-
$RGA5_{NB\text{-}ARC}$	pSC199	pDO138	pDO150	pSC270	pDO126	-	-
$RGA5_{LRR\text{-}Cter}$	pSC203	pDO139	pDO151	pSC272	pDO127	-	-
$RGA5_{\Delta CC}$	pSC201	-	-	pSC271	PDO159		
RGA4	pSC41	-	-			pCV094	pSC061
$RGA4_{CC\text{-}NB}$	pSC179	-	-	-	-	pSC218	pSC225

RGA4 $_{\text{CC-NB-ARC}}$ 547	pSC181	-	-	-	-	pSC219	pSC226
RGA4 $_{\text{NB}}$	pSC183	-	-	-	-	pSC220	pSC227
RGA4 $_{\text{NB-ARC}}$	pSC185	-	-	-	-	pSC221	pSC228
RGA4 $_{\Delta\text{CC}}$	pSC187	-	-	-	-	pSC222	pSC229
RGA4 $_{\text{LRR-Cter}}$	pSC189	-	-	-	-	pSC223	pSC230
RGA4 $_{\Delta\text{Cter}}$	pSC192	-	-	-	-	pSC224	pSC231

*AVR-Pia without tag pSC95

BIBLIOGRAPHY

Aarts, N., Metz, M., Holub, E., Staskawicz, B.J., Daniels, M.J., and Parker, J.E. (1998). Different requirements for *EDS1* and *NDR1* by disease resistance genes define at least two R gene-mediated signaling pathways in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. **95**: 10306–10311.

Ade, J., DeYoung, B.J., Golstein, C., and Innes, R.W. (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. Proc. Natl. Acad. Sci. U. S. A. **104**: 2531–2536.

Ashikawa, I., Hayashi, N., Yamane, H., Kanamori, H., Wu, J., Matsumoto, T., Ono, K., and Yano, M. (2008). Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer Pikm-specific rice blast resistance. Genetics **180**: 2267–2276.

Balmuth, A. and Rathjen, J.P. (2007). Genetic and molecular requirements for function of the Pto/Prf effector recognition complex in tomato and *Nicotiana benthamiana*. Plant J. **51**: 978–990.

Barthe, P., Ropars, V., and Roumestand, C. (2006). DYNAMOF: a program for the dynamics analysis of relaxation data obtained at multiple magnetic fields. Comptes Rendus Chim. **9**: 503–513.

Bekal, S., Niblack, T.L., and Lambert, K.N. (2003). A chorismate mutase from the soybean cyst nematode Heterodera glycines shows polymorphisms that correlate with virulence. Mol. Plant. Microbe. Interact. **16**: 439–446.

Bendahmane, A., Farnham, G., Moffett, P., and Baulcombe, D.C. (2002). Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the *Rx* locus of potato. Plant J. **32**: 195–204.

Bentham, A., Burdett, H., Anderson, P.A., Williams, S.J., and Kobe, B. (2016). Animal NLRs provide structural insights into plant NLR function. Ann. Bot.: 1–14.

Bernoux, M., Burdett, H., Williams, S.J., Zhang, X., Chen, C., Newell, K., Lawrence, G., Kobe, B., Ellis, J.G., Anderson, P., and Dodds, P.N. (2016). Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. Plant Cell: TPC2015-00303-RA.

Bernoux, M., Ellis, J.G., and Dodds, P.N. (2011a). New insights in plant immunity signaling activation. Curr. Opin. Plant Biol. **14**: 512–8.

Bernoux, M., Ve, T., Williams, S., Warren, C., Hatters, D., Valkov, E., Zhang, X., Ellis, J.G., Kobe, B., and Dodds, P.N. (2011b). Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. Cell Host Microbe **9**: 200–211.

Böhm, H., Albert, I., Fan, L., Reinhard, A., and Nürnberger, T. (2014). Immune receptor complexes at the plant cell surface. Curr. Opin. Plant Biol. **20C**: 47–54.

Bohm, H., Albert, I., Oome, S., Raaymakers, T.M., Van den Ackerveken, G., and Nurnberger, T. (2014). A Conserved Peptide Pattern from a Widespread Microbial Virulence Factor Triggers Pattern-Induced Immunity in Arabidopsis. PLoS Pathog. **10(11)**: e1004491.

Boller, T. and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu. Rev. Plant Biol. **60**: 379–406.

Bonardi, V., Tang, S.J., Stallmann, A., Roberts, M., Cherkis, K., and Dangl, J.L. (2011). Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. Proc. Natl. Acad. Sci. U. S. A. **108**: 16463–16468.

Boutemy, L.S., King, S.R.F., Win, J., Hughes, R.K., Clarke, T.A., Blumenschein, T.M.A., Kamoun, S., and Banfield, M.J. (2011). Structures of Phytophthora RXLR effector proteins: A conserved but adaptable fold underpins functional diversity. J. Biol. Chem. **286**: 35834–35842.

Boyd, L.A., Ridout, C., O'Sullivan, D.M., Leach, J.E., and Leung, H. (2013). Plant-pathogen interactions: Disease resistance in modern agriculture. Trends Genet. **29**: 233–240.

Brown, J.K.M. (2015). Durable Resistance of Crops to Disease: A Darwinian Perspective. Annu. Rev. Phytopathol. **53**: 150615190330008.

Brown, J.K.M. and Tellier, A. (2011). Plant-parasite coevolution: bridging the gap between genetics and ecology. Annu. Rev. Phytopathol. **49**: 345–67.

Brutus, A., Sicilia, F., Macone, A., Cervone, F., De Lorenzo, G., and Lorenzo, G. De (2010). A domain swap approach reveals a role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. Proc. Natl. Acad. Sci. U. S. A. **107**: 9452–7.

Buell, C.R. (2002). Interactions between Xanthomonas species and Arabidopsis thaliana. Arab. B. **14**: 1.

Burch-Smith, T.M., Schiff, M., Caplan, J.L., Tsao, J., Czymmek, K., and Dinesh-Kumar, S.P. (2007). A novel role for the TIR domain in association with pathogen-derived elicitors. PLoS Biol. **5**: 0501–0514.

Cao, Y., Liang, Y., Tanaka, K., Nguyen, C.T., Jedrzejczak, R.P., Joachimiak, A., and Stacey, G. (2014). The kinase LYK5 is a major chitin receptor in Arabidopsis and forms a chitin-induced complex with related kinase CERK1. Elife **3**: e03766.

Caplan, J.L., Mamillapalli, P., Burch-Smith, T.M., Czymmek, K., and Dinesh-Kumar, S.P. (2008). Chloroplastic Protein NRIP1 Mediates Innate??Immune Receptor Recognition of a Viral Effector. Cell **132**: 449–462.

Cesari, S. et al. (2013). The rice resistance protein pair RGA4/RGA5 recognizes the Magnaporthe oryzae effectors AVR-Pia and AVR1-CO39 by direct binding. Plant Cell **25**: 1463–81.

Césari, S., Bernoux, M., Moncuquet, P., Kroj, T., and Dodds, P.N. (2014). A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. Front. Plant Sci. **5**: 606.

Césari, S., Kanzaki, H., Fujiwara, T., Bernoux, M., Chalvon, V., Kawano, Y., Shimamoto, K., Dodds, P., Terauchi, R., and Kroj, T. (2014). The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease resistance. EMBO J.: 1–19.

Chanclud, E., Kisiala, A., Emery, N.R.J., Chalvon, V., Ducasse, A., Romiti-Michel, C., Gravot, A., Kroj, T., and Morel, J.B. (2016). Cytokinin Production by the Rice Blast Fungus Is a Pivotal Requirement for Full Virulence. PLoS Pathog. **12**: 1–25.

Chang, C., Yu, D., Jiao, J., Jing, S., Schulze-Lefert, P., and Shen, Q.-H. (2013). Barley MLA immune receptors directly interfere with antagonistically acting transcription factors to initiate disease resistance signaling. Plant Cell **25**: 1158–73.

Chen, D., Chen, X., MA, B., WANG, Y., Zhu, L., and Li, S. (2010). Genetic Transformation of Rice with Pi-d2 Gene Enhances Resistance to Rice Blast Fungus Magnaporthe oryzae. Rice Sci. **17**: 19–27.

Chen, S., Songkumarn, P., Venu, R.C., Gowda, M., Bellizzi, M., Hu, J., Liu, W., Ebbole, D., Meyers, B., Mitchell, T., and Wang, G.-L. (2013). Identification and characterization of in planta-expressed secreted effector proteins from Magnaporthe oryzae that induce cell death in rice. Mol. Plant. Microbe. Interact. **26**: 191–202.

Chen, X. et al. (2006). A B-lectin receptor kinase gene conferring rice blast resistance. Plant J. **46**: 794–804.

Chiapello, H. et al. (2015). Deciphering Genome Content and Evolutionary Relationships of Isolates from the Fungus *Magnaporthe oryzae* Attacking Different Host Plants. Genome Biol. Evol. **7**: 2896–912.

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**: 803–14.

Choi, J., Park, J., Kim, D., Jung, K., Kang, S., and Lee, Y.-H. (2010). Fungal secretome database: integrated platform for annotation of fungal secretomes. BMC Genomics **11**: 105.

Choi, J., Tanaka, K., Cao, Y., Qi, Y., Qiu, J., Liang, Y., Lee, S.Y., and Stacey, G. (2014). Identification of a plant receptor for extracellular ATP. Science **343**: 290–4.

Chung, E.H., Da Cunha, L., Wu, A.J., Gao, Z., Cherkis, K., Afzal, A.J., MacKey, D., and Dangl, J.L. (2011). Specific threonine phosphorylation of a host target by two unrelated type III effectors activates a host innate immune receptor in plants. Cell Host Microbe **9**: 125–136.

Civá, P., Craig, H., Cox, C.J., and Brown, T.A. (2015). Three geographically separate domestications of Asian rice. **1**: 1–5.

Clarke, C.R., Chinchilla, D., Hind, S.R., Taguchi, F., Miki, R., Ichinose, Y., Martin, G.B., Leman, S., Felix, G., and Vinatzer, B.A. (2013). Allelic variation in two distinct Pseudomonas syringae flagellin epitopes modulates the strength of plant immune responses but not bacterial motility. New Phytol. **200**: 847–860.

Collier, S.M., Hamel, L., and Moffett, P. (2011). Cell Death Mediated by the N-Terminal Domains of a Unique and Highly Conserved Class of NB-LRR Protein. Mol. Plant. Microbe. Interact. **24**: 918–931.

Collier, S.M. and Moffett, P. (2009). NB-LRRs work a “bait and switch” on pathogens. Trends Plant Sci. **14**: 521–9.

Cook, D.E., Mesarich, C.H., and Thomma, B.P.H.J. (2014). Understanding Plant Immunity as a Surveillance System to Detect Invasion. Annu. Rev. Phytopathol. **53**: 150605182533006.

Costanzo, S. and Jia, Y. (2009). Alternatively spliced transcripts of Pi-ta blast resistance gene in *Oryza sativa*. Plant Sci. **177**: 468–478.

Couch, B.C., Fudal, I., Lebrun, M.H., Tharreau, D., Valent, B., Van Kim, P., Nottenghem, J.L., and Kohn, L.M. (2005). Origins of host-specific populations of the blast pathogen Magnaporthe oryzae in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. Genetics **170**: 613–630.

Couto, D. and Zipfel, C. (2016). Regulation of pattern recognition receptor signalling in plants. Nat.

Rev. Immunol.

Črnigoj Kristan, K., Viero, G., Dalla Serra, M., Maček, P., and Anderluh, G. (2009). Molecular mechanism of pore formation by actinoporins. Toxicon **54**: 1125–1134.

Dangl, J.L., Horvath, D.M., and Staskawicz, B.J. (2013). Pivoting the plant immune system from dissection to deployment. Science **341**: 746–51.

Dangl, J.L. and Jones, J.D.G. (2001). Defence Responses To Infection. Nature **411**: 826–833.

Dangl, J.L. and McDowell, J.M. (2006). Two modes of pathogen recognition by plants. Proc. Natl. Acad. Sci. U. S. A. **103**: 8575–6.

Danot, O. (2015). How "arm-twisting" by the inducer triggers activation of the MalT transcription factor, a typical signal transduction ATPase with numerous domains (STAND). Nucleic Acids Res. **43**: 3089–3099.

Day, B. (2005). Molecular Basis for the RIN4 Negative Regulation of RPS2 Disease Resistance. Plant Cell Online **17**: 1292–1305.

Dean, R. a et al. (2005). The genome sequence of the rice blast fungus Magnaporthe grisea. Nature **434**: 980–6.

Dean, R., Van Kan, J.A.L., Pretorius, Z.A., Hammond-Kosack, K.E., Di Pietro, A., Spanu, P.D., Rudd, J.J., Dickman, M., Kahmann, R., Ellis, J., and Foster, G.D. (2012). The Top 10 fungal pathogens in molecular plant pathology. Mol. Plant Pathol. **13**: 414–430.

Delteil, A., Gobbato, E., Cayrol, B., Estevan, J., Michel-Romiti, C., Dievart, A., Kroj, T., and Morel, J.-B. (2016). Several wall-associated kinases participate positively and negatively in basal defense against rice blast fungus. BMC Plant Biol. **16**: 17.

Djamei, A. et al. (2011). Metabolic priming by a secreted fungal effector. Nature **1**: 0–5.

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. Proc. Natl. Acad. Sci. U. S. A. **103**: 8888–93.

Dodds, P.N. and Rathjen, J.P. (2010). Plant immunity: towards an integrated view of plant-pathogen interactions. Nat. Rev. Genet. **11**: 539–548.

Doehlemann, G., Van Der Linde, K., Aßmann, D., Schwammbach, D., Hof, A., Mohanty, A., Jackson, D., and Kahmann, R. (2009). Pep1, a secreted effector protein of Ustilago maydis, is required for successful invasion of plant cells. PLoS Pathog. **5**.

Dong, S. et al. (2014). Effector specialization in a lineage of the Irish potato famine pathogen. Science **343**: 552–5.

Dong, X. and Kahmann, R. (2009). Battle for survival: plants and their allies and enemies. Curr. Opin. Plant Biol. **12**: 387–9.

Dou, D., Kale, S.D., Wang, X., Jiang, R.H.Y., Bruce, N.A., Arredondo, F.D., Zhang, X., and Tyler, B.M. (2008). RXLR-Mediated Entry of Phytophthora sojae Effector Avr1b into Soybean Cells Does Not Require Pathogen-Encoded Machinery. Plant Cell Online **20**: 1930–1947.

Duplessis, S. (2011). Obligate biotrophy features unraveled by the genomic analysis of rust fungi.

Proc. Natl. Acad. Sci. U. S. A.: 1–23.

Eitas, T.K. and Dangl, J.L. (2010). NB-LRR proteins: Pairs, pieces, perception, partners, and pathways. Curr. Opin. Plant Biol. **13**: 472–477.

Eitas, T.K., Nimchuk, Z.L., and Dangl, J.L. (2008). Arabidopsis TAO1 is a TIR-NB-LRR protein that contributes to disease resistance induced by the Pseudomonas syringae effector AvrB. Proc. Natl. Acad. Sci. U. S. A. **105**: 6475–6480.

Ellis, J.G. (2016). Integrated decoys and effector traps: how to catch a plant pathogen. BMC Biol. **14**: 13.

Engelhardt, S., Boevink, P.C., Armstrong, M.R., Ramos, M.B., Hein, I., and Birch, P.R.J. (2012). Relocalization of Late Blight Resistance Protein R3a to Endosomal Compartments Is Associated with Effector Recognition and Required for the Immune Response. Plant Cell **24**: 5142–5158.

van Esse, H.P., Bolton, M.D., Stergiopoulos, I., de Wit, P.J.G.M., and Thomma, B.P.H.J. (2007). The chitin-binding Cladosporium fulvum effector protein Avr4 is a virulence factor. Mol. Plant. Microbe. Interact. **20**: 1092–1101.

van Esse, H.P., van't Klooster, J.W., Bolton, M.D., Yadeta, K.A., van Baarlen, P., Boeren, S., Vervoort, J., de Wit, P.J.G.M., and Thomma, B.P.H.J. (2008). The Cladosporium fulvum Virulence Protein Avr2 Inhibits Host Proteases Required for Basal Defense. Plant Cell **20**: 1948–1963.

Faivre-Rampant, O., Thomas, J., Allègre, M., Morel, J.B., Tharreau, D., Nottéghem, J.L., Lebrun, M.H., Schaffrath, U., and Piffanelli, P. (2008). Characterization of the model system rice-Magnaporthe for the study of nonhost resistance in cereals. New Phytol. **180**: 899–910.

Fellbrich, G., Romanski, A., Varet, A., Blume, B., Brunner, F., Engelhardt, S., Felix, G., Kemmerling, B., Krzymowska, M., and N??rnberger, T. (2002). NPP1, a Phytophthora-associated trigger of plant defense in parsley and Arabidopsis. Plant J. **32**: 375–390.

Frost, D., Way, H., Howles, P., Luck, J., Manners, J., Hardham, A., Finnegan, J., and Ellis, J. (2004). Tobacco transgenic for the flax rust resistance gene L expresses allele-specific activation of defense responses. Mol. Plant. Microbe. Interact. **17**: 224–32.

Fujisaki, K., Abe, Y., Ito, A., Saitoh, H., Yoshida, K., Kanzaki, H., and Kanzaki, E. (2015). Rice Exo70 interacts with a fungal effector , AVR-Pii , and is required for AVR-Pii-triggered immunity. The Plant J. **83**: 875–887.

Gao, Z., Chung, E., Eitas, T.K., and Dangl, J.L. (2011). Plant intracellular innate immune receptor Resistance to Pseudomonas syringae pv. maculicola 1 (RPM1) is activated at, and functions on, the plasma membrane. Proc. Natl. Acad. Sci. **108**: 8914–8914.

Garris, A.J., Tai, T.H., Coburn, J., Kresovich, S., and McCouch, S. (2005). Genetic structure and diversity in Oryza sativa L. Genetics **169**: 1631–1638.

Giannakopoulou, A., Steele, J.F.C., Segretin, M.E., Bozkurt, T.O., Zhou, J., Robatzek, S., Banfield, M.J., Pais, M., and Kamoun, S. (2015). Tomato I2 immune receptor can be engineered to confer partial resistance to the oomycete Phytophthora infestans in addition to the fungus Fusarium oxysporum. Mol. Plant. Microbe. Interact. **28**: 1–45.

Gijzen, M. and Nurnberger, T. (2006). Nep1-like proteins from plant pathogens: Recruitment and diversification of the NPP1 domain across taxa. Phytochemistry **67**: 1800–1807.

Giraldo, M.C., Dagdas, Y.F., Gupta, Y.K., Mentlak, T. a, Yi, M., Martinez-Rocha, A.L., Saitoh, H., Terauchi, R., Talbot, N.J., and Valent, B. (2013). Two distinct secretion systems facilitate tissue invasion by the rice blast fungus Magnaporthe oryzae. Nat. Commun. **4**: 1996.

Giraldo, M.C. and Valent, B. (2013). Filamentous plant pathogen effectors in action. Nat. Rev. Microbiol. **11**: 800–14.

Godfrey, D., Böhlenius, H., Pedersen, C., Zhang, Z., Emmersen, J., and Thordal-Christensen, H. (2010). Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. BMC Genomics **11**: 317.

Göhre, V. and Robatzek, S. (2008). Breaking the barriers: microbial effector molecules subvert plant immunity. Annu. Rev. Phytopathol. **46**: 189–215.

Gómez-Gómez, L. and Boller, T. (2000). FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. Mol. Cell **5**: 1003–1011.

Griebel, T., Maekawa, T., and Parker, J.E. (2014). NOD-like receptor cooperativity in effector-triggered immunity. Trends Immunol. **35**: 562–570.

Gutierrez, J.R., Balmuth, A.L., Ntoukakis, V., Mucyn, T.S., Gimenez-Ibanez, S., Jones, A.M.E., and Rathjen, J.P. (2010). Prf immune complexes of tomato are oligomeric and contain multiple Pto-like kinases that diversify effector recognition. Plant J. **61**: 507–518.

Guttman, D., Mchardy, A.C., and Schulze-lefert, P. (2014). Microbial genome-enabled insights into plant – microorganism interactions. Nat. Publ. Gr. **15**: 797–813.

Guyon, K., Balagué, C., Roby, D., and Raffaele, S. (2014). Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. BMC Genomics **15**: 336.

Hacquard, S., Kracher, B., Maekawa, T., Vernaldi, S., Schulze-Lefert, P., and Ver Loren van Themaat, E. (2013). Mosaic genome structure of the barley powdery mildew pathogen and conservation of transcriptional programs in divergent hosts. Proc. Natl. Acad. Sci. U. S. A. **110**: E2219-28.

Hamel, L.-P., Sekine, K.-T., Wallon, T., Sugiwaka, Y., Kobayashi, K., and Moffett, P. (2016). The chloroplastic protein THF1 interacts with the coiled-coil domain of the disease resistance protein N' and regulates light-dependent cell death. Plant Physiol.: pp.00234.2016.

Hammond-kosack, K.E. and Jones, J.D.G. (1996). Resistance Gene-Dependent Plant Defense Responses.Plant Cell **8**: 1773–1791.

Hao, W., Collier, S.M., Moffett, P., and Chai, J. (2013). Structural basis for the interaction between the potato virus X resistance protein (Rx) and its cofactor ran GTPase-activating protein 2 (RanGAP2). J. Biol. Chem. **288**: 35868–35876.

Harkenrider, M., Sharma, R., De Vleesschauwer, D., Tsao, L., Zhang, X., Chern, M., Canlas, P., Zuo, S., and Ronald, P.C. (2016). Overexpression of Rice Wall-Associated Kinase 25 (OsWAK25) alters resistance to bacterial and fungal pathogens. PLoS One **11**: 1–16.

Harris, C.J., Slootweg, E.J., Goverse, A., and Baulcombe, D.C. (2013). Stepwise artificial evolution of a plant disease resistance gene. Proc. Natl. Acad. Sci. U. S. A. **110**: 21189–94.

Heidrich, K., Blanvillain-Baufumé, S., and Parker, J.E. (2012). Molecular and spatial constraints on NB-LRR receptor signaling. Curr. Opin. Plant Biol. **15**: 385–391.

Heidrich, K., Wirthmueller, L., Tasset, C., and Pouzet, C. (2011). Reports 10. **2473**: 1401–1405.

Hirsch, A., Lum, M., and Downie, J. (2001). What makes the rhizobia-legume symbiosis so special? Plant Physiol. **127**: 1484–1492.

Hogenhout, S. a, Van der Hoorn, R. a L., Terauchi, R., and Kamoun, S. (2009). Emerging concepts in effector biology of plant-associated organisms. Mol. Plant. Microbe. Interact. **22**: 115–22.

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**: 2009–17.

Hossain, M. (2007). Technological progress for sustaining food-population balance: Achievement and challenges. Agric. Econ. **37**: 161–172.

Howles, P., Lawrence, G., Finnegan, J., Mcfadden, H., Ayliffe, M., and Dodds, P. (2005). Autoactive Alleles of the Flax L6 Rust Resistance Gene Induce Non-Race-Specific Rust Resistance Associated with the Hypersensitive Response. **18**: 570–582.

Hu, W., Lv, Y., Lei, W., Li, X., Chen, Y., Zheng, L., Xia, Y., and Shen, Z. (2014). Cloning and characterization of the Oryza sativa wall-associated kinase gene OsWAK11 and its transcriptional response to abiotic stresses. Plant Soil **384**: 335–346.

Hu, Z. et al. (2013). Crystal structure of NLRC4 reveals its autoinhibition mechanism. Science . **341**: 172–5.

Hu, Z., Zhou, Q., Zhang, C., Fan, S., Cheng, W., Zhao, Y., Shao, F., Wang, H.-W., Sui, S.-F., and Chai, J. (2015). Structural and biochemical basis for induced self-propagation of NLRC4. Science. **350**: 399–404.

Inoue, H., Hayashi, N., Matsushita, A., Xinqiong, L., Nakayama, A., Sugano, S., Jiang, C.-J., and Takatsuji, H. (2013a). Blast resistance of CC-NB-LRR protein Pb1 is mediated by WRKY45 through protein-protein interaction. Proc. Natl. Acad. Sci. U. S. A. **110**: 9577–82.

Inoue, H., Hayashi, N., Matsushita, A., Xinqiong, L., Nakayama, A., Sugano, S., Jiang, C.-J., and Takatsuji, H. (2013b). Blast resistance of CC-NB-LRR protein Pb1 is mediated by WRKY45 through protein-protein interaction. Proc. Natl. Acad. Sci. U. S. A. **110**: 9577–82.

Jacob, F., Takaki, M., and Vernaldi, S. (2013). Evolution and conservation of plant NLR functions. Front. Immunol.: 211.

Jia, Y. and Martin, R. (2008). Identification of a new locus, Ptr(t), required for rice blast resistance gene Pi-ta-mediated resistance. Mol. Plant. Microbe. Interact. **21**: 396–403.

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 J. **19**: 4004–4014.

Jiang, N. et al. (2012). Molecular mapping of the Pi2/9 allelic gene Pi2-2 conferring broad-spectrum resistance to Magnaporthe oryzae in the rice cultivar Jefferson. Rice **5**: 29.

Joly, D.L., Feau, N., Tanguay, P., and Hamelin, R.C. (2010). Comparative analysis of secreted protein evolution using expressed sequence tags from four poplar leaf rusts (Melampsora spp.). BMC Genomics **11**: 422.

Jones, J.D.G. and Dangl, J.L. (2006). The plant immune system. Nature **444**: 323–9.

de Jonge, R., Bolton, M.D., and Thomma, B.P.H.J. (2011). How filamentous pathogens co-opt plants: the ins and outs of fungal effectors. Curr. Opin. Plant Biol. **14**: 400–6.

de Jonge, R., van Esse, H.P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., van der Krol, S., Shibuya, N., Joosten, M.H. a J., and Thomma, B.P.H.J. (2010). Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. Science **329**: 953–5.

Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K., Minami, E., and Shibuya, N. (2006). Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. Proc. Natl. Acad. Sci. U. S. A. **103**: 11086–11091.

Kale, S.D. et al. (2010). External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. Cell **142**: 284–95.

Kamoun, S. (2006). A catalogue of the effector secretome of plant pathogenic oomycetes. Annu. Rev. Phytopathol. **44**: 41–60.

Kanzaki, H., Yoshida, K., Saitoh, H., Fujisaki, K., Hirabuchi, A., Alaux, L., Fournier, E., Tharreau, D., and Terauchi, R. (2012). Arms race co-evolution of Magnaporthe oryzae AVR-Pik and rice Pik genes driven by their physical interactions. Plant J. **72**: 894–907.

Katagiri, F., Thilmony, R., and He, S.Y. (2002). The Arabidopsis Thaliana-Pseudomonas Syringae Interaction. Arab. B. **1**: e0039.

Kawano, Y., Akamatsu, A., Hayashi, K., Housen, Y., Okuda, J., Yao, A., Nakashima, A., Takahashi, H., Yoshida, H., Wong, H.L., Kawasaki, T., and Shimamoto, K. (2010). Activation of a Rac GTPase by the NLR family disease resistance protein Pit plays a critical role in rice innate immunity. Cell Host Microbe **7**: 362–375.

Kawano, Y. and Shimamoto, K. (2013). Early signaling network in rice PRR-mediated and R-mediated immunity. Curr. Opin. Plant Biol. **16**: 496–504.

Khan, M., Subramaniam, R., and Desveaux, D. (2016). Of guards , decoys , baits and traps : pathogen perception in plants by type III effector sensors. Curr. Opin. Microbiol. **29**: 49–55.

Khang, C.H., Berruyer, R., Giraldo, M.C., Kankanala, P., Park, S.-Y., Czymmek, K., Kang, S., and Valent, B. (2010). Translocation of Magnaporthe oryzae effectors into rice cells and their subsequent cell-to-cell movement. Plant Cell **22**: 1388–403.

Kim, H.-S., Desveaux, D., Singer, A.U., Patel, P., Sondek, J., and Dangl, J.L. (2005). The Pseudomonas syringae effector AvrRpt2 cleaves its C-terminally acylated target, RIN4, from Arabidopsis membranes to block RPM1 activation. Proc. Natl. Acad. Sci. U. S. A. **102**: 6496–501.

Kim, S.H., Qi, D., and Helm, M. (2016). Using decoys to expand the recognition specificity of a plant disease resistance protein. Plant Immunity.**351:**6274

Kleemann, J., Rincon-Rivera, L.J., Takahara, H., Neumann, U., van Themaat, E.V.L., van der Does, H.C., Hacquard, S., Stuber, K., Will, I., Schmalenbach, W., Schmelzer, E., and O'Connell, R.J. (2012). Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen colletotrichum higginsianum. PLoS Pathog. **8**.

Koeck, M., Hardham, A.R., and Dodds, P.N. (2011). The role of effectors of biotrophic and hemibiotrophic fungi in infection. Cell. Microbiol. **13**: 1849–1857.

Kofoed, E.M. and Vance, R.E. (2011). Innate immune recognition of bacterial ligands by NAIPs

determines inflammasome specificity. Nature **477**: 592–5.

Kombrink, A. and Thomma, B.P.H.J. (2013). LysM Effectors : Secreted Proteins Supporting Fungal Life. PLoS Pathog. **9**.

Krasileva, K. V, Dahlbeck, D., and Staskawicz, B.J. (2010). Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. Plant Cell **22**: 2444–2458.

Kroj, T., Chanclud, E., Michel-Romiti, C., Grand, X., and Morel, J.B. (2016). Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. New Phytol. **210**: 618–626.

Küfner, I., Ottmann, C., Oecking, C., and Nürnberger, T. (2009). Cytolytic toxins as triggers of plant immune response. Plant Signal. Behav. **4**: 977–9.

Kunze, G. (2004). The N Terminus of Bacterial Elongation Factor Tu Elicits Innate Immunity in Arabidopsis Plants. Plant Cell Online **16**: 3496–3507.

Lacomme, C. and Santa Cruz, S. (1999). Bax-induced cell death in tobacco is similar to the hypersensitive response. Proc. Natl. Acad. Sci. U. S. A. **96**: 7956–7961.

Le Roux, C. et al. (2015). A Receptor Pair with an Integrated Decoy Converts Pathogen Disabling of Transcription Factors to Immunity. Cell **161**: 1074–1088.

Lee, S.K. et al. (2009). Rice Pi5-mediated resistance to Magnaporthe oryzae requires the presence of two coiled-coil-nucleotide-binding-leucine-rich repeat genes. Genetics **181**: 1627–1638.

Lehmann, S., Serrano, M., L'Haridon, F., Tjamos, S.E., and Metraux, J.P. (2015). Reactive oxygen species and plant resistance to fungal pathogens. Phytochemistry **112**: 54–62.

Leipe, D.D., Koonin, E. V., and Aravind, L. (2004). STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: Multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. J. Mol. Biol. **343**: 1–28.

Li, W. et al. (2009). The Magnaporthe oryzae avirulence gene AvrPiz-t encodes a predicted secreted protein that triggers the immunity in rice mediated by the blast resistance gene Piz-t. Mol. Plant. Microbe. Interact. **22**: 411–20.

Li, Z., Yin, Z., Fan, Y., Xu, M., Kang, Z., and Huang, L. (2015). Candidate effector proteins of the necrotrophic apple canker pathogen Valsa mali can suppress BAX-induced PCD. Front. Plant Sci. **6**: 579.

Lieberherr, D. (2005). A Sphingolipid Elicitor-Inducible Mitogen-Activated Protein Kinase Is Regulated by the Small GTPase OsRac1 and Heterotrimeric G-Protein in Rice. Plant Physiol. **138**: 1644–1652.

Liu, T., Liu, Z., Song, C., Hu, Y., Han, Z., She, J., Fan, F., Wang, J., Jin, C., Chang, J., Zhou, J.-M.J.-M., and Chai, J. (2012). Chitin-Induced Dimerization Activates a Plant Immune Receptor. Science. **336**: 1160–4.

Liu, W. and Wang, G.-L. (2016). Plant innate immunity in rice: a defense against pathogen infection. Natl. Sci. Rev.: nww015.

Loutre, C., Wicker, T., Travella, S., Galli, P., Scofield, S., Fahima, T., Feuillet, C., and Keller, B. (2009). Two different CC-NBS-LRR genes are required for Lr10-mediated leaf rust resistance in

tetraploid and hexaploid wheat. Plant J. **60**: 1043–1054.

Lukasik, E. and Takken, F.L.W. (2009). STANDing strong, resistance proteins instigators of plant defence. Curr. Opin. Plant Biol. **12**: 427–36.

Lukasik-Shreepaathy, E., Slootweg, E., Richter, H., Goverse, A., Cornelissen, B.J.C., and Takken, F.L.W. (2012). Dual regulatory roles of the extended N terminus for activation of the tomato MI-1.2 resistance protein. Mol. Plant. Microbe. Interact. **25**: 1045–57.

Luo, S., Zhang, Y., Hu, Q., Chen, J., Li, K., Lu, C., Liu, H., Wang, W., and Kuang, H. (2012). Dynamic Nucleotide-Binding Site and Leucine-Rich Repeat-Encoding Genes in the Grass Family. Plant Physiol. **159**: 197–210.

Ma, L.-J. et al. (2010). Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature **464**: 367–373.

Macho, A.P. and Zipfel, C. (2014). Plant PRRs and the activation of innate immune signaling. Mol. Cell **54**: 263–272.

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.

Maekawa, T. et al. (2011). Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. Cell Host Microbe **9**: 187–199.

Malik, S. and Van der Hoorn, R. (2016). Commentary Inspirational decoys : a new hunt for effector targets.New Phyt: 2014–2016.

Maqbool, a, Saitoh, H., Franceschetti, M., Stevenson, C., Uemura, a, Kanzaki, H., Kamoun, S., Terauchi, R., and Banfield, M. (2015). Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. Elife **4**: 1–24.

Mentlak, T. a., Kombrink, a., Shinya, T., Ryder, L.S., Otomo, I., Saitoh, H., Terauchi, R., Nishizawa, Y., Shibuya, N., Thomma, B.P.H.J., and Talbot, N.J. (2012). Effector-Mediated Suppression of Chitin-Triggered Immunity by Magnaporthe oryzae Is Necessary for Rice Blast Disease. Plant Cell **24**: 322–335.

Mestre, P. and Baulcombe, D.C. (2006). Elicitor-mediated oligomerization of the tobacco N disease resistance protein. Plant Cell **18**: 491–501.

Meyers, B.C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R.W. (2003). Genome-wide analysis of NBS-LRR – encoding genes in Arabidopsis. Plant Cell **15**: 809–834.

Meyers, B.C., Morgante, M., and Michelmore, R.W. (2002). TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in Arabidopsisand other plant genomes. Plant J **32**: 77–92.

Michelmore, R.W., Christopoulou, M., and Caldwell, K.S. (2013). Impacts of resistance gene genetics, function, and evolution on a durable future. Annu. Rev. Phytopathol. **51**: 291–319.

Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H., and Shibuya, N. (2007). CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. **104**: 19613–8.

Mosquera, G., Giraldo, M.C., Khang, C.H., Coughlan, S., and Valent, B. (2009). Interaction transcriptome analysis identifies Magnaporthe oryzae BAS1-4 as Biotrophy-associated secreted proteins in rice blast disease. Plant Cell **21**: 1273–1290.

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.

Mueller, A.N., Ziemann, S., Treitschke, S., Amann, D., and Doehlemann, G. (2013). Compatibility in the Ustilago maydis-Maize Interaction Requires Inhibition of Host Cysteine Proteases by the Fungal Effector Pit2. PLoS Pathog. **9**.

Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., and Narusaka, Y. (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. Plant J. **60**: 218–226.

Nei, M. and Rooney, A.P. (2005). Concerted and birth-and-death evolution of multigene families. Annu. Rev. Genet. **39**:121–52.

Nemri, A., Saunders, D.G.O., Anderson, C., Upadhyaya, N.M., Win, J., Lawrence, G.J., Jones, D.A., Kamoun, S., Ellis, J.G., and Dodds, P.N. (2014). The genome sequence and effector complement of the flax rust pathogen Melampsora lini. Front. Plant Sci. **5**: 98.

Ntoukakis, V., Saur, I.M.L., Conlan, B., and Rathjen, J.P. (2014). The changing of the guard: The Pto/Prf receptor complex of tomato and pathogen recognition. Curr. Opin. Plant Biol. **20**: 69–74.

O'Connell, R.J. et al. (2012). Lifestyle transitions in plant pathogenic Colletotrichum fungi deciphered by genome and transcriptome analyses. Nat. Genet. **44**: 1060–5.

Van Ooijen, G., Mayr, G., Kasiem, M.M. a, Albrecht, M., Cornelissen, B.J.C., and Takken, F.L.W. (2008). Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. J. Exp. Bot. **59**: 1383–1397.

Oome, S. and Van den Ackerveken, G. (2014). Comparative and functional analysis of the widely occurring family of Nep1-like proteins. Mol. Plant. Microbe. Interact. **27**: 1–51.

Oome, S., Raaymakers, T.M., Cabral, A., Samwel, S., Böhm, H., Albert, I., Nürnberger, T., and Van den Ackerveken, G. (2014). Nep1-like proteins from three kingdoms of life act as a microbe-associated molecular pattern in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. **111**: 16955–60.

Orbach, M.J. (2000). A Telomeric Avirulence Gene Determines Efficacy for the Rice Blast Resistance Gene Pi-ta. Plant Cell Online **12**: 2019–2032.

Ottmann, C., Luberacki, B., Küfner, I., Koch, W., Brunner, F., Weyand, M., Mattinen, L., Pirhonen, M., Anderluh, G., Seitz, H.U., Nürnberger, T., and Oecking, C. (2009). A common toxin fold mediates microbial attack and plant defense. Proc. Natl. Acad. Sci. U. S. A. **106**: 10359–64.

Padmanabhan, M.S., Ma, S., Burch-Smith, T.M., Czymmek, K., Huijser, P., and Dinesh-Kumar, S.P. (2013). Novel Positive Regulatory Role for the SPL6 Transcription Factor in the N TIR-NB-LRR Receptor-Mediated Plant Innate Immunity. PLoS Pathog. **9**.

Panstruga, R. and Dodds, P.N. (2009). Terrific protein traffic: the mystery of effector protein delivery by filamentous plant pathogens. Science **324**: 748–50.

Park, C.H. et al. (2016). The E3 Ligase APIP10 Connects the Effector AvrPiz-t to the NLR Receptor Piz-t in Rice. PLoS Pathog. **12**: e1005529.

Park, C.-H., Chen, S., Shirsekar, G., Zhou, B., Khang, C.H., Songkumarn, P., Afzal, A.J., Ning, Y., Wang, R., Bellizzi, M., Valent, B., and Wang, G.-L. (2012). The Magnaporthe oryzae effector AvrPiz-t targets the RING E3 ubiquitin ligase APIP6 to suppress pathogen-associated molecular pattern-triggered immunity in rice. Plant Cell **24**: 4748–62.

Parker, J.E., Feys, B.J., van der Biezen, E. a, Noël, L., Aarts, N., Austin, M.J., Botella, M. a, Frost, L.N., Daniels, M.J., and Jones, J.D. (2000). Unravelling R gene-mediated disease resistance pathways in Arabidopsis. Mol. Plant Pathol. **1**: 17–24.

Peart, J.R., Mestre, P., Lu, R., Malcuit, I., and Baulcombe, D.C. (2005). NRG1, a CC-NB-LRR protein, together with N, a TIR-NB-LRR protein, mediates resistance against tobacco mosaic virus. Curr. Biol. **15**: 968–973.

Pedersen, C. et al. (2012). Structure and evolution of barley powdery mildew effector candidates. BMC genomics.**13:** 694

Pendleton, A.L., Smith, K.E., Feau, N., Martin, F.M., Grigoriev, I. V., Hamelin, R., Nelson, C.D., Burleigh, J.G., and Davis, J.M. (2014). Duplications and losses in gene families of rust pathogens highlight putative effectors. Front. Plant Sci. **5**: 299.

Pennisi, E. (2010). Armed and dangerous. Science **327**: 804–5.

Pfund, C., Tans-Kersten, J., Dunning, F.M., Alonso, J.M., Ecker, J.R., Allen, C., and Bent, A.F. (2004). Flagellin is not a major defense elicitor in Ralstonia solanacearum cells or extracts applied to Arabidopsis thaliana. Mol. Plant. Microbe. Interact. **17**: 696–706.

Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., and Kahmann, R. (2015). Fungal Effectors and Plant Susceptibility. Annu. Rev. Plant Biol. **66**: 513–545.

LPruitt, R.N. et al. (2015). The rice immune receptor XA21 recognizes a tyrosine-sulfated protein from a Gram-negative bacterium. Sci. Adv. **1**: e1500245.

Qi, D., DeYoung, B.J., and Innes, R.W. (2012). Structure-function analysis of the coiled-coil and leucine-rich repeat domains of the RPS5 disease resistance protein. Plant Physiol. **158**: 1819–32.

Qi, D. and Innes, R. (2013). Recent Advances in Plant NLR Structure, Function, Localization, and Signaling. Front. Immunol.: 221.

Rairdan, G.J., Collier, S.M., Sacco, M. a, Baldwin, T.T., Boettrich, T., and Moffett, P. (2008). The coiled-coil and nucleotide binding domains of the Potato Rx disease resistance protein function in pathogen recognition and signaling. Plant Cell **20**: 739–751.

Rao, Y., Li, Y., and Qian, Q. (2014). Recent progress on molecular breeding of rice in China. Plant Cell Rep. **33**: 551–564.

Rehmany, A.P. (2005). Differential Recognition of Highly Divergent Downy Mildew Avirulence Gene Alleles by RPP1 Resistance Genes from Two Arabidopsis Lines. Plant Cell **17**: 1839–1850.

Ribot, C., Hirsch, J., Balzergue, S., Tharreau, D., Nottéghem, J.L., Lebrun, M.H., and Morel, J.B. (2008). Susceptibility of rice to the blast fungus, Magnaporthe grisea. J. Plant Physiol. **165**: 114–124.

Robatzek, S., Bittel, P., Chinchilla, D., K??chner, P., Felix, G., Shiu, S.H., and Boller, T. (2007). Molecular identification and characterization of the tomato flagellin receptor LeFLS2, an orthologue of Arabidopsis FLS2 exhibiting characteristically different perception specificities. Plant Mol. Biol. **64**: 539–547.

Rooney, H., Klooster, J., van der Hoorn, R., Joosten, M., Jones, J., and de Wit, P. (2005). Cladosporium Avr2 Inhibits Tomato Rcr3 Protease Required for Cf-2 – Dependent Disease Resistance. Science **308**: 1783–1786.

Rouxel, T. et al. (2011). Effector diversification within compartments of the Leptosphaeria maculans genome affected by Repeat-Induced Point mutations. Nat. Commun. **2**: 202.

Rovenich, H., Boshoven, J.C., and Thomma, B.P.H.J. (2014). Filamentous pathogen effector functions: Of pathogens, hosts and microbiomes. Curr. Opin. Plant Biol. **20**: 96–103.

Saitoh, H. et al. (2012). Large-scale gene disruption in Magnaporthe oryzae identifies MC69, a secreted protein required for infection by monocot and dicot fungal pathogens. PLoS Pathog. **8**: e1002711.

Sánchez-Vallet, A., Saleem-Batcha, R., Kombrink, A., Hansen, G., Valkenburg, D.J., Thomma, B.P.H.J., and Mesters, J.R. (2013). Fungal effector Ecp6 outcompetes host immune receptor for chitin binding through intrachain LysM dimerization. Elife **2013**: 1–16.

Sarris, P.F. et al. (2015). A Plant Immune Receptor Detects Pathogen Effectors that Target WRKY Transcription Factors. Cell **161**: 1089–1100.

Sarris, P.F., Cevik, V., Dagdas, G., Jones, J.D.G., and Krasileva, K. V (2016). Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. BMC Biol. **14**: 8.

Saunders, D.G.O., Win, J., Cano, L.M., Szabo, L.J., Kamoun, S., and Raffaele, S. (2012). Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. PLoS One **7**.

Saur, I.M.-L., Conlan, B.F., and Rathjen, J.P. (2015). The N-Terminal Domain of the Tomato Immune Protein Prf Contains Multiple Homotypic and Pto Kinase Interaction Sites. J. Biol. Chem. **290**: jbc.M114.616532.

Schilling, L., Matei, A., Redkar, A., Walbot, V., and Doehlemann, G. (2014). Virulence of the maize smut Ustilago maydis is shaped by organ-specific effectors. Mol. Plant Pathol. **15**: 780–789.

Schornack, S., van Damme, M., Bozkurt, T.O., Cano, L.M., Smoker, M., Thines, M., Gaulin, E., Kamoun, S., and Huitema, E. (2010). Ancient class of translocated oomycete effectors targets the host nucleus. Proc. Natl. Acad. Sci. U. S. A. **107**: 17421–17426.

Schumann, F.H., Riepl, H., Maurer, T., Gronwald, W., Neidig, K.P., and Kalbitzer, H.R. (2007). Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. J. Biomol. NMR **39**: 275–289.

Segretin, M.E., Pais, M., Franceschetti, M., Chaparro-Garcia, A., Bos, J.I., Banfield, M.J., and Kamoun, S. (2014). Single amino acid mutations in the potato immune receptor R3a expand response to Phytophthora effectors. Mol. Plant. Microbe. Interact. **27**: 624–637.

Selin, C., de Kievit, T.R., Belmonte, M.F., and Fernando, W.G.D. (2016). Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. Front. Microbiol. **7**: 1–21.

Shao, F., Golstein, C., Ade, J., Stoutemyer, M., Dixon, J.E., and Innes, R.W. (2003). Cleavage of Arabidopsis PBS1 by a bacterial type III effector. Science **301**: 1230–1233.

Sharpee, W.C. and Dean, R.A. (2016). Form and function of fungal and oomycete effectors. Fungal Biol. Rev.

Shen, Q.H., Saijo, Y., Mauch, S., Biskup, C., Bieri, S., Keller, B., Seki, H., Ulker, B., Somssich, I.E., and Schulze-Lefert, P. (2007). Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. Science. **315**: 1098–1103.

Shimizu, T., Nakano, T., Takamizawa, D., Desaki, Y., Ishii-Minami, N., Nishizawa, Y., Minami, E., Okada, K., Yamane, H., Kaku, H., and Shibuya, N. (2010). Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. Plant J. **64**: 204–214.

Sinapidou, E., Williams, K., Nott, L., Bahkt, S., T??r, M., Crute, I., Bittner-Eddy, P., and Beynon, J. (2004). Two TIR:NB:LRR genes are required to specify resistance to Peronospora parasitica isolate Cala2 in Arabidopsis. Plant J. **38**: 898–909.

Slootweg, E. et al. (2010). Nucleocytoplasmic distribution is required for activation of resistance by the potato NB-LRR receptor Rx1 and is balanced by its functional domains. Plant Cell **22**: 4195–215.

Slootweg, E.J., Spiridon, L.N., Roosien, J., Butterbach, P., Pomp, R., Westerhof, L., Wilbers, R., Bakker, E., Bakker, J., Petrescu, A.-J., Smant, G., and Goverse, A. (2013). Structural determinants at the interface of the ARC2 and leucine-rich repeat domains control the activation of the plant immune receptors Rx1 and Gpa2. Plant Physiol. **162**: 1510–28.

Song, J., Win, J., Tian, M., Schornack, S., Kaschani, F., Ilyas, M., van der Hoorn, R. a L., and Kamoun, S. (2009). Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3. Proc. Natl. Acad. Sci. U. S. A. **106**: 1654–1659.

Soyer, J.L., El Ghalid, M., Glaser, N., Ollivier, B., Linglin, J., Grandaubert, J., Balesdent, M.H., Connolly, L.R., Freitag, M., Rouxel, T., and Fudal, I. (2014). Epigenetic Control of Effector Gene Expression in the Plant Pathogenic Fungus Leptosphaeria maculans. PLoS Genet. **10**.

Spanu, P.D. et al. (2010). Genome Expansion and Gene Loss in. Science (80-.). **330**: 1543–1546.

Sperschneider, J., Gardiner, D.M., Dodds, P.N., Tini, F., Covarelli, L., Singh, K.B., Manners, J.M., and Taylor, J.M. (2015). Methods E FFECTOR P : predicting fungal effector proteins from secretomes using machine learning.New Phytologist.

Steinbrenner, A.D., Goritschnig, S., and Staskawicz, B.J. (2015). Recognition and Activation Domains Contribute to Allele-Specific Responses of an Arabidopsis NLR Receptor to an Oomycete Effector Protein. PLOS Pathog. **11**: e1004665.

Stergiopoulos, I., Kourmpetis, Y.A.I., Slot, J.C., Bakker, F.T., De Wit, P.J.G.M., and Rokas, A. (2012). In silico characterization and molecular evolutionary analysis of a novel superfamily of fungal effector proteins. Mol. Biol. Evol. **29**: 3371–3384.

Stergiopoulos, I. and de Wit, P.J.G.M. (2009). Fungal effector proteins. Annu. Rev. Phytopathol. **47**: 233–63.

Sukarta, O.C.A., Slootweg, E.J., and Goverse, A. (2016). Structure-Informed Insights for Nlr Functioning in Plant Immunity. Semin. Cell Dev. Biol. **56**: 134–149.

Sun, W., Dunning, F.M., Pfund, C., Weingarten, R., and Bent, A.F. (2006). Within-species flagellin polymorphism in Xanthomonas campestris pv campestris and its impact on elicitation of Arabidopsis FLAGELLIN SENSING2-dependent defenses. Plant Cell **18**: 764–79.

Swiderski, M.R., Birker, D., Jones, J.D.G., Centre, J.I., and Colney, P. (2009). The TIR Domain of TIR-NB-LRR Resistance Proteins Is a Signaling Domain Involved in Cell Death Induction. Mol. Plant. Microbe. Interact. **22**: 157–165.

Takemoto, D., Rafiqi, M., Hurley, U., Lawrence, G.J., Bernoux, M., Hardham, A.R., Ellis, J.G., Dodds, P.N., and Jones, D. a (2012). N-terminal motifs in some plant disease resistance proteins function in membrane attachment and contribute to disease resistance. Mol. Plant. Microbe. Interact. **25**: 379–92.

Takken, F.L.W. and Goverse, A. (2012). How to build a pathogen detector: Structural basis of NB-LRR function. Curr. Opin. Plant Biol. **15**: 375–384.

Takken, F.L.W. and Tameling, W.I.L. (2009). To nibble at plant resistance proteins. Science **324**: 744–746.

Talbot, N.J. and Kershaw, M.J. (2009). The emerging role of autophagy in plant pathogen attack and host defence. Curr. Opin. Plant Biol. **12**: 444–50.

Tameling, W.I.L., Nooijen, C., Ludwig, N., Boter, M., Slootweg, E., Goverse, A., Shirasu, K., and Joosten, M.H. a J. (2010). RanGAP2 mediates nucleocytoplasmic partitioning of the NB-LRR immune receptor Rx in the Solanaceae, thereby dictating Rx function. Plant Cell **22**: 4176–4194.

Tanaka, S., Brefort, T., Neidig, N., Djamei, A., Kahnt, J., Vermerris, W., Koenig, S., Feussner, K., Feussner, I., and Kahmann, R. (2014). A secreted Ustilago maydis effector promotes virulence by targeting anthocyanin biosynthesis in maize. Elife **3**: e01355.

Tenthorey, J.L., Kofoed, E.M., Daugherty, M.D., Malik, H., and Vance, R.E. (2014). Molecular Basis for Specific Recognition of Bacterial Ligands by NAIP/NLRC4 Inflammasomes. Mol. Cell **54**: 17–29.

Tharreau, D., Fudal, I., Andriantsimialona, D., Santoso, Utami, D., Fournier, E., Lebrun, M.-H., and Nottéghem, J.-L. (2009). World population structure and migration of the rice blast fungus, Magnaporthe oryzae. In Advances in genetics genomics and control of rice blast disease, pp. 209–215.

Thines, M. (2014). Gene loss rather than gene gain is associated with a host jump from monocots to dicots in the Smut Fungus Melanopsichium pennsylvanicum. Genome Biol. Evol. **6**: 2034–2049.

Thomas A. Mentlak, Nicholas J. Talbot, and T.K. (2012). Fungus, Effector Translocation and Delivery by the Rice Blast Oryzae, Magnaporthe. Eff. Plant Microbe Interact.

Thomma, B.P.H.J., Nürnberger, T., and Joosten, M.H. a J. (2011). Of PAMPs and effectors: the blurred PTI-ETI dichotomy. Plant Cell **23**: 4–15.

Torres, M.A., Jones, J.D.G., and Dangl, J.L. (2006). Reactive oxygen species signaling in response to pathogens. Plant Physiol. **141**: 373–378.

Ueda, H., Yamaguchi, Y., and Sano, H. (2006). Direct interaction between the tobacco mosaic virus helicase domain and the ATP-bound resistance protein, N factor during the hypersensitive response in tobacco plants. Plant Mol. Biol. **61**: 31–45.

Valent, B. and Khang, C.H. (2010). Recent advances in rice blast effector research. Curr. Opin. Plant Biol. **13**: 434–41.

VanLoon, L.C. (1997). Induced resistance in plants and the role of pathogenesis-related proteins. Eur. J. Plant Pathol. **103**: 753–765.

Vinatzer, B. a, Monteil, C.L., and Clarke, C.R. (2014). Harnessing population genomics to understand how bacterial pathogens emerge, adapt to crop hosts, and disseminate. Annu. Rev. Phytopathol. **52**: 19–43.

Vleeshouwers, V.G. a a et al. (2008). Effector genomics accelerates discovery and functional profiling of potato disease resistance and phytophthora infestans avirulence genes. PLoS One **3**: e2875.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). nucleotide binding fold Figure IA). Approximately 220 / of amino acid residues along that useATP Homology between a and (3 and otherenzymes. EMBO J. **I**: 945–951.

Wan, J., Zhang, X.-C., Neece, D., Ramonell, K.M., Clough, S., Kim, S.-Y., Stacey, M.G., and Stacey, G. (2008). A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. Plant Cell **20**: 471–481.

Wang, C.-I.A.C.-I. a 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**: 2898–912.

Wang, G., Roux, B., Zhou, J., Noe, L.D., Wang, G., Roux, B., Feng, F., Guy, E., Li, L., Li, N., Zhang, X., and Lautier, M. (2015). The Decoy Substrate of a Pathogen Effector and a Pseudokinase Specify Pathogen-Induced Modified- Self Recognition and Immunity in Plants. Cell Host Microbe. **18:** 285–295.

Weaver, L.M., Swiderski, M.R., Li, Y., Jones, J.D.G., Centre, J.I., Colney, P., and Nr, N. (2006). The Arabidopsis thaliana TIR-NB-LRR R-protein , RPP1A ; protein localization and constitutive activation of defence by truncated alleles in tobacco and Arabidopsis Plant J. **2**: 829–840.

William Sharpee, Y. and Macchi, Mihwa Yi, William Franck, Alex Eyre, Laura Okagaki, B.V. and R.D. (2016). Identification and Characterization of Suppressors of Plant Cell Death (SPD) Genes from Magnaporthe oryzae. Mol. Plant Pathol.

Williams, S.J. et al. (2014). Structural basis for assembly and function of a heterodimeric plant immune receptor. Science **344**: 299–303.

Williams, S.J., Sornaraj, P., DeCourcy-Ireland, E., Menz, R.I., Kobe, B., Ellis, J.G., Dodds, P.N., and Anderson, P.A. (2011). An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. Mol. Plant-Microbe Interact. **24**: 897–906.

Willmann, R., Lajunen, H.M., Erbs, G., Newman, M., Kolb, D., and Tsuda, K. (2011). Mediate Bacterial Peptidoglycan Sensing and Immunity To Bacterial Infection. Proc. Natl. Acad. Sci. **108**: 19824–19829.

Win, J., Bos, J.I.B., Liu, H., Damme, M. Van, Morgan, W., Choi, D., Vossen, E.A.G. Van Der, Vleeshouwers, V.G.A.A., and Kamoun, S. (2009). In Planta Expression Screens of Phytophthora infestans RXLR Effectors Reveal Diverse Phenotypes , Including Activation of the Solanum bulbocastanum Disease Resistance Protein Rpi-blb2. Plant Cell. **21**: 2928–2947.

Win, J., Krasileva, K. V., Kamoun, S., Shirasu, K., Staskawicz, B.J., and Banfield, M.J. (2012). Sequence divergent RXLR effectors share a structural fold conserved across plant pathogenic oomycete species. PLoS Pathog. **8**: 8–11.

Wirthmueller, L., Maqbool, A., and Banfield, M.J. (2013). On the front line : structural insights into plant – pathogen interactions. Nature reviews. Microbiology. **11**: 761–776.

Wirthmueller, L., Zhang, Y., Jones, J.D.G., and Parker, J.E. (2007). Nuclear Accumulation of the Arabidopsis Immune Receptor RPS4 Is Necessary for Triggering EDS1-Dependent Defense. Curr. Biol. **17**: 2023–2029.

Wit, P, Mehrabi, R., Burg, H., and Stergiopoulos, I. (2009). Fungal effector proteins : past , present and future. Mol. Plant Path.**10**: 735–747.

Wouw, A. and Howlett, B.J. (2011). Fungal pathogenicity genes in the age of " omics ." Mol. Plant Path. **12**: 507–514.

Yang, D.L., Yang, Y., and He, Z. (2013a). Roles of plant hormones and their interplay in rice immunity. Mol. Plant **6**: 675–685.

Yang, S., Feng, Z., Zhang, X., Jiang, K., Jin, X., Hang, Y., Chen, J.Q., and Tian, D. (2006). Genome-wide investigation on the genetic variations of rice disease resistance genes. Plant Mol. Biol. **62**: 181–193.

Yang, S., Li, J., Zhang, X., Zhang, Q., Huang, J., Chen, J.-Q., Hartl, D.L., and Tian, D. (2013b). Rapidly evolving R genes in diverse grass species confer resistance to rice blast disease. Proc. Natl. Acad. Sci. U. S. A. **110**: 18572–7.

Yoshida, K., Saitoh, H., Fujisawa, S., Kanzaki, H., Matsumura, H., Yoshida, K., Tosa, Y., Chuma, I., Takano, Y., Win, J., Kamoun, S., and Terauchi, R. (2009). Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen Magnaporthe oryzae. Plant Cell. **21**: 1573–91.

Yoshida, K., Saunders, D.G.O., Mitsuoka, C., Natsume, S., Kosugi, S., Saitoh, H., Inoue, Y., Chuma, I., Tosa, Y., Cano, L.M., Kamoun, S., and Terauchi, R. (2016). Host specialization of the blast fungus Magnaporthe oryzae is associated with dynamic gain and loss of genes linked to transposable elements. BMC Genomics **17**: 370.

Zhong, Z. et al. (2016). Directional Selection from Host Plants Is a Major Force Driving Host Specificity in Magnaporthe Species. Sci. Rep. **6**: 25591.

Zipfel, C. (2008). Pattern-recognition receptors in plant innate immunity. Curr. Opin. Immunol. **20**: 10–6.

Zou, H., Henzel, W.J., Liu, X., Lutschg, A., and Wang, X. (1997). Apaf-1, a human protein homologous to c.elegans CED-4, participates in cytochrome c dependent activation of caspace 3. Cell **90**: 405–413.

RESUME

Etude des bases moléculaires de la reconnaissance de l'effecteur fongique AVR-Pia par le récepteur immunitaire du riz RGA5

Mots clés: Récepteurs immunitaires de type NLR, Immunité des plantes, Effecteurs, Magnaporthe oryzae, Riz

Les maladies des plantes causées par les champignons sont un problème majeur en agriculture. Pour les contrôler, les gènes de résistance (*R*) qui permettent de développer des variétés de plantes résistantes sont des éléments clés. La majorité des gènes *R* codent pour des protéines NLRs caractérisées par la présence d'un domaine de liaison aux nucléotides (NB-ARC) et un domaine de répétitions riches en leucines (LRR). Ces protéines agissent comme des récepteurs immunitaires intracellulaires et reconnaissent des facteurs de virulence des agents pathogènes appelés effecteurs. Les champignons phytopathogènes possèdent de vastes répertoires d'effecteurs qui contiennent des centaines de protéines sécrétés, de petites tailles et sans similarités de séquence entre elles.

La première question abordée dans ma thèse concerne l'origine de l'immense diversité des effecteurs fongiques. Une analyse structurale a identifié une famille d'effecteurs de séquences différentes mais qui possèdent une structure conservée. Cette famille a été appelée MAX-effectors (***M****agnaporthe* **A**vrs and To**x**B like) et elle est particulièrement importante chez *Magnaporthe oryzae*, l'agent causal de la pyriculariose du riz. Par des analyses d'expression, j'ai confirmé que la majorité des effecteurs MAX de *M. oryzae* sont spécifiquement exprimés durant la phase précoce de l'infection, suggérant une fonction importante durant la colonisation de la plante. Les effecteurs MAX constituent la première famille d'effecteurs fongiques définis par leur structure. Cette étude apporte donc de nouvelles pistes pour l'identification d'effecteurs chez les champignons et contribue à une meilleure compréhension de l'évolution des effecteurs. En effet, le scénario observé chez les effecteurs MAX suggère que beaucoup d'effecteurs fongiques appartiennent à un nombre restreint de familles d'effecteurs définies par leur structure.

La seconde question que j'ai abordée durant ma thèse est le mécanisme moléculaire de la reconnaissance des effecteurs par les NLRs. J'ai abordé cette question en étudiant la reconnaissance de l'effecteur AVR-Pia par le couple de NLRs RGA4/RGA5. Des travaux précédents ont montré que RGA5 agit comme récepteur et se lie directement à AVR-Pia tandis que RGA4 agit comme élément de signalisation constitutivement actif, qui, en absence de l'agent pathogène, est réprimé par RGA5. Un domaine de RGA5, normalement absent chez les protéines NLR et similaire à la chaperonne du cuivre ATX1 (domaine RATX1), interagit physiquement avec AVR-Pia. Il a été suggéré que ce domaine RATX1 puisse agir comme un leurre de la cible de virulence d'AVR-Pia. Ce leurre, intégré dans la structure de RGA5, permettrait de « piéger » l'effecteur par interaction directe et jouerait donc un rôle crucial dans sa reconnaissance spécifique. Grâce à une analyse structurale détaillée d'AVR-Pia j'ai pu confirmer le rôle central de l'interaction AVR-Pia-RATX1 dans la reconnaissance de cet effecteur ce qui conforte le modèle du « leurre intégré ». De plus, j'ai caractérisé la

surface d'interaction avec laquelle AVR-Pia lie le domaine RATX1. De plus, j'ai détecté des interactions entre AVR-Pia et d'autres parties de RGA5, indépendantes du domaine RATX1, notamment les domaines NB-ARC et LRR. Ceci a permis de développer un modèle qui explique comment la liaison d'un effecteur à un récepteur NLR comportant un leurre intégré par différentes interactions indépendantes conduit à une reconnaissance très sensible et spécifique qui est peu affectée par des mutations ponctuelles de l'effecteur. En résumé, cette étude a produit des connaissances nouvelles sur la fonction des récepteurs des plantes de type NLRs et sur leur capacité à reconnaitre des effecteurs. Ceci contribue à une meilleure compréhension du système immunitaire des plantes, ce qui est un élément important pour l'obtention de cultures durablement résistantes aux maladies.

SUMMARY

Functional analysis of the AVR-Pia fungal effector recognition by the rice immune receptor RGA5

Key words: NLR immune receptors, Plant immunity, Effectors, Magnaporthe oryzae, Rice

Plant diseases caused by fungi constitute a worldwide threat to food security and disease resistance (*R*) genes that allow to breed resistant crops are key elements for efficient disease control. The vast majority of *R* genes code for NLR multi domain proteins characterized by nucleotide-binding and leucine-rich repeat domains and acting as intracellular immune receptors for pathogen-secreted virulence factors termed effectors. Phytopathogenic fungi possess huge effector repertoires that are dominated by hundreds of sequence-unrelated small secreted proteins.

The first question I addressed in my PhD thesis is: how is the tremendous diversity of fungal effectors generated? A structural analysis had identified the family of sequence-unrelated but structurally conserved MAX-effectors (***M****agnaporthe* **A**vrs and To**x**B like) that has expanded specifically in *Magnaporthe oryzae* the causal agent of rice blast disease. By expression analysis, I confirmed that the majority of *M. oryzae* MAX-effectors are expressed specifically during early infection suggesting important functions during host colonization. MAX effectors are the first structurally defined family of effectors in fungi and this study gives therefore news clues for the identification of candidate effectors in fungi and constitutes a crucial step towards a better understanding of effector evolution. In fact, the scenario observed for MAX-effectors leads to the hypothesis that the enormous number of sequence-unrelated fungal effectors belong in fact to a restricted set of structurally conserved effector families.

The second question I investigated in my PhD thesis is: what are the molecular mechanisms of effector recognition by NLR immune receptors? I addressed this question by studying recognition of the *M. oryzae* effector AVR-Pia by the rice NLR pair RGA4/RGA5. Previous work has shown that RGA5 acts as a receptor that binds directly to AVR-Pia while RGA4 acts as a constitutively active signaling protein that is, in the absence of pathogen, repressed by RGA5. This functional interaction involves formation of an RGA4/RGA5 receptor complex. By protein-protein interaction studies, I showed that complex formation involves interactions between the RA4 and RGA5 NB-ARC and LRR domains, in addition to previously identified interactions between the coiled-coil domains. AVR-Pia recognition seems not to induce dissociation of the RGA4/RGA5 complex but a ternary RGA4/RGA5/AVR-Pia complex could also not be detected consistently. How effector recognition is translated into receptor complex activation remains therefore to be elucidated in more detail in the future.

Previous work has shown that a domain of RGA5 normally not present in NLRs and related to the copper chaperone ATX1 (RATX1 domain) interacts physically with AVR-Pia and may be crucial for effector recognition. The RATX1 domain was hypothesized to mimic the true host targets of AVR-Pia leading to the development of the 'integrated decoy' model that states

that unconventional domains in NLRs act as decoys in the recognition of effector proteins. By detailed structure-informed analysis of AVR-Pia, I could confirm the pivotal role of the AVR-Pia-RATX1 interaction for effector recognition lending important support to the integrated decoy model. In addition, I could precisely characterize the interaction surface with which AVR-Pia binds to the RGA5 RATX1 domain. Finally, I detected interactions of AVR-Pia with other parts of RGA5, in particular the NB-ARC and the LRR domains. Based on these results, I developed a model that explains how such binding to several independent sites in NLRs leads to high overall affinity and robust effector recognition that is resilient to effector mutations. Taken together, this study provides important novel insight into NLR function and effector recognition and contributes by this to a better understanding of plant immunity which is crucial for generating durable disease resistance in crops.