The role and evolution of fungal effectors in plant pathogenesis

Ronnie de Jonge

Thesis committee

Promoter
Prof. dr. ir. P.J.G.M. de Wit
Professor of Phytopathology
Wageningen University

Co-promoter
Dr. ir. B.P.H.J. Thomma
Associate professor, Laboratory of Phytopathology
Wageningen University

Other members
Prof. dr. J.H.S.G.M. de Jong, Wageningen University
Prof. dr. S.C. de Vries, Wageningen University
Prof. dr. B.J.C. Cornelissen, University of Amsterdam
Prof. dr. W.J. Stiekema, University of Amsterdam

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences

The role and evolution of fungal effectors in plant pathogenesis

Ronnie de Jonge

Thesis

submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 9 November 2012
at 4 p.m. in the Aula

Ronnie de Jonge
The role and evolution of fungal effectors in plant pathogenesis,
148 pages

PhD thesis, Wageningen University, Wageningen, NL (2012)
With references, with summaries in Dutch and English

ISBN 978-94-6173-391-7

Table of contents

Chapter 1

General introduction and outline of the thesis

de Jonge R, Bolton MD and Thomma BPHJ (2011) **How filamentous pathogens co-opt plants: the ins and outs of fungal effectors.** *Curr Opin Plant Biology* **14**: 400-406

General introduction

Abstract

Research on effectors secreted by pathogens during host attack has dominated the field of molecular plant–microbe interactions over recent years. Functional analysis of type III secreted effectors injected by pathogenic bacteria into host cells has significantly advanced the field and demonstrated that many function to suppress host defense. Fungal and oomycete effectors are delivered outside the host plasma membrane, and although research has lagged behind on bacterial effectors, we are gradually learning more and more about the functions of these effectors. While some function outside the host cell to disarm defense, others exploit host cellular uptake mechanisms to suppress defense or liberate nutrients intracellularly. Comparative genomics suggests that the organization of effector genes drives effector evolution in many pathogen genomes.

Introduction

Inheritance of plant immunity to pathogens is controlled by corresponding gene pairs, as plants carry resistance (*R*) genes that interact with pathogen avirulence (*Avr*) genes in a gene-for-gene manner. Since direct interaction between R and Avr proteins could often not be demonstrated experimentally, it was recognized that R proteins may also monitor the state of host components targeted by pathogen Avr molecules to establish disease. Presently, the term 'effector' is commonly used for these molecules [1]. Similar morphological growth characteristics, virulence mechanisms, and infection strategies are generally shared in the taxonomically distinct fungi and oomycetes, despite differences in physiology, biochemistry, and genetics. Both types of pathogens target effectors to the apoplast or cytoplasm where they function to modulate host physiology, often through suppression of host defenses, or to protect the pathogen from host defense responses employed to halt pathogen growth. In this review, we focus on recent progress in research on the function and evolution of effectors from filamentous plant pathogens, guided by the consecutive stages occurring during disease establishment (Figure 1).

Effector production

Fungal effector genes are typically not, or lowly, expressed in axenic cultures, but are induced upon host colonization. Since some effector genes are induced by nitrogen starvation *in vitro*, nitrogen limitation was proposed as an *in planta* trigger of their induction. However, nitrogen availability may not be limited in plants, and many *in planta*-induced effector genes do not respond to nitrogen deprivation [2]. Thus, the *in planta* signals that trigger induction of effector genes presently remain largely unknown.

Transcriptional regulators important for early infection stages were recently identified. In the root invading fungus *Fusarium oxysporum* f. sp. *lycopersici* (*Fol*), the transcriptional regulator Sge1 is required for *in planta* expression of various effector genes [3]. Interestingly, *SGE1* orthologs occur widely in fungi and include master regulators of morphological switching in dimorphic fungi. Recently, the *Magnaporthe*

oryzae zinc finger transcription factor MoCRZ1 was found to regulate various virulence factors [4]. *MoCRZ1* is important for virulence on rice, and homologs were identified as pathogenicity regulators in various fungi [5,6]. Intriguingly, MoCRZ1 also regulates genes involved in vesicle-mediated secretion, potentially implicating MoCRZ1 in effector secretion [4].

Recently it was elegantly demonstrated that pathogens may tailor their effectors to individual host tissues. A gene expression study in *Ustilago maydis*-infected maize tissues revealed differential timing and organ-specific expression of particular effector proteins. Subsequent inactivation of effector clusters revealed differential impact on pathogenicity in various plant tissues. The data suggest that *U. maydis* employs universal effectors for establishment of host compatibility, followed by deployment of effectors with organ-specific properties to redirect physiology [7].

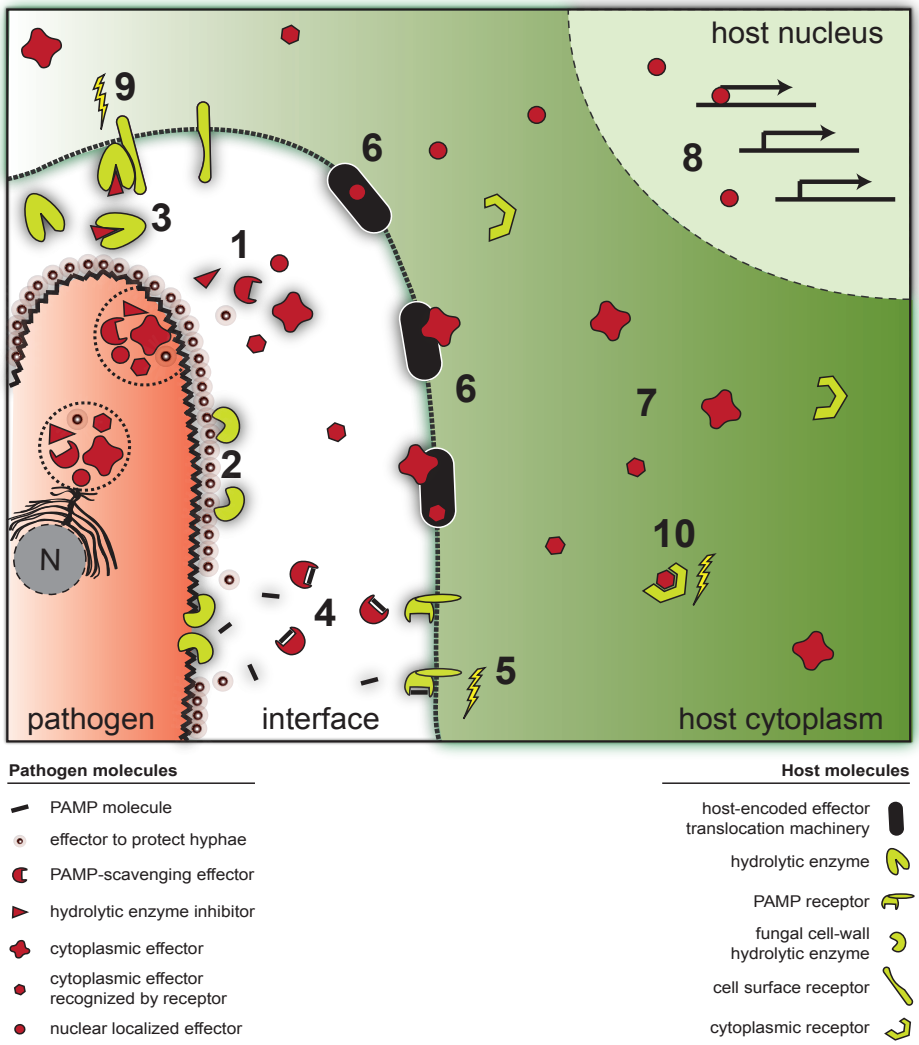


Figure 1. The role of effectors in the interactions between fungi and their host plants. Fungi secrete effectors in the interface between pathogen and host after host penetration (1). Some effectors contribute to fungal virulence by shielding hyphae against hydrolytic host defense enzymes in the host–pathogen interface (2), by inactivating these enzymes (3), or by scavenging potential PAMP molecules (4) that may alarm host defense (5). Many effectors do not remain in the host–pathogen interface but are translocated to the host cytoplasm without the use of pathogen-encoded translocation machinery (6). Although the molecular mechanism explaining how translocated effectors contribute to fungal virulence largely remains obscure, some of them are expected to affect cytoplasmic processes related to host defense (7). Recent evidence suggests some effectors are translocated to the nucleus where they may regulate transcription of target genes (8). Host recognition of filamentous pathogen effectors occurs through cell surface receptors in the host–pathogen interface (9), or in the host cytoplasm through NB-LRR-type receptors (10).

Effector delivery

Filamentous pathogen effector proteins are typically produced in the endoplasmic reticulum and secreted through Golgi-derived vesicles. While fungi limited to the extracellular space presumably secrete effectors mainly at hyphal tips, rusts, downy, and powdery mildews deliver their effectors via haustoria [8]. These structures were originally only considered as feeding structures that invaginate the host plasma membrane and are surrounded by an extracellular matrix. A remarkable structure for effector delivery is described as the biotrophic interfacial complex of *M. oryzae*. Upon penetration of rice cells by this fungus, the penetration peg differentiates into a primary hypha that invaginates the host plasma membrane [9]. At the tip of the entering hypha, effectors are secreted where the biotrophic interfacial complex develops. Subsequently, the hypha differentiates into a bulbous pseudohypha and continues to grow into neighboring plant cells while the biotrophic interfacial complex remains at the same position and delivers newly synthesized effectors [10].

Effectors with apoplastic functions

Cell-wall-degrading enzymes (CWDEs) are relatively well-characterized apoplastic effectors. Comparative genomics demonstrates that CWDE catalogs differ significantly between fungal pathogens [11,12]. Although in several fungi the sucrose nonfermenting 1 protein (SNF1) regulates CWDE expression and *SNF1* mutants display impaired virulence, functional redundancy complicates investigations into the contribution of individual CWDEs in virulence [13]. As CWDEs are also produced by saprophytic fungi, they are likely recruited as pathogenicity factors in pathogenic species that evolved from saprophytes, but do not determine host range or host specificity.

A second group of apoplastic effectors are the necrosis and ethylene-inducing protein (NEP1)-like proteins (NLPs) present in many pathogenic bacteria, fungi, and oomycetes that generally induce cell death in dicotyledonous plants through plasma membrane permeabilization [14]. Curiously, pathogens of monocotyledonous plants also carry *NLP* genes, but their role in pathogenicity remains obscure since they do not elicit necrosis. For example, heterologous expression of the single *NLP* gene from *Mycosphaerella graminicola* (MgNLP) did not induce cell death or elicit immune responses in wheat leaves, and gene knockouts did not affect virulence on wheat. However, MgNLP induced cell death in Arabidopsis leaves [15].

Perhaps the most intriguing apoplastic effectors are generally referred to as small cysteine-rich secreted proteins with unknown function. These effectors are generally species-specific or even isolate-specific. For few, their role in virulence has recently been elucidated. Several of these effectors from *Cladosporium fulvum*, but also from the oomycete *Phytophthora infestans*, have been characterized as inhibitors of extracellular host proteases important for basal defense [16,17,18]. Others appear to play key roles in protecting the fungus from chitin-triggered host defenses [19,20]. Plants produce apoplastic exochitinases that are not detrimental to fungal growth, but release chitin oligosaccharides from fungal cell walls that act as recognition patterns for host defense receptors. The rice lysin motif (LysM)-containing chitin oligosaccharide elicitor-binding protein (CEBiP) was characterized that, together with the LysM-

containing chitin elicitor receptor kinase-1 (OsCERK1), is required for chitin-triggered immune responses [21,22]. These responses include vacuolar accumulation of basic endochitinases that act as powerful antifungal agents once they are released. Orthologous chitin receptors are found in other plant species, including Arabidopsis [23]. Upon stomatal entry, *C. fulvum* secretes a repertoire of effector proteins that include the chitin-binding effectors Avr4 and Ecp6. Avr4 contributes to virulence by binding to fungal cell walls through an invertebrate chitin-binding domain in order to protect hyphae from host chitinases [19]. In contrast, Ecp6 sequesters chitin oligosaccharides through its LysM domains in order to prevent the activation of plant immune receptors [20]. Ecp6 homologous LysM effectors widely occur in fungi [24,25], suggesting that scavenging of chitin oligosaccharides is a conserved strategy of fungal pathogens to avoid detection [20]. Interestingly, although the secretion of chitinases by the plant is a widespread strategy in antimicrobial defense, Avr4 homologs appear restricted to only few *C. fulvum*-related fungi [26]. Possibly, in other pathogens LysM effectors may also be able to protect fungal hyphae against plant chitinases [25].

Effector uptake into host cells

Although effectors are delivered apoplastically, many appear to be subsequently translocated into the host cytoplasm. Initial evidence for cytoplasmic translocation derives from cytoplasmic R proteins that recognize fungal effectors. Flax rust (*Melampsora lini*) Avr effectors induced cell death in plants containing cytoplasmic R proteins, and direct interaction between the effectors and corresponding R proteins was demonstrated [27]. Recently, host cell internalization of haustorial effectors in the absence of *M. lini* was demonstrated, showing that pathogen-encoded components are not required for translocation [28]. Cytoplasmic recognition of effectors occurs in host cells for other pathogens also [1,8]. Interestingly, some *M. oryzae* biotrophic interfacial complex-secreted effectors autonomously move from the cytoplasm of invaded cells into neighboring cells, possibly preparing these for fungal invasion [10].

Ground-breaking work on a possible mechanism of effector uptake was recently reported [29]. Many predicted oomycete effectors contain an N-terminal RxLR motif [30] that was proposed to mediate autonomous effector uptake [31,32]. It is proposed that RxLR motifs enable oomycete effectors to bind to host cell surface phosphatidylinositol-3-phosphate (PI3P) and subsequently enter host cells though vesicle-mediated endocytosis [29]. Similarly, the N-termini of various fungal effectors were reported to carry degenerate RxLR motifs that bind to PI3P and mediate effector translocation, although this may not be a universal means of effector uptake [29,33]. Furthermore, whether effector uptake mediated by PI3P binding is functionally involved in the physiology of plant infection by fungi and oomycetes presently remains unknown. Conceivably, lipid-targeting may be one of several means for effectors to enter host cells since different uptake mechanisms are likely to exist to prevent hosts from intercepting effector trafficking.

Evidence for another conserved oomycete host translocation motif was provided for crinkler effectors, many of which appear to be targeted to the host nucleus [34]. Furthermore, powdery mildew and rust fungi encode small secreted proteins that share an N-terminal Y/F/WxC motif that is not found in effectors from non-haustorial fungi or oomycetes, and it is tempting to speculate that this motif mediates

translocation of fungal haustorial effectors [35].

Effectors with cytoplasmic functions

In contrast to many bacterial type III effectors that suppress host defense responses [1], the function of few cytoplasmic fungal effectors has been elucidated. Houterman *et al*. showed that the *Fol* effector Avr1 (Six4) suppresses resistance mediated by the tomato cytoplasmic R protein I-2 [36]. Recently, 'SWEET' sugar efflux transporters were identified in plants [37]. Several pathogens, including fungi with diverse feeding styles, induce expression of distinct *SWEET* genes, and *SWEET* induction by pathogenic bacteria was type III secretion dependent. Moreover, direct binding of a type III effector to a *SWEET* promoter was demonstrated, suggesting that sugar efflux is hijacked by cytoplasmic pathogen effectors in order to release nutrients [37].

Previous studies identified fungal hexose transporters in obligate biotrophs that were assumed to act in concert with fungal cell-wall-derived invertases to take up glucose or fructose after sucrose hydrolysis [38]. However, a recent study identified a plasma membrane-localized sucrose transporter in the smut fungus *U. maydis* that is specifically produced during plant infection, required for virulence, and able to outcompete plant transporters. In this way, *U. maydis* can utilize sucrose without prior extracellular hydrolysis by invertases [39]. Direct utilization of sucrose circumvents invertase-induced changes in apoplastic glucose concentrations known to induce defense [38].

Effector evolution

Effector genes are frequently under selection pressure, illustrating the coevolutionary arms race between host and pathogen [40-45]. They are often located at genomic sites that promote evolution through mutation or recombination. Comparative genomics among *Aspergillus* spp. revealed the accumulation of species-specific genes in chromosomal islands enriched for transposons [46]. Tomato pathogenic *Fol* strains contain a transposon-enriched pathogenicity chromosome that can be exchanged between isolates [12]. The extreme impact of transposons is illustrated in the size-expanded genomes of obligate powdery mildew pathogens that are largely composed of transposons. These pathogens lost many genes that are dispensable for obligate biotrophy, likely explaining why they can no longer grow in the absence of their host [11]. Intriguingly, of the ~250 effector genes identified in barley powdery mildew, only a handful are shared with pea and Arabidopsis powdery mildews, illustrating extreme host adaptation [11].

Repeat-induced point mutation (RIP) is a fungal defense mechanism to protect genomes against transposable elements by accumulating mutations in repetitive DNA. In a large-scale study of *Leptosphaeria maculans* field isolates, a transposon-enriched cluster of effector genes was found to be degenerated by RIP, presumably as consequence of imposed selection pressure through the introduction of resistant canola varieties with matching R genes [41]. Intriguingly, one-third of the *L. maculans* genome is composed of AT-rich blocks that contain effector genes and transposons that are both affected by RIP [42]. Taken together, transposon and RIP activity orchestrate rapid effector diversification, and aid in the rapid generation of effector variants that escape host recognition [42]. Transposon activity appears to play an important role in

effector evolution in oomycete genomes as well [43].

Effector discovery

Pathogen effector catalogs are highly lineage-specific and determination of effector catalogs is a challenge. Typical effector calling based on the presence of signal peptides and absence of transmembrane domains has resulted in the prediction of catalogs that often contain up to hundreds of potential effectors for individual pathogens. To enhance prediction accuracy, secreted protein prediction pipelines have been developed that combine different algorithms [47,48]. However, more sophisticated methods are required to identify the most relevant effectors for disease establishment within large effector catalogs. A rather obvious criterion is whether candidate effector expression *in planta* can be detected [11,35,44]. Furthermore, several studies have now shown that signatures of positive selection can be used to pinpoint candidate effector genes in sequenced genomes [40-45]. Comparative genomics on the related maize pathogens *U. maydis* and *Sporisorium reilianum* identified regions of low sequence conservation that primarily encode clusters of secreted effectors in otherwise well-conserved syntenic genomes. Interestingly, this effector differentiation suggests that the two maize pathogens target different host molecules. Furthermore, mutational analysis of several effector clusters confirmed their role in virulence [49].

Effector recognition

As discussed above, successful pathogens exploit effectors to subvert their hosts, resulting in effector-triggered susceptibility (ETS). Plants have responded by evolving R proteins that recognize effectors and activate effector-triggered immunity (ETI). Necrotrophic fungal pathogens were considered rather nonspecific in their host attack. However, many necrotrophic pathogens evolved mechanisms to attack plants in sophisticated ways, even exploiting host resistance mechanisms [50,51,52]. Since various effectors (toxins) interact with disease resistance protein analogs, it is now suggested that necrotrophic pathogens deliberately activate host ETI responses directed against biotrophic pathogens to establish ETS [52].

Nowadays, cultivar-specific resistance activated by species-specific, race-specific or strain-specific effectors is generally discriminated from immune responses triggered by pathogen-associated molecular patterns (PAMPs) that are conserved throughout classes of microbes. However, some pathogens deploy evolutionarily ancient and well-conserved effectors that are instrumental for pathogenicity, forcing plants to evolve recognition of these molecules to become resistant to these pathogens. Essentially, such effectors now act as PAMPs that blur the PAMP-effector dichotomy and illustrate a continuum between immune responses triggered by PAMPs and by effectors. Ultimately, plant resistance is determined by immune receptors that recognize appropriate ligands, the nature and intrinsic function of which is not relevant as long as they accurately betray the microbial invader to the plant [53].

Conclusions

Although all fungal effectors are delivered to the apoplast, they can be divided into two groups: those

that remain in the apoplast and those that translocate into host cells. Recently, a mechanism for effector uptake has been proposed, but the universality of this mechanism is not certain and other means of effector uptake are likely to exist. Without doubt, the major challenge for the future will be to assign biological functions to the increasing number of effector molecules identified in fungal genomes. Typical effector calling based on motifs for extracellular secretion has resulted in the prediction of catalogs containing hundreds of effectors for individual pathogen strains. More sophisticated methods of effector discovery are required to identify those that make major contributions to virulence. Comparative genomics upon resequencing of multiple isolates of a single species or related species with overlapping or differential host ranges can identify signs of evolutionary pressure on specific genes that may be of interest to focus research efforts. Ultimately, understanding the function of individual pathogen effectors is expected to provide new avenues for disease control.

Box 1

Outstanding questions:

- Why do pathogens employ highly lineage-specific effector catalogs while many of their host targets appear to be conserved across host species?
- Which are the targets of filamentous pathogen effectors?
- Why do genomes of filamentous pathogens often encode hundreds of effector proteins?
- Which are the *in planta* triggers of effector gene expression and how are these triggers perceived?
- What is the role of NLP effectors in biotrophic pathogens and the role of LysM effectors in non-pathogenic fungi?
- Do mycorrhizal fungi utilize effectors that target host defense to establish symbioses?

Acknowledgements

The authors thank Pierre de Wit and Matthieu Joosten for critically reading the manuscript.

References

1. Dodds PN, Rathjen JP, **Plant immunity: towards an integrated view of plant-pathogen interactions.** *Nat Rev Genet*, 2010, **11**:539-548.
2. Bolton MD, Thomma BPHJ, **The complexity of nitrogen metabolism and nitrogen-regulated gene expression in plant pathogenic fungi.** *Physiol Mol Pathol*, 2008, **72**:104-110.
3. Michielse CB, *et al*., **The nuclear protein Sge1 of *Fusarium oxysporum* is required for parasitic growth.** *PLoS Pathog,* 2009, **5**:e1000637.
4. Kim S, *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, **20**:e1000909.

5. Schumacher J, de Larrinoa IF, Tudzynski B, **Calcineurin-responsive zinc finger transcription factor CRZ1 of *Botrytis cinerea* is required for growth, development, and full virulence on bean plants.** *Eukaryot Cell*, 2008 **7**:584-601.
6. Soriani FM, *et al*., **Functional characterization of the *Aspergillus fumigatus* CRZ1 homologue, CrzA.** *Mol Microbiol*, 2008, **67**:1274-1291.
7. Skibbe DS, Doehlemann G, Fernandes J, Walbot V, **Maize tumors caused by *Ustilago maydis* require organ-specific genes in host and pathogen.** *Science*, 2010, **328:**89-92.
8. Panstruga R, Dodds PN, **Terrific protein traffic: the mystery of effector protein delivery by filamentous plant pathogens.** *Science*, 2009, **324**:748-750.
9. Kankanala P, Czymmek K, Valent B, **Roles for rice membrane dynamics and plasmodesmata during biotrophic invasion by the blast fungus.** *Plant Cell*, 2007, **19**:706-724.
10. Khang CH *et al*., **Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement.** *Plant Cell*, 2010, **22**:1388-1403.
11. Spanu PD, *et al*., **Genome expansion and gene loss in powdery mildew reveal tradeoffs in extreme parasitism.** *Science*, 2010, **330**:1543-1546.
12. Ma LJ, *et al*., **Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*.** *Nature*, 2010, **464**:367-373.
13. Tonukari NJ, Scott-Craig JS, Walton JD, **The *Cochliobolus carbonum SNF1* gene is required for cell wall-degrading enzymes expression and virulence on maize.** *Plant Cell*, 2000, **12**:237-248.
14. Ottmann C, *et al*., **A common toxic fold mediates microbial attack and plant defense**. *Proc Natl Acad Sci USA*, 2009, **106**:10359-10364.
15. Motteram J, *et al*., **Molecular characterization and functional analysis of MgNLP, the sole NPP1 domain-containing protein, from the fungal wheat leaf pathogen *Mycosphaerella graminicola*.** *Mol Plant Microbe Interact*, 2009, **22**:790-799.
16. van Esse HP, *et al*., **The *Cladosporium fulvum* virulence protein Avr2 inhibits host proteases required for basal defense.** *Plant Cell*, 2008, **20**:1948-1963.
17. Shabab M, *et al*., **Fungal effector protein AVR2 targets diversifying defense-related cys proteases of tomato.** *Plant Cell*, 2008, **20**:1169-1183.
18. Song J, *et al*., **Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3.** *Proc Natl Acad Sci USA*, 2009, **106**:1654-1659.
19. van den Burg HA, *et al*., ***Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection.** *Mol Plant Microbe Interact*, 2006, **19**:1420-1430.
20. de Jonge R, *et al*., **Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants.** *Science*, 2010, **329**:953-955.
21. Kaku H, *et al*., **Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor.** *Proc Natl Acad Sci USA*, 2006, **103**:11086-11091.
22. Shimizu T, *et al*., **Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice.** *Plant J*, 2010, **64**:204-214.
23. Miya A, *et al*., **CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis.** *Proc Natl Acad Sci USA*, 2007, **104**:19613-19618.
24. Bolton MD, *et al*., **The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species.** *Mol Microbiol*, 2008, **69**:119-136.
25. de Jonge R, Thomma BPHJ, **Fungal LysM effectors: extinguishers of host immunity?** *Trends Microbiol*, 2009, **17**:151-157.
26. Stergiopoulos I, *et al*., **Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots.** *Proc Natl Acad Sci USA*, 2010, **107**:7610-7615.

27. Ellis JG, Dodds PN, Lawrence GJ, **Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions.** *Annu Rev Phytopathol*, 2007, **45**:289-306.

28. Rafiqi M, *et al*., **Internalization of flax rust avirulence proteins into flax and tobacco cells can occur in the absence of the pathogen.** *Plant Cell*, 2010, **22**:2017-2032.

29. Kale SD, *et al*., **External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells.** *Cell*, 2010, **142**:981-983.

30. Rehmany AP, *et al*., **Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two Arabidopsis lines.** *Plant Cell*, 2005, **17**:1839-1850.

31. Whisson SC, *et al*., **A translocation signal for delivery of oomycete effector proteins into host plant cells.** *Nature*, 2007, **450**:115-118.

32. Dou D, *et al*., **RXLR mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery.** *Plant Cell*, 2008, **20**:1930-1947.

33. Gan PH, *et al*., **Lipid binding activities of flax rust AvrM and AvrL567 effectors.** *Plant Signal Behav*, 2010, **5**:10.

34. Schornack S, *et al*., **Ancient class of translocated oomycete effectors targets the host nucleus.** *Proc Natl Acad Sci USA*, 2010, **107**:17421-17426.

35. Godfrey D, *et al*., **Powdery mildew fungal effectors candidates share N-terminal Y/F/WxC-motif.** *BMC Genomics*, 2010, **11**:317.

36. Houterman PM, Cornelissen BJC, Rep M, **Suppression of plant resistance gene-based immunity by a fungal effector.** *PLoS Pathog*, 2008, **4**:e1000061.

37. Chen LQ, *et al*., **Sugar transporters for intercellular exchange and nutrition of pathogens.** *Nature*, 2010, **468**:527-532.

38. Bolton MD, **Primary metabolism and plant defense-Fuel for the fire.** *Mol Plant Microbe Interact*, 2009, **22**:487-497.

39. Wahl R, Wippel K, Goos S, Kämper J, Sauer N, **A novel high-affinity sucrose transporter is required for virulence of the plant pathogen *Ustilago maydis*.** *PLoS Pathog*, 2010, **8**:e1000303.

40. Stukenbrock EH, McDonald BA, **Population genetics of fungal and oomycete effectors involved in gene-for-gene interactions.** *Mol Plant Microbe Interact*, 2009, **22**:371-380.

41. van de Wouw AP, *et al*., **Evolution of linked avirulence effectors in *Leptosphaeria maculans* is affected by genomic environment and exposure to resistance genes in host plants.** *PLoS Pathog*, 2010, **6**:e1001180.

42. Rouxel T, *et al*., **Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations.** *Nat Commun*, 2011, **2:**202.

43. Raffaele S, Win J, Cano LM, Kamoun S, **Analyses of genome architecture and gene expression reveal novel candidate virulence factors in the secretome of *Phytophthora infestans*.** *BMC Genomics*, 2010, **11:**637.

44. Raffaele S, *et al*., **Genome evolution following host jumps in the Irish potato famine pathogen lineage.** *Science*, 2010, **330**:1540-1543.

45. Stukenbrock EH, *et al*., **Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen *Mycosphaerella graminicola*.** *PLoS Genet*, 2010, **6**:e1001189.

46. Fedorova ND, *et al*., **Genomics islands in the pathogenic filamentous fungus *Aspergillus fumigatus*.** *PLoS Genet*, 2008, **4**:e1000046.

47. Lum G, Min XJ, **FunSecKB: the Fungal Secretome KnowledgeBase.** *Database*, 2011, **2011:**bar001.

48. Choi J, *et al*., **Fungal secretome database: integrated platform for annotation of fungal secretomes.** *BMC Genomics*, 2010, **11**:105.

49. Schirawski J, *et al*., **Pathogenicity determinants in smut fungi revealed by genome comparison.** *Science,* 2010, **330**:1546-1548.

50. Lorang JM, Sweat TA, Wolpert TJ, **Plant disease susceptibility conferred by a "resistance" gene.** *Proc Natl Acad Sci USA*, 2007, **104**:14681-14866.

51. Nagy ED, Bennetzen JL, **Pathogen corruption and site-directed recombination at a plant disease resistance gene cluster.** *Genome Res*, 2008, **18**:1918-1923.

52. Faris JD, *et al*., **A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens.** *Proc Natl Acad Sci USA*, 2010, **107**:13544-13549.

53. Thomma BPHJ, Nürnberger T, Joosten MHAJ, **Of PAMPs and effectors: The blurred PTI-ETI dichotomy.** *Plant Cell*, 2011, **23**:4-15.

Outline of the thesis

Plant pathogens deliver effectors into the apoplast or cytoplasm of their hosts where they function to modulate host physiology, often through suppression of host defences, or to protect the pathogen from host defence responses employed to halt pathogen colonization. In this thesis I studied the role and evolution of effectors of fungal plant pathogens, with an emphasis on the tomato pathogenic fungi *Cladosporium fulvum* and *Verticillium dahliae*.

During tomato leaf colonization, the biotrophic fungus *C. fulvum* secretes effector proteins into the apoplast. Several of these effectors have been characterized and most of them show no significant homology to each other or to other fungal proteins. However, the Ecp6 effector contains three lysin motifs (LysM domains) that are recognized as carbohydrate-binding modules, and Ecp6 orthologues were identified in various fungal species. In **chapter 2**, we undertook a survey of publicly available sequence data of 70 fungal species to investigate the occurrence and diversity of secreted LysM-containing proteins in the fungal kingdom. The largest group of LysM-containing proteins were found to contain only LysM domains in addition to a signal peptide for secretion and were designated LysM effectors. They are found in species with diverse lifestyles, including pathogens of plants and animals, symbionts and saprophytes. We hypothesize that LysM effectors likely interact with fungal cell wall chitin to prevent elicitation of host defence responses (**chapter 2**).

To unravel the mechanism by which LysM effector Ecp6 contributes to *C. fulvum* virulence on tomato, we first examined the binding affinity of Ecp6 for various insoluble polysaccharides in pull down assays (**chapter 3**). Ecp6 precipitated only with chitin and not with other polysaccharides tested. To identify the number of potential substrates, we hybridized Ecp6 to a glycan array with over 400 glycan substrates and found that Ecp6 interacts only with chitin oligosaccharides. By isothermal titration calorimetry we examined the affinity of Ecp6 with soluble chitin oligosaccharides and revealed that Ecp6 binding to chitin oligosaccharides follows a "one binding site" model and demonstrated that Ecp6 contains three binding sites for chitin oligosaccharides with a degree of GlcNAc polymerization of 4, 5 and 6, which matches with the three LysM domains in Ecp6. We furthermore show that Ecp6 does not protect fungal hyphae against hydrolysis by chitinases, but rather prevents chitin-mediated elicitation of host immune responses. Consistent with a role in suppression of chitin-triggered immunity, Ecp6 is able to successfully compete for binding of chitin oligosaccharides to a plant chitin receptor.

In tomato, resistance against race 1 isolates of the vascular wilt fungi *V. dahliae* and *V. albo-atrum* is governed by the *Ve1* resistance gene. However, the corresponding *Verticillium* effector remained unknown thus far, and various attempts to clone this effector were unsuccessful. In this chapter, we describe a comparative genomics approach to clone the effector that triggers Ve1-mediated resistance (**chapter 4**). A single sequence stretch of approximately 50 Kb was identified that only occurs in race 1 strains, and subsequent transcriptome sequencing of *Verticillium*-infected *Nicotiana benthamiana* plants revealed only one highly expressed ORF within this region. Functional analyses confirmed that this ORF activates Ve1-mediated resistance, and the corresponding effector was called *Ave1* (for *Avirulence on Ve1* tomato). We

furthermore demonstrated that Ave1 contributes to fungal virulence on plants that lack *Ve1*. Interestingly, although orthologous proteins were found in a few fungi, numerous orthologs were found in plants, suggesting that *Verticillium* acquired *Ave1* from plants through horizontal gene transfer. Interestingly, various Ave1 orthologs can activate Ve1-mediated resistance, including the ortholog from the tomato pathogen *Fusarium oxysporum* f. sp. *lycopersici*. Consequently, Ve1 was found to mediate resistance not only against *Verticillium*, but also against *Fusarium*.

Similar to the region carrying *Ave1*, genomic regions with limited distribution were frequently found among the sequenced strains, and amounted up to 1 Mb of sequence for each of the strains (**chapter 5**). By comparative genomics we discovered that extensive chromosomal rearrangements established highly dynamic "plastic" regions in the genomes of *Verticillium* strains, and are responsible for genetic variation between these strains. We found that plastic regions are enriched for *in planta*-induced genes, including effector genes such as *Ave1* and a LysM effector that were both shown to contribute to virulence. We suggest that chromosomal plasticity functions as a mechanism to adapt to changing environments in asexually reproducing organisms such as *Verticillium*.

The final chapter (**chapter 6**) discusses the power of next-generation genomics related to the identification and characterization of effectors from filamentous plant pathogens. The current methods to identify effectors and the application of population genomics to study the diversity and evolution of candidate effectors in order to identify the effectors that are most critical for host colonization are described. Also strategies to implement next-generation genomics to establish durable resistance and study emerging plant diseases are discussed.

Chapter 2

Fungal LysM effectors - Extinguishers of host immunity?

de Jonge R and Thomma BPHJ (2009) **Fungal LysM effectors - Extinguishers of host immunity?** *Trends Microbiol* 17:151-157

Abstract

Lysin motifs (LysMs) have been recognized in prokaryotes and plants as carbohydrate-binding protein modules. Recently, a novel virulence factor with LysMs was characterized from the plant pathogenic fungus *Cladosporium fulvum*. Here, we present a survey of public sequence data of 70 fungal species to demonstrate that putatively secreted LysM-containing proteins are widespread in the fungal kingdom, as they are found in mammalian and plant pathogenic species, in addition to saprophytes. We propose that these putative LysM effectors might have a role in sequestration of chitin oligosaccharides – breakdown products of fungal cell walls that are released during invasion and act as triggers of host immunity – to dampen host defence.

The Lysin motif

The Lysin motif (LysM) was originally identified as a protein domain in a *Bacillus* phage lysozyme, an enzyme that degrades bacterial cell walls by catalysing hydrolysis of glycosidic bonds between the peptidoglycan building blocks N-acetylmuramic acid and N-acetyl-D-glucosamine (GlcNAc) [1]. In addition to lysozymes, LysMs were found in various prokaryotic enzymes, including chitinases that hydrolyse glycosidic bonds in chitin, the long-chain GlcNAc polymer that is the main constituent of fungal cell walls [2].

In addition to prokaryotic enzymes, LysMs also occur in eukaryotic proteins [2,3]. In plants, cell surface receptors with extracellular LysMs have been identified, some of which have a role in the interaction with microbes [4,5]. LysM-containing receptor-like kinases (LysM-RLKs) are composed of one to three extracellular LysMs, a single-pass transmembrane domain and an intracellular kinase domain; they were first identified in legume plants as receptors for nodulation (Nod) factors. Nod factors are lipochitin oligosaccharide signalling molecules that are secreted by symbiotic nitrogen-fixing *Rhizobium* bacteria to initiate symbiosis with their hosts [6-8]. Interestingly, genomes of non-legume plants also harbour *LysM-RLK* genes, indicating their involvement in processes other than establishing symbiosis [5,9]. Indeed, in Arabidopsis, a LysM-RLK was identified as a receptor that activates an immune response to fungal pathogens upon perception of chitin oligosaccharides, breakdown products of fungal cell walls that are released during invasion [10,11]. Previously, a similar cell surface immune receptor for perception of chitin oligosaccharides was identified in rice that contained extracellular LysMs but lacked a cytoplasmic kinase domain [12].

Interestingly, new functions seem to emerge for LysM-containing proteins in fungi. In particular, the LysM protein Ecp6 that is secreted by the tomato pathogenic fungus *Cladosporium fulvum* during colonization of its host was recently identified and found to act as a virulence factor [13]. Here, we report that LysM-containing proteins are widespread among fungi of diverse taxa and lifestyles. Remarkably, based on their overall domain architecture, most of these proteins do not contain any recognizable motif other than a varying number of LysMs. Because it has been demonstrated that LysMs can bind chitin oligosaccharides, we propose that many fungal LysM proteins might be involved in sequestering chitin oligosaccharides to prevent elicitation of host immune responses or attraction of mycoparasites (see Glossary), and/or protection of fungal hyphae against chitinases secreted by competitors.

The structure of LysMs

A typical LysM (Pfam PF01476; see Box 1) consists of 44 amino acids of which the first 16 residues and, albeit to a lesser extent, the final 10 residues are the most conserved (Figure S1). Although eukaryotic LysMs contain multiple conserved cysteine residues that are probably involved in disulphide bridges to increase stability of the domain, prokaryotic LysMs contain extensive secondary structures and hydrogen bond networks to provide the overall tertiary structure [3,14]. The tertiary LysM structure was determined for an *Escherichia coli* membrane-bound lytic murein transferase, the *Bacillus subtilis* YkuD protein, and a human protein of unknown function [15,16] (Sasagawa, A. *et al*., unpublished). These LysMs show a βααβ secondary structure, with the two α-helices packed on the same side of the two-stranded antiparallel β-sheet (Figure 1). Multiple LysMs within the same protein are often separated by Ser, Pro and Thr rich regions that can form a flexible region. Additionally, in plants, a conserved Cys-X-Cys motif is often found in this region [7,8], which might be involved in disulphide bridge formation to provide stability to the extracellular region of the protein.

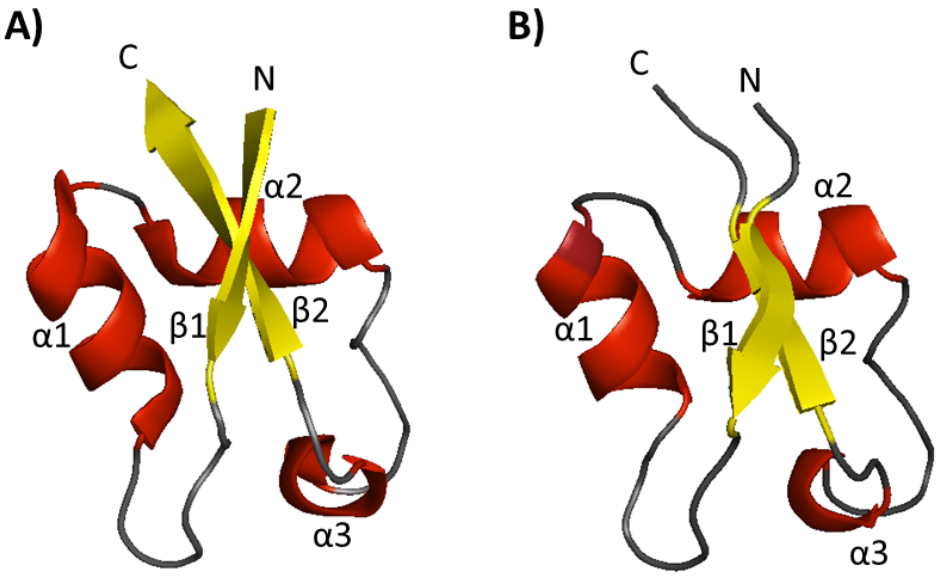


Figure 1. Tertiary structures of LysM domains. A) LysM domain of the human hypothetical protein SB145 (PDB code: 2DJP). B) Predicted tertiary structure for the first LysM domain of the Ecp6 effector from *Cladosporium fulvum*. This model was generated using the nFOLD3 Protein Fold Recognition Server [46] and the atomic coordinates of the LysM domain of SB145. N- and C-terminal ends are indicated, as are the secondary structure elements (α-helices and β-strands, respectively).

Fungal LysM proteins

Presently, over 4,000 LysM proteins are deposited in public databases, and pre-calibrated profile hidden markov models (HMMs) (Box 1) have been created to recognize LysMs in protein sequences. The first identified fungal LysM protein was a chitinase from the yeast *Kluyveromyces lactis* that contains subunits of a killer toxin and inhibits growth of sensitive yeast cells [3,17]. Rqecently, from the tomato pathogenic fungus *C. fulvum* [18], a secreted protein called extracellular protein 6 (Ecp6) was identified, containing three LysMs [13]. The Ecp6 effector was found to be specifically produced during host colonization and required for full pathogen virulence.

Only a few fungal LysM proteins have been characterized in addition to several chitinases and *C. fulvum* Ecp6. To assess the occurrence of LysM proteins in fungi, we composed a database with ~650,000

predicted proteins from 62 publicly available fungal genomes and expressed sequence tag (EST) collections from eight additional species (Table S1). The 70 species encompassed pathogenic and non-pathogenic fungi with a high degree of diversity in growth form, life style and environmental niches, and included members of the ascomycetes, basidiomycetes, zygomycetes, chytridiomycetes and the microsporidia (Table 1). This protein database was queried for LysM proteins (Box 1), and in total 403 protein sequences were identified (Table 1).

Box 1

Sequence analysis of fungal LysM proteins

Profile hidden markov models (HMMs) contain conserved patterns and information for all positions in a multiple sequence alignment of homologous proteins or specific domains [47]. We used the LysM profile HMM in the Pfam database [48] (http://pfam.sanger.ac.uk) to detect LysMs (E-value < 0.001) in ~650,000 predicted fungal proteins with HMMER 2.3.2 [49] (http://hmmer.janelia.org), resulting in the identification of 403 LysM proteins. Subsequent protein domain annotations were performed locally using Pfam (release 23) and pfam_scan.pl (http://www.sanger.ac.uk/Users/sgj/code/pfam/scripts/search/pfam_scan.pl), resulting in the identification of various additional protein domains (Figure 2 in main text). Finally, signal peptides were predicted using the SignalP 3.0 server [50] (http://www.cbs.dtu.dk/services/SignalP/) and cross-verified using WoLF PSORT [51] (http://wolfpsort.org/) to predict subcellular localization. Potential nuclear localization signals and nucleotide-binding regions were identified using predictNLS software [52] (http://cubic.bioc.columbia.edu/predictNLS/) and the Protein Function Prediction Server [53], respectively. The results of our analysis are summarized in Table 1 and Table S1 in supplementary material.

Classification of fungal LysM proteins

We grouped the fungal LysM proteins into five different types, based on their overall domain architecture (Box 1 and Figure 2). The largest group consists of 302 proteins that do not contain any recognizable motif other than a varying number of LysMs (type A; Figure 2). Most of these proteins (121) contain only one LysM, and a decreasing number of proteins contain an increasing number of LysMs, up to seven repeats. *C. fulvum* Ecp6 belongs to this group and contains three LysMs [13]. Hereafter, the type-A proteins are further referred to as putative 'LysM effectors' because of the presence of a secretion motif and the hypothesis that many of them might have a role in the infection process (discussed later and in Box 2).

The second largest group (type B; Figure 2) contains 74 proteins all with homology to chitinases. Besides LysMs, these proteins contain a Cys-rich chitin-binding domain, first identified in wheat germ agglutinin (Pfam ID: PF00187), and an enzymatic domain (Pfam ID: PF00704) responsible for the hydrolytic activity.

The third group of LysM proteins consists of 10 proteins that harbour a CyanoVirin-N homology domain (CVNH; Pfam ID: PF08881) with a nested LysM (type C; Figure 2) [19]. CVNH domains are homologous to the carbohydrate-binding antiviral protein cyanovirin-N (CV-N), a 11 kDa protein that was originally isolated from culture extracts of the cyanobacterium *Nostoc ellipsosporum* in a screen to find novel anti-HIV agents [20]. The antiviral activity of CV-N is mediated through high-affinity interactions with GlcNAc-derived carbohydrate moieties of the HIV envelope glycoprotein gp120, thus hampering virus-to-cell or cell-to-cell fusion [20-22]. By now, CVNH domains are recognized as anti-HIV domains that frequently occur in eukaryotic proteins [19]. Recently, CVNHs from the plant *Ceratopteris richardii* and the fungi *Tuber borchii* and *Neurospora crassa* were shown to bind various carbohydrates [23]. Furthermore, all 10 proteins in this group contain in their N-terminal part a domain of unknown function with similarity to a *Rickettsia* 17 kDa surface antigen (Pfam ID: PF05433).

The remaining 17 LysM proteins contain a variety of domains in combination with the LysM(s). Three of these LysM proteins contain one or two chitin recognition domains (Pfam ID: PF00187) (type D; Figure 2). Another three LysM proteins are putative N-acetylmuramoyl-L-alanine amidases (type E; Figure 2) containing an enzyme domain to cleave the amide bond between N-acetylmuramoyl and L-amino acids in peptidoglycans of bacterial cell walls (Pfam ID: PF01510). The remaining 11 LysM proteins have unique domain compositions (data not shown).

Interestingly, several of the domains that were identified in combination with LysMs in these

Table 1. Assessment for the presence of LysM proteins and LysM effectors in the fungal kingdom

Taxonomy	Species tested	Growth form	Pathogenicity	LysM proteins	LysM effectors
Ascomycetes – Dothideomycetes	8	Filamentous	Plant pathogens	36	32
Ascomycetes – Eurotiomycetes	19	Filamentous	(Opportunistic) mammalian pathogens and saprophytes	125	95
Ascomycetes – Leotiomycetes	3	Filamentous	Plant pathogens	14	11
Ascomycetes – Saccharomycetes	10	Dimorphic	Opportunistic mammalian pathogens and plant pathogens	6	3
Ascomycetes – Sordariomycetes	18	Filamentous	Plant pathogens and saprophytes	172	113
Basidiomycetes – Agaricomycetes	4	Filamentous	Saprophytes	15	15
Basidiomycetes – Homobasidiomycetes	1	Filamentous	Saprophyte	7	7
Basidiomycetes – Pucciniomycetes	1	Filamentous	Plant pathogen	0	0
Basidiomycetes – Urediniomycetes	1	Unicellular	Saprophyte	0	0
Basidiomycetes – Ustilaginomycetes	1	Filamentous	Plant pathogen	2	2
Chytridiomycetes – Chytridiomycetes	1	Unicellular	Amphibian pathogen	2	2
Microsporidia	1	Unicellular	Opportunistic mammalian pathogen	0	0
Zygomycetes – Mucorales	1	Filamentous	Opportunistic mammalian pathogen	11	11
Zygomycetes – Zygomycetes	1	Filamentous	Saprophyte	13	11

different protein types have, like the LysMs themselves, been implicated in carbohydrate binding, namely the chitin-binding domain (found in types B and D), the CVNH domain (found in type C), and the amidase domain (found in type E).

Subcellular localization of fungal LysM proteins

C. fulvum Ecp6 was found to share substantial homology to the *Colletotrichum* intracellular hypha 1 (CIH1) glycoprotein identified in the plant pathogenic fungus *Colletotrichum lindemuthianum* [24]. Although a role in pathogenicity has not been demonstrated, during colonization of bean plants CIH1 was shown to be secreted and accumulate in the walls of intracellular hyphae and the interfacial matrix, which separates the hyphae from the host plasma membrane [24].

In our sequence analysis, we found that all putative LysM effectors (type A) seem to contain signal peptides for secretion (Box 1), as *C. fulvum* Ecp6 and *C. lindemuthianum* CIH1 do. Additionally, ~20% of LysM-containing chitinases (type B) also have signal peptides for secretion. Fungal chitinases function in various processes that include cell wall modification and degradation, such as spore germination, hyphal branching and tip growth, but also mycoparasitism [25]. The latter has been shown for the extracellular chitinase Cht42 from *Trichoderma virens* that is important for the biocontrol of *Rhizoctonia solani* [26]. No signal peptides were identified in the CVNH-LysM proteins (type C), and the presence of several nuclear localization signals (NLSs) upstream of the CVNH domain support nuclear localization of these proteins. In addition, we identified a nucleotide-binding site in a putative region for transcriptional regulation (GO terms GO0000166 and GO0006355, respectively) that co-localizes with the predicted *Rickettsia* 17 kDa surface antigen domain. LysM proteins types D and E are predicted to be secreted proteins.

Distribution of LysM effectors across fungi

Genes encoding LysM proteins occur in saprophytic fungi, which grow on decaying organic matter, in addition to true and opportunistic mammalian and plant pathogens (Table 1). The number of LysM-encoding genes identified in these species greatly varies, and amounts to 19 genes for the saprophytic organisms *Aspergillus nidulans* and *Podospora anserina*. Interestingly, in all species that carry LysM-encoding genes, except for *Histoplasma capsulatum* and *Trichoderma reesii*, the genes encoding putative LysM effectors (type A) are the largest LysM group. Moreover, for 18 species, these LysM effectors are the only LysM proteins encoded by the genome (Table S1).

Of the 70 species tested, LysM-encoding genes are not found in 14 species. However, for five of these species, only relatively small EST collections were available which does not exclude presence of such genes in their genomes. Nonetheless, no LysM-encoding genes are found in nine whole genome sequences; those of the opportunistic mammalian pathogens *Encephalitozoon cuniculi*, *Candida albicans* strain WO1, *C. lusitaniae* and *C. parapsilosis*, the non-pathogenic, free living yeast *Sporobolomyces roseus*, the baker's yeast *Saccharomyces cerevisiae* and the plant pathogens *Ashbya gossypii*, *Blumeria graminis* and *Puccinia graminis*. Taxonomically these species are rather diverse with six ascomycetes, two basidiomycetes and a

microsporidial species, unicellular and filamentous growth forms.

The structure of fungal LysMs

To study the conservation between LysMs of the putative LysM effectors, a multiple sequence alignment of the 669 LysMs found in the 302 LysM effectors was generated and used to build a HMM sequence logo (Box 1 and Figure S1). When comparing this HMM model to the published HMM model that is mainly based on prokaryotic LysMs (Pfam ID: PF01476) it seemed that, like in bacteria, the first 16 amino acid residues are most conserved, whereas the final 10 residues seem less conserved for fungal LysMs. The most remarkable difference, however, is the presence of two highly conserved cysteine residues at positions 9 and 44 in the fungal LysMs. Moreover, two additional, albeit less-conserved, cysteine residues are found at positions 32 and 34 in the fungal LysM HMM model. It has been shown for several *C. fulvum* effectors that the formation of disulphide bridges between cysteine residues is required for stability upon secretion in the host [18,27,28], which might be true also for fungal LysM effectors.

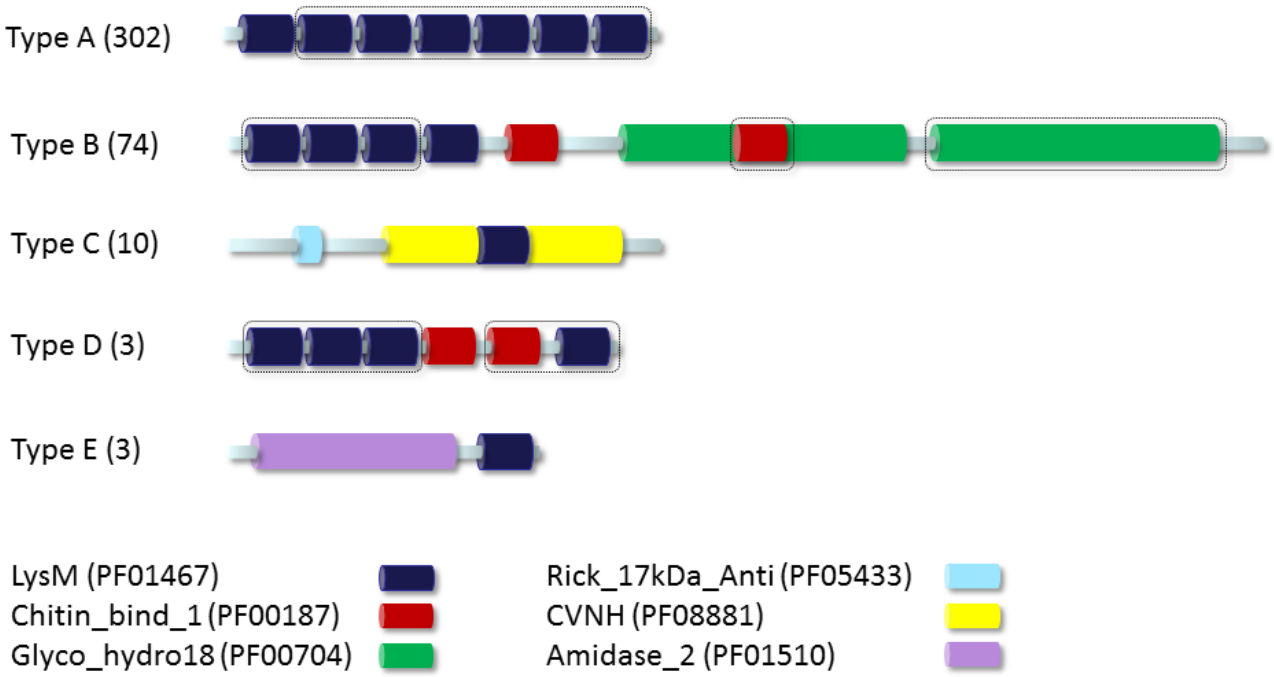


Figure 2. Classification of fungal LysM proteins according to domain architecture. Pfam protein domains were detected in 403 fungal LysM proteins that were subsequently grouped according to their overall domain architecture into 5 types (type A-E). The number of proteins that belong to each type is indicated in brackets. Boxed areas indicate that the respective domain occurs in variable numbers in different proteins belonging to that type.

The biological role of LysM effectors

As noted earlier, LysMs have been implicated in chitin binding. It has previously been shown that the *C. fulvum* effector Avr4, which shares no homology with Ecp6, is a chitin-binding lectin that contains an invertebrate chitin-binding domain [28,29]. Interaction with other cell wall polysaccharides was not observed, and it was proposed that Avr4 could shield chitin on the fungal cell wall, thus preventing degradation by plant chitinases [28,29]. This hypothesis was further substantiated by the finding that Avr4 protected chitin against hydrolysis by plant chitinases *in vitro*, and protected fungi containing exposed

chitin in their cell walls against deleterious concentrations of plant chitinases [30]. Through this effect, Avr4 was shown to contribute to full *C. fulvum* virulence [31]. Through its chitin-binding activity, Ecp6 could be a functional homologue of Avr4 to shield fungal hyphae from chitinases [13]. However, the marked decrease in *C. fulvum* virulence upon silencing of the *Avr4* gene argues against a role of Ecp6 as a functional homologue of Avr4 [31]. Alternatively, Ecp6 might act as a 'stealth factor' by shielding fungal hyphae in a similar fashion as has been suggested for hydrophobins [32], or to avoid recognition by the plant by sequestering chitin monomers or oligomers; these act as elicitors of defence responses once they are released by the activity of plant chitinases (Box 2). The large amounts of Ecp6 that are secreted by *C. fulvum* in the apoplast of infected tomato leaves [13] might support a role as scavenger of these elicitors. A similar role was considered for Avr4, but because of the lower affinity for chito-oligomers when compared to plant receptors, it was concluded that Avr4 could not effectively sequester chitin oligomers to prevent interaction with plant receptors [29,30].

LysM effectors of (opportunistic) mammalian pathogens might have a similar role in suppression of host defences as proposed for the *C. fulvum* LysM effector Ecp6, as it was recently demonstrated that chitin acts as a pathogen-associated molecular pattern (PAMP) that activates host immune responses also in mammals [33].

Intriguingly, putative LysM effectors are not found in oomycetes, a group of 'fungus-like' mycelial organisms of the kingdom Straminopila that contain cellulose cell walls. Although the presence of GlcNAc in oomycete cell walls was recently demonstrated, it corresponds to non-crystalline chitin oligosaccharides associated with glucans, rather than to chitin, and only constitutes a minor constituent of the cell walls [34]. Thus, sequestration of chitin oligosaccharides might be of limited importance to oomycetes.

Most of the 100,000 fungal species are saprophytic organisms that grow on decaying organic matter. Pathogenic species are scattered throughout all taxonomic groups of fungi and are often closely related to non-pathogenic species [35,36], strongly indicating that pathogenicity has evolved at multiple instances during fungal evolution [36-38]. So far, attempts to uncover gene catalogues associated with pathogenic species have met limited success [35-37,39]. Similarly, presence of genes coding for putative LysM effectors is not unique to pathogenic species (Table 1). This seems to be a general phenomenon because many pathogen effectors assembled in the Pathogen-Host Interaction database are found in pathogenic species and saprophytes [37]. In addition to LysM effectors, another example is formed by the Nep1-like proteins (NLPs), a family of microbial proteins that are secreted by plant pathogenic oomycetes, fungi and bacteria [40,41]. NLPs trigger plant defence responses that culminate in cell death [42] and have been shown to contribute to the virulence of necrotizing fungal and bacterial pathogens similarly to host nonselective toxins [43-45]. Like LysM effectors, NLPs occur in both non-pathogenic and pathogenic microbial species, albeit these proteins are more frequently associated to pathogens [37,40].

The observation that putative LysM effectors are preserved in saprophytic species indicates that these proteins also have a role in a saprophytic lifestyle. LysM proteins might be involved in the protection of fungal hyphae against chitinases secreted by competitor microbes. Alternatively, sequestration of chitin oligosaccharides might prevent attraction of mycoparasites. This would then also explain why, in contrast to

Box 2

A hypothesis for Ecp6 function

Basal plant immunity is activated upon recognition of invariant microbial non-self molecules that are released during host colonization, also known as pathogen-associated molecular patterns (PAMPs) [54,55]. Cell surface receptors that recognize the presence of PAMPs activate PAMP-triggered immunity (PTI), including cell wall fortifications, production of reactive oxygen species and the release of antimicrobial compounds such as chitinases, to halt pathogen ingress [54-56]. Successful pathogens overcome PTI by delivery of effector molecules that perturb host defences [57]. Some plants evolved disease resistance proteins to detect these effectors and activate effector-triggered immunity (ETI) leading to localized cell death to restrict the pathogen at the penetration site [57,58].

The role of three effectors (Avr4, Avr2, Ecp6) of the plant pathogen *C. fulvum*, causal agent of tomato leaf mould, has recently been revealed. The Avr4 effector contains an invertebrate chitin-binding domain [28,29] and contributes to virulence by shielding hyphae against hydrolysis by tomato chitinases [30,31]. By contrast, Avr2 inhibits extracellular tomato proteases that act in host defence [59]. Finally, the role of the LysM effector Ecp6 is presently enigmatic.

Like Avr4, Ecp6 might protect fungal cell walls from hydrolysis by plant chitinases. Interestingly, the Avr4 effector is detected by the tomato Cf-4 resistance protein that activates immunity against Avr4-producing *C. fulvum* strains. In turn, *C. fulvum* strains with mutant Avr4 isoforms evolved that circumvent Cf-4 resistance whereas the mutant Avr4 isoforms retained chitin-binding activity [28]. This suggests that Ecp6 and Avr4 are not functional homologues.

A favoured hypothesis is that Ecp6 contributes to virulence by preventing activation of PTI. During growth, fungal organisms release chitin oligosaccharides that act as PAMPs (Figure Ia) [60,61]. The plant cells perceive these chitin oligosaccharides thanks to cell surface receptors with extracellular LysMs; upon ligand binding, the receptors initiate an immune response aimed at the arrest of tissue colonization by the pathogen (Figure Ib) [10-12]. The presence of LysMs in Ecp6 indicates that this effector might compete with host receptors for chitin oligosaccharides. Ecp6 might be secreted by *C. fulvum* to sequester chitin oligosaccharides released in the interface between fungus and host to avoid binding of the PAMP to the host immune receptor, and thus prevent activation of host immune responses (Figure Ic). The large amounts of Ecp6 that are deposited by *C. fulvum* during infection in the tomato apoplast, the extracellular space surrounding the mesophyll cells, might support a role as scavenger of chitin oligosaccharide elicitors [13]. Alternatively, Ecp6 might interact with chitin oligosaccharides that have bound to host immune receptors; in this case, Ecp6 might block or prevent conformational changes of the receptors, or their interaction with other components required for the receptor complex to activate immune signalling.

Box 2: continued

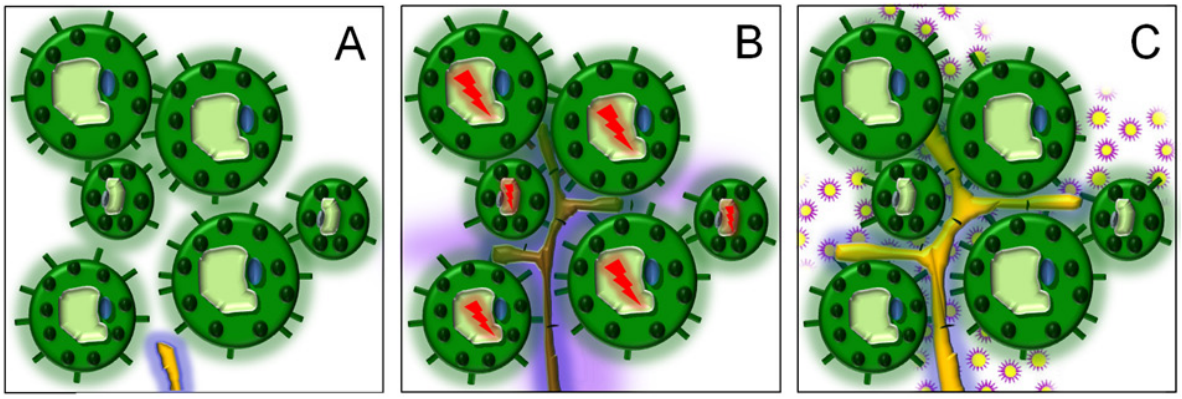


Figure I. LysM effectors as extinguishers of host immunity. A) Plant cells contain cell surface receptors (indicated by green sticks on the surface) to monitor the presence of potential pathogens (such as fungi, yellow) by means of pathogen-associated patterns (such as chitin, purple halo). B) Perception of chitin oligosaccharides by chitin receptors results in the activation of an immune response (red lightning), and the pathogen (brown) is halted. C) LysM effectors (yellow spheres) would contribute to pathogen virulence by sequestering chitin oligosaccharides that thus no longer activate host chitin receptors. As a result, the host immune response is extinguished.

the NLP family, the LysM family is not expanded in pathogenic species when compared with non-pathogenic species.

Concluding remarks and future directions

Here, we have shown that putatively secreted LysM-containing proteins are widespread in the fungal kingdom, and that most of the LysM proteins do not contain any recognizable motif other than their LysM(s). LysM domains have been shown to bind carbohydrates, including chitin oligosaccharides. Therefore, we propose that these putative LysM effectors might have a role in sequestration of chitin oligosaccharides, breakdown products of fungal cell walls that are released during invasion and act as triggers of host immunity. Future experiments should be directed to demonstrate that Ecp6 indeed binds chitin and functions to dampen host defences (Box 3).

Box 3

Outstanding questions:

- What are the targets of fungal LysM effectors? Do these effectors bind chitin oligosaccharides?
- Can fungal LysM effectors protect fungal cell walls against hydrolysis by chitinases?
- If Ecp6 binds chitin oligosaccharides, is the affinity sufficient to compete with plant receptors for these oligosaccharides?
- Does the secretion of fungal LysM effectors dampen host defence responses during host colonization?
- What is the role of LysM effectors in non-pathogenic fungal species?

References

1. Garvey KJ, *et al*., **Nucleotide sequence of Bacillus phage phi 29 genes 14 and 15: homology of gene 15 with other phage lysozymes.** *Nucl Acid Res,* 1986, **14**:10001-10008.
2. Buist G, *et al*., **LysM, a widely distributed protein motif for binding to (peptido)glycans.** *Mol Microbiol*, 1999, **68**:838-847.
3. Ponting CP, *et al*., **Eukaryotic signalling domain homologues in Archae and Bacteria. Ancient ancestry and horizontal gene transfer.** *J Mol Biol*, 1999, **289:**729-745.
4. Knogge W, Scheel D, **LysM receptors recognize friend and foe.** *Proc Natl Acad Sci USA*, 2006, **103**:10829-10830.
5. Zhang X-C, *et al*., **Molecular evolution of lysin motif-type receptor-like kinases in plants.** *Plant Physiol*, 2007, **144**:623-636.
6. Limpens E, *et al.,* **LysM domain receptor kinases regulating rhizobial nod factor-induced infection.** *Science*, 2003, **302**:630-633.
7. Madsen EB, *et al.,* **A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals.** *Nature*, 2003, **425**:637-640.
8. Radutoiu S, *et al.,* **Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases.** *Plant Cell*, 2004, **16**:1220-1234.
9. Shiu, S-H, *et al*., **Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice.** *Plant Cell*, 2004, **16**:1220–1234.
10. Wan J, *et al*., **A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in *Arabidopsis*.** *Plant Cell*, 2008, **20**:471-481.
11. Miya A, *et al*., **CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*.** *Proc Natl Acad Sci USA*, 2007, **104**:19613-19618.
12. Kaku H, *et al*., **Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor.** *Proc Natl Acad Sci USA*, 2006, **103**:11086-11091.
13. Bolton MD, *et al*., **The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species.** *Mol Microbiol*, 2008, **69**:119-136.
14. Ohnuma T, *et al*., **LysM domains from *Pteris ryukyuensis* Chitinase-A: a stability study and characterization of the chitin-binding site.** *J Biol Chem*, 2007, **283**:5178-5187.
15. Bateman A, Bycroft M, **The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD).** *J Mol Biol*, 2000, **299**:1113-1119.
16. Bielnicki J, *et al*., ***B. subtilis* ykuD protein at 2.0 Å resolution: insights into the structure and function of a novel, ubiquitous family of bacterial enzymes.** *Proteins*, 2006, **62**:144-151.
17. Stark MJR, Boyd A, **The killer toxin of *Kluyveromyces lactis*: characterization of the toxin subunits and identification of the genes which encode them.** *EMBO J*, 1986, **5**:1995-2002.
18. Thomma BPHJ *et al*., ***Cladosporium fulvum* syn. *Passalora fulva*, a highly specialized plant pathogen as a model for functional studies on plant pathogenic *Mycosphaerellaceae*.** *Mol Plant Pathol*, 2005, **6**:379-393.
19. Percudani R, *et al*., **The anti-HIV cyanovirin-N domain is evolutionarily conserved and occurs as a protein module in eukaryotes.** *Proteins*, 2005, **60**:670-678.
20. Boyd MR, *et al*., **Discovery of cyanovirin-N, a novel human immunodeficiency virus- inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development.** *Antimicrob Agents Chemother*, 1997, **41**:1521-1530.
21. Bewley CA, Otero-Quintero S, **The potent Anti-HIV protein Cyanovirin-N contains two novel carbohydrate binding sites that selectively bind to Man8 D1D3 and Man9 with nanomolar affinity: implications for binding to the HIV envelope**

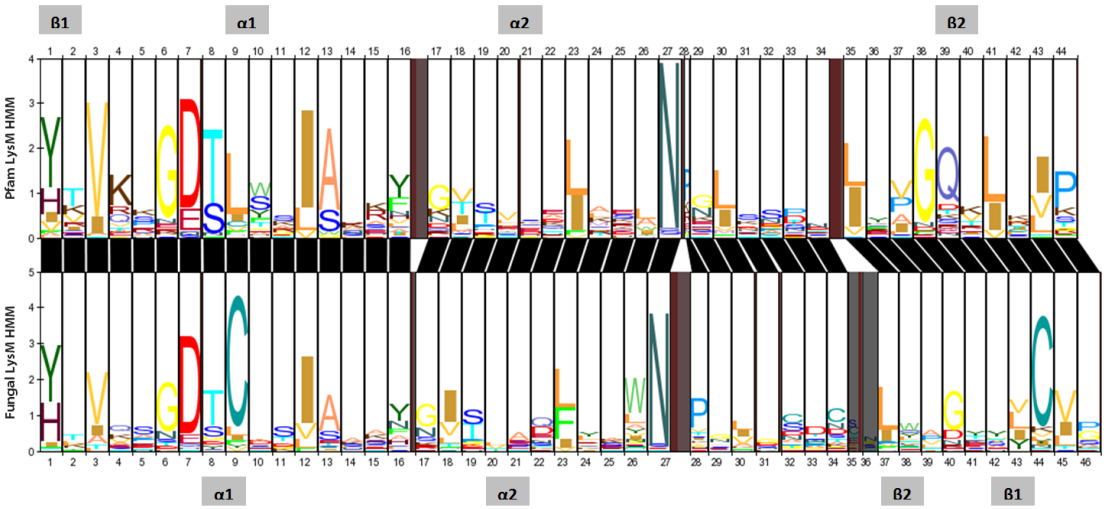
protein gp120. *J Am Chem Soc*, 2001, **123**:3892-3902.

22. Botos I, Wlodawer A, **Cyanovirin-N - a sugar-binding antiviral protein with a new twist.** *Cell Mol Life Sci*, 2003, **60**:277-287.
23. Koharudin LMI, *et al*., **The evolutionarily conserved family of Cyanovirin-N homologs: structures and carbohydrate specificity.** *Structure*, 2008, **16**:570-584.
24. Perfect SE, *et al*., **Expression cloning of a fungal proline-rich glycoprotein specific to the biotrophic interface formed in the *Colletotrichum*-bean interaction.** *Plant J*, 1998, **15**:273-279.
25. Adams DJ, **Fungal cell wall chitinases and glucanases.** *Microbiology*, 2004, **150**:2029-2035.
26. Baek JM, *et al*., **The role of an extracellular chitinase from *Trichoderma virens* Gv29-8 in the biocontrol of *Rhizoctonia solani*.** *Curr Genet*, 1999, **35**:41-50.
27. van den Hooven HW, *et al*., **Disulfide bond structure of the Avr9 elicitor of the fungal tomato pathogen *Cladosporium fulvum*: evidence for a cystine knot.** *Biochemistry Mosc*, 2001, **40**:3458-3466.
28. van den Burg HA, *et al*., **Natural disulfide bond-disrupted mutants of AVR4 of the tomato pathogen *Cladosporium fulvum* are sensitive to proteolysis, circumvent Cf-4-mediated resistance, but retain their chitin binding ability.** *J Biol Chem*, 2003, **278**:27340-27346.
29. van den Burg HA, *et al*., **Binding of the AVR4 elicitor of *Cladosporium fulvum* to chitotriose units is facilitated by positive allosteric protein-protein interactions: the chitin-binding site of AVR4 represents a novel binding site on the folding scaffold shared between the invertebrate and the plant chitin-binding domain.** *J Biol Chem*, 2004, **279**:16786-16796.
30. van den Burg HA, *et al*., ***Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection.** *Mol Plant Microbe Interact*, 2006, **19**:1420-1430.
31. van Esse HP, *et al*., **The chitin-binding *Cladosporium fulvum* effector protein Avr4 is a virulence factor.** *Mol Plant Microbe Interact*, 2007, **20**:1092-1101.
32. Whiteford JR, Spanu PD, **Hydrophobins and the interactions between fungi and plants.** *Mol Plant Pathol*, 2002, **3**:391-400.
33. Da Silva CA, *et al*., **TLR-2 and IL-17A in chitin-induced macrophage activation and acute Inflammation.** *J Immunol*, 2008, **181**:4279-4286.
34. Badreddine I, *et al*., **Cell wall chitosaccharides are essential components and exposed patterns of the phytopathogenic oomycete *Aphanomyces euteiches*.** *Eukaryot Cell*, 2008, **7**:1980-1993.
35. Lievens B, *et al*., **A robust identification and detection assay to discriminate the cucumber pathogens *Fusarium oxysporum* f. sp. *cucumerinum* and f. sp. *radicis-cucumerinum*.** *Environ Microbiol*, 2007, **9**:2145-2161.
36. Lievens B, *et al*., **Recent developments in the molecular discrimination in *formae speciales* of *Fusarium oxysporum*.** *Pest Manag Sci*, 2008, **64**:781-788.
37. Soanes DM, *et al*., **Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis.** *PLoS One*, 2008, **3**:e2300.
38. Tunlid A, Talbot NJ, **Genomics of parasitic and symbiotic fungi.** *Curr Opin Microbiol*, 2002, **5**:513-519.
39. Idnurm A, Howlett BJ, **Pathogenicity genes of phytopathogenic fungi.** *Mol Plant Pathol*, 2001, **2**:241-255.
40. Gijzen M, Nürnberger T, **Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa.** *Phytochemistry*, 2006, **67**:1800-1807.
41. Pemberton CL, Salmond GPC, **The Nep1-like proteins a growing family of microbial elicitors of plant necrosis.** *Mol Plant Pathol*, 2004, **5**:353-359.
42. Qutob D, *et al*., **Phytotoxicity and innate immune responses induced by Nep1-like proteins.** *Plant Cell*, 2006, **18**:3721-3744.
43. Amsellem Z, *et al*., **Engineering hypervirulence in a mycoherbicidal fungus for efficient weed control.** *Nat Biotech*, 2002, **20**:1035-1039.
44. Mattinen L, *et al*., **Identification and Characterization of Nip, Necrosis-inducing virulence protein of *Erwinia carotovora***

subsp. *carotovora*. *Mol Plant Microbe Interact*, 2004, **17**:1366-1375.

45. Pemberton CL, *et al*., **Novel quorum-sensing-controlled genes in *Erwinia carotovora* subsp. *carotovora*: Identification of a fungal elicitor homologue in a soft-rotting bacterium.** *Mol Plant Microbe Interact*, 2005, **18**:343-353.
46. Jones DT, *et al*., **Prediction of novel and analogous folds using fragment assembly and fold recognition.** *Proteins*, 2005, **61**:143-151.
47. Krogh A, *et al*., **Hidden Markov models in computational biology: applications to protein modeling.** *J Mol Biol*, 1994, **235**:1501-1531.
48. Finn RD, *et al*., **The Pfam protein families database.** *Nucl Acids Res*, 2008, **36**:281-288.
49. Eddy SE, **Profile hidden Markov models.** *Bioinformatics*, 1998, **14**:755-763.
50. Bendtsen JD, *et al*., **Improved Prediction of Signal Peptides: SignalP 3.0.** *J Mol Biol*, 2004, **340**:783-795.
51. Horton P, *et al*., **WoLF PSORT: protein localization predictor.** *Nucleic Acids Res*, 2007, **35**:W585-W587.
52. Cokol M, *et al*., **Finding nucelar localization signals.** *EMBO Rep*, 2000, **5**:411-415.
53. Hawkins T, *et al*., **Enhanced automated function prediction using distantly related sequences and contextual association by PFP.** *Protein Sci*, 2006, **15**:1550-1556.
54. Ausubel FM, **Are innate immune signaling pathways in plants and animals conserved?** *Nat Immunol*, 2005, **6**:973-979.
55. Nürnberger T, *et al*., **Innate immunity in plants and animals: striking similarities and obvious differences.** *Immunol Rev*, 2004, **198**:249-266.
56. Zipfel C, Felix G, **Plants and animals: a different taste for microbes?** *Curr Opin Plant Biol*, 2005, **8**:353-360.
57. Göhre V, Robatzek S, **Breaking the barriers: microbial effector molecules subvert plant immunity.** *Annu Rev Phytopathol*, 2008, **46**:189-215.
58. Bent AF, Mackey D, **Elicitors, Effectors, and R genes: the new paradigm and a lifetime supply of questions.** *Annu Rev Phytopathol*, 2007, **45**:399-436.
59. van Esse HP, *et al*., **The *Cladosporium fulvum* virulence protein Avr2 inhibits host proteases required for basal defense.** Plant Cell, 2008, **20**:1948-1963.
60. Boller T, **Chemoperception of microbial signals in plant cells.** *Annu Rev Plant Physiol Plant Mol Biol*, 1995, **46**:189-214.
61. Felix G, *et al*., **Specific perception of subnanomolar concentrations of chitin fragments by tomato cells: induction of extracellular alkalinization, changes in protein phosphorylation, and establishment of a refractory state.** *Plant J*, 1993, **4**:307-316.
62. Schuster-Bockler B, Bateman A, **Visualizing profile-profile alignment: pairwise HMM logos.** *Biofinformatics*, 2005, **21**:2912-2913.
63. Schuster-Bockler B, *et al*., **HMM logos for visualization of protein families.** *BMC Biofinformatics*, 2004, **5**:7.

Supplementary data



Supplemental Figure 1. Pairwise alignment of the fungal LysM HMM sequence logo with the traditional LysM HMM sequence logo. The fungal LysM HMM sequence logo (lower logo) is based on a multiple sequence alignment of 669 LysMs found in the 302 LysM effectors and aligned to the published HMM model (Pfam ID: PF01476; upper logo) using the LogoMat-P software [62, 63]. Secondary structure elements are indicated in grey boxes.

Supplementary Table 1: Fungal species assessed for the presence of LysM proteins and LysM effectors in this study.

SPECIES	TAXONOMY	GROWTH FORM	PATHO-GENICITY	LysM PROTEINS	LysM EFFECTORS	TRANSCRIPTS QUERIED	SOURCE
Alternaria brassicicola	Ascomycetes – Dothideomycetes	Filamentous	Plant pathogen	0	0	3220	COGEME
Ashbya gossypii	Ascomycetes – Saccharomycetes	Dimorphic	Plant pathogen	0	0	4718	EMBL
Aspergillus clavatus	Ascomycetes – Eurotiomycetes	Filamentous	Opportunistic mammalian pathogen	6	6	9121	Broad - FGI
Aspergillus flavus	Ascomycetes – Eurotiomycetes	Filamentous	Opportunistic mammalian pathogen	10	8	12604	Broad - FGI
Aspergillus fumigatus	Ascomycetes – Eurotiomycetes	Filamentous	Opportunistic mammalian pathogen	9	6	9887	Broad - FGI
Aspergillus nidulans	Ascomycetes – Eurotiomycetes	Filamentous	Saprophyte	19	15	10701	Broad - FGI
Aspergillus niger	Ascomycetes – Eurotiomycetes	Filamentous	Saprophyte	12	11	11200	Broad - FGI
Aspergillus oryzae	Ascomycetes – Eurotiomycetes	Filamentous	Saprophyte	10	7	12336	Broad - FGI
Aspergillus terreus	Ascomycetes – Eurotiomycetes	Filamentous	Opportunistic mammalian pathogen	16	11	10406	Broad - FGI
Batrachochytrium dendrobatidis	Chytridiomycetes – Chytridiomycetes	Unicellular	Amphibian pathogen	2	2	8794	Broad - FGI
Blumeria graminis	Ascomycetes – Leotiomycetes	Filamentous	Plant pathogen	0	0	20246	BLUGEN
Botrytis cinerea	Ascomycetes – Leotiomycetes	Filamentous	Plant pathogen	6	4	16448	Broad - FGI
Candida albicans SC5314	Ascomycetes – Saccharomycetes	Dimorphic	Opportunistic mammalian pathogen	1	0	6090	Broad - FGI
Candida albicans WO1	Ascomycetes – Saccharomycetes	Dimorphic	Opportunistic mammalian pathogen	0	0	6160	Broad - FGI
Candida guilliermondii	Ascomycetes – Saccharomycetes	Dimorphic	Opportunistic mammalian pathogen	1	0	5920	Broad - FGI
Candida lusitaniae	Ascomycetes – Saccharomycetes	Dimorphic	Opportunistic mammalian pathogen	0	0	5941	Broad - FGI
Candida parapsilosis	Ascomycetes – Saccharomycetes	Dimorphic	Opportunistic mammalian pathogen	0	0	5733	Broad - FGI
Candida tropicalis	Ascomycetes – Saccharomycetes	Dimorphic	Opportunistic mammalian pathogen	2	2	6258	Broad - FGI
Chaetomium globosum	Ascomycetes – Sordariomycetes	Filamentous	Saprophyte	17	10	11124	Broad - FGI
Cladosporium fulvum	Ascomycetes – Dothideomycetes	Filamentous	Plant pathogen	1	1	512	COGEME
Coccidioides immitis H538.4	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	3	2	10608	Broad - FGI
Coccidioides immitis RMSCC 2394	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	3	2	10403	Broad - FGI
Coccidioides immitis RMSCC 3703	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	2	2	10423	Broad - FGI
Coccidioides immitis RS	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	3	2	10355	Broad - FGI
Coccidioides posadasii RMSCC 3488	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	3	2	9897	Broad - FGI
Coccidioides posadasii Silveira	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	3	2	10060	Broad - FGI

Cochliobolus heterostrophus	Ascomycetes – Dothideomycetes	Filamentous	Plant pathogen	13	12	9633	JGI - DOE
Colletotrichum gloeosporioides	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	0	0	1413	COGEME
Coprinopsis cinereus	Basidiomycetes – Agaricomycetes	Filamentous	Saprophyte	5	5	13544	Broad - FGI
Cryphonectria parasitica	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	0	0	2184	COGEME
Cryptococcus neoformans serotype A	Basidiomycetes – Agaricomycetes	Dimorphic	Opportunistic mammalian pathogen	3	3	7302	Broad - FGI
Debaryomyces hansenii	Ascomycetes – Saccharomycetes	Dimorphic	Opportunistic mammalian pathogen	2	1	6312	Broad - FGI
Encephalitozoon cuniculi	Microsporidia	Unicellular	Opportunistic mammalian pathogen	0	0	1997	EMBL
Fusarium graminearum	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	13	9	13332	Broad - FGI
Fusarium oxysporum	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	16	12	17735	Broad - FGI
Fusarium sporotrichioides	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	0	0	3435	COGEME
Fusarium verticillioides	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	17	9	14179	Broad - FGI
Glomerella cingulata	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	0	0	783	COGEME
Histoplasma capsulatum	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	1	0	9349	Broad - FGI
Laccaria bicolor	Basidiomycetes – Agaricomycetes	Filamentous	Symbiont	4	4	20614	JGI - DOE
Leptosphaeria maculans	Ascomycetes – Dothideomycetes	Filamentous	Plant pathogen	1	1	898	COGEME
Lodderomyces elongisporus	Ascomycetes – Saccharomycetes	Dimorphic	Opportunistic mammalian pathogen	0	0	5802	Broad - FGI
Magnaporthe grisea	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	10	9	11074	Broad - FGI
Mycosphaerella fijiensis	Ascomycetes – Dothideomycetes	Filamentous	Plant pathogen	4	3	10327	JGI - DOE
Mycosphaerella graminicola	Ascomycetes – Dothideomycetes	Filamentous	Plant pathogen	5	5	11395	JGI - DOE
Nectria haematococca	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	18	13	15707	JGI - DOE
Neosartorya fischeri	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	13	11	10406	Broad - FGI
Neurospora crassa	Ascomycetes – Sordariomycetes	Filamentous	Saprophyte	8	5	9826	Broad - FGI
Ophiostoma novo-ulmi	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	1	1	434	COGEME
Paracoccidioides brasiliensis Pb01	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	2	2	9132	Broad - FGI
Paracoccidioides brasiliensis Pb03	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	1	1	7875	Broad - FGI
Paracoccidioides brasiliensis Pb18	Ascomycetes – Eurotiomycetes	Filamentous	Mammalian pathogen	1	1	8741	Broad - FGI
Phanerochaete crysosporium	Basidiomycetes – Homo-basidiomycetes	Filamentous	Saprophyte	7	7	10048	JGI - DOE
Phycomyces blakesleeanus	Zygomycetes – Zygomycetes	Filamentous	Saprophyte	13	11	14792	JGI - DOE
Podospora anserina	Ascomycetes – Sordariomycetes	Filamentous	Saprophyte	19	10	10614	CNRS
Postia placenta	Basidiomycetes – Agaricomycetes	Filamentous	Saprophyte	3	3	17173	JGI - DOE
Puccinia graminis	Basidiomycetes – Pucciniomycetes	Filamentous	Plant pathogen	0	0	20567	Broad - FGI
Pyrenophora tritici-repentis	Ascomycetes – Dothideomycetes	Filamentous	Plant pathogen	7	6	12171	Broad - FGI

Rhizopus oryzae	Zygomycetes – Mucorales	Filamentous	Opportunistic mammalian pathogen	11	11	17467	Broad - FGI
Saccharomyces cerevisiae	Ascomycetes – Saccharomycetes	Dimorphic	Saprophyte	0	0	5695	Broad - FGI
Sclerotinia sclerotiorum	Ascomycetes – Leotiomycetes	Filamentous	Plant pathogen	8	7	14522	Broad - FGI
Sporobolomyces roseus	Basidiomycetes – Urediniomycetes	Unicellular	Saprophyte	0	0	5536	JGI - DOE
Stagonospora nodorum	Ascomycetes – Dothideomycetes	Filamentous	Plant pathogen	5	4	16597	Broad - FGI
Trichoderma atroviride	Ascomycetes – Sordariomycetes	Filamentous	Saprophyte	12	10	11100	JGI - DOE
Trichoderma reesei	Ascomycetes – Sordariomycetes	Filamentous	Saprophyte	8	3	9129	JGI - DOE
Trichoderma virens	Ascomycetes – Sordariomycetes	Filamentous	Saprophyte	18	10	11643	JGI - DOE
Uncinocarpus reesii	Ascomycetes – Eurotiomycetes	Filamentous	Saprophyte	8	4	7798	Broad - FGI
Ustilago maydis	Basidiomycetes – Ustilaginomycetes	Filamentous	Plant pathogen	2	2	6522	Broad - FGI
Verticillium albo-atrum	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	8	6	10535	Broad - FGI
Verticillium dahliae	Ascomycetes – Sordariomycetes	Filamentous	Plant pathogen	7	6	10221	Broad - FGI

Chapter 3

Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants

de Jonge R, van Esse HP, Kombrink A, Shinya T, Desaki Y, Bours R, van der Krol S, Shibuya N, Joosten MHAJ and Thomma BPHJ (2010) **Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants.** *Science* **329**:953-955.

Abstract

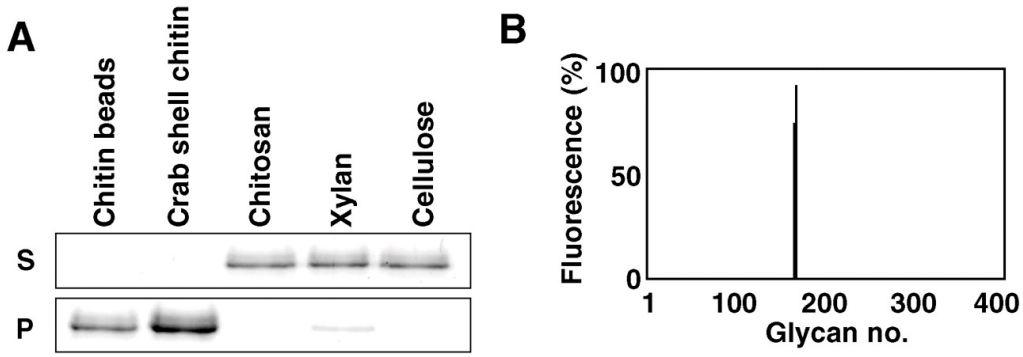
Multicellular organisms activate immunity upon recognition of pathogen-associated molecular patterns (PAMPs). Chitin is the major component of fungal cell walls, and chitin oligosaccharides act as PAMPs in plant and mammalian cells. Microbial pathogens deliver effector proteins to suppress PAMP-triggered host immunity and to establish infection. Here, we show that the LysM domain–containing effector protein Ecp6 of the fungal plant pathogen *Cladosporium fulvum* mediates virulence through perturbation of chitin-triggered host immunity. During infection, Ecp6 sequesters chitin oligosaccharides that are released from the cell walls of invading hyphae to prevent elicitation of host immunity. This may represent a common strategy of host immune suppression by fungal pathogens, because LysM effectors are widely conserved in the fungal kingdom.

Multicellular organisms activate immune responses upon recognition of microbe-derived nonself components. These responses are mediated by pattern recognition receptors (PRRs), cell surface receptors that recognize invariant structures, usually originating from microbial surfaces that are essential for microbial survival and not present in the host. These microbial structures are known as pathogen-associated molecular patterns (PAMPs) [1–4]. Well-known PAMPs include bacterial lipopolysaccharides, peptidoglycans, flagellin, and fungal cell wall–derived glucans and mannans [4–6]. Also chitin, an unbranched β-1-4–linked N-acetyl-glucosamine (GlcNAc) homopolymer that is the major structural component of fungal cell walls, acts as a PAMP in many organisms [6–8]. In the plant species rice and Arabidopsis, single PRRs were shown to be required for the activation of host immunity upon chitin perception [9–12]. Mutants in these receptors are compromised in their response to chitin and are impaired in their defense against chitin-containing fungal pathogens, which indicates that perception of chitin fragments plays a pivotal role in resistance of plants to fungal pathogens. Both chitin receptors of rice and Arabidopsis were shown to contain extracellular LysM domains that generally occur in glycan-binding proteins [9–13].

Cladosporium fulvum is a fungal pathogen that causes leaf mold of tomato (*Solanum lycopersicum*) [14]. During colonization of the intercellular spaces of the leaves, the fungus secretes effector proteins to establish disease, one of which, Avr4, is a chitin-binding lectin that protects fungal cell walls against hydrolysis by plant chitinases [15,16]. Recently, the in planta abundantly secreted *C. fulvum* LysM domain–containing effector Ecp6 was identified and shown to be required for full virulence [17,18].

We first examined the affinity of Ecp6 for insoluble polysaccharides because the presence of three LysM domains in Ecp6 suggested that it has glycan-binding activity. Ecp6 showed specific affinity for chitin, as it coprecipitated with insoluble chitin (chitin beads and crab shell chitin), but not with chitosan (i.e., deacetylated chitin) or the plant cell wall polysaccharides xylan and cellulose [19] (Figure 1A). To further examine Ecp6 substrate specificity, a glycan array was used to test the affinity of Ecp6 for more than 400 different glycan substrates [20]. Ecp6 only bound to the three chitin oligosaccharides present on the array, $(GlcNAc)_3$, $(GlcNAc)_5$, and $(GlcNAc)_6$, but not to any other glycan, including the N-linked glycan chitobiose (Figure 1B and Table S1). Therefore, we conclude that Ecp6 is a highly specific chitin-binding LysM effector.

The affinity of Ecp6 for soluble chitin oligosaccharides was determined by isothermal titration calorimetry (ITC). The binding curves for chitin tetra-, penta- and hexamer oligosaccharides [$(GlcNAc)_4$, $(GlcNAc)_5$, and $(GlcNAc)_6$, respectively] obeyed a "one binding site" model, revealing three binding sites for these oligosaccharides per Ecp6 molecule, which matches with the three LysM domains in Ecp6 (Figure S1, A to C). The $(GlcNAc)_8$ binding curve deviated from this model, which suggested that the size of this octamer allows it to interact with multiple LysM domains simultaneously (Figure S1D). The dissociation constant (K_d) for the various GlcNAc oligosaccharides was similar and decreased from 11.5 to 3.7 µM between $(GlcNAc)_4$ and $(GlcNAc)_8$ (Figure 1C), which showed that Ecp6 had high affinity for chitin oligosaccharides of various lengths. It was previously determined that the invertebrate (CBM14) chitin-binding domain of Avr4 exclusively interacts with $(GlcNAc)_3$ repeats and that the Avr4 K_d decreased from 1.3 mM to 6.3 µM between $(GlcNAc)_4$ and $(GlcNAc)_6$ [21]. This shows that, in contrast to Ecp6, Avr4 has low affinity for short-chain chitin oligosaccharides.



C

Degree of GlcNAc polymerization	*n*	K_b (M^{-1})	K_d (µM)	ΔH (Kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})
4	3.44	868	11.5	-11.0	-14.0
5	3.25	1,160	8.6	-11.4	-14.8
6	3.29	1,560	6.4	-11.2	-13.7
8	1.76	2,760	3.7	-11.3	-13.6

Figure 1. *C. fulvum* Ecp6 binds chitin. (A) Affinity precipitation [19] of Ecp6 with insoluble chitin. Ecp6 protein remaining in concentrated supernatant (S) and the insoluble polysaccharide pellet (P) after SDS-polyacrylamide gel electrophoresis and Coomassie staining. Ecp6 is specifically precipitated with chitin (beads) and chitin from crab shells, but not with other insoluble polysaccharides of plants (xylan, cellulose) and fungi (chitosan). The figure is representative of three independent experiments. (B) Glycan array analysis of Ecp6. Relative fluorescence upon scanning of a glycan array that contains probes for 406 glycans (see Table S1 for identities) after Ecp6 hybridization. Ecp6 only hybridizes to probes 170 to 172 representing $(GlcNAc)_6$, $(GlcNAc)_5$, and $(GlcNAc)_3$,respectively. (C) Thermodynamics of binding of chitin oligosaccharides with different degrees of polymerization to Ecp6. The number of calculated binding sites per protein molecule is represented by *n*.

Avr4 fully protects fungal cell walls against hydrolysis by plant chitinases at a concentration of 10 µM [15,16]. Despite its chitin-binding activity, however, 10 µM or 100 µM Ecp6 failed to protect the fungus *Trichoderma viride* against hydrolysis by crude extracts of tomato leaves containing intracellular basic chitinases (Figure 2A and Figure S2 [19]). We conclude that Ecp6 effector function did not involve prevention of the hydrolysis of fungal cell walls by plant chitinases.

Apart from hydrolyzing fungal cells, host chitinases cause the release of chitin oligosaccharide

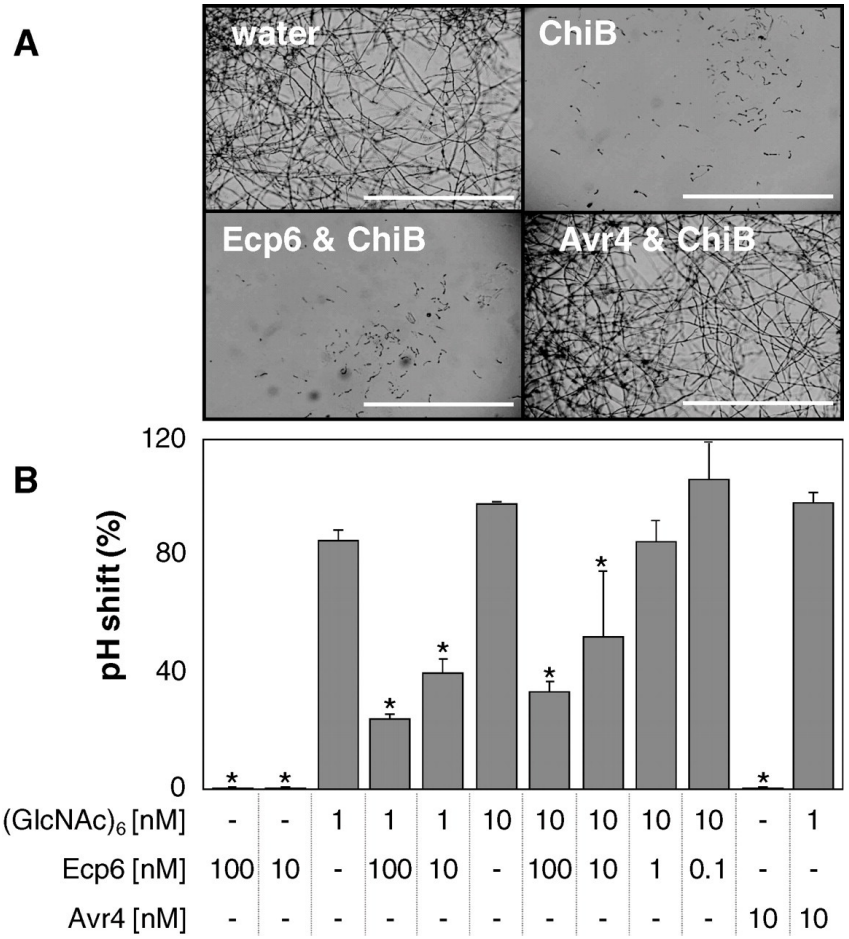


Figure 2. Ecp6 cannot protect fungal hyphae from hydrolysis by tomato chitinases but inhibits chitin-induced medium alkalinization of tomato cell suspensions. (A) Micrographs of *Trichoderma viride* taken 24 hours after the addition of water (w), crude extract of tomato leaves containing intracellular, basic chitinases (ChiB), pretreatment with 10 µM Ecp6 followed by addition of tomato extract (Ecp6 & ChiB), and pretreatment with 10µM Avr4 followed by addition of tomato extract (Avr4 & ChiB). Scalebars, 50 µm. The figure is representative of three independent experiments. (B) Medium alkalinization of tomato cell suspensions induced by chitin oligosaccharides is inhibited by Ecp6. The ΔpH_{max} determined after treatment of tomato cell suspensions with mixtures of chitin oligosaccharides $(GlcNAc)_6$ and Ecp6, after normalization to the response upon treatment with 10 nM $(GlcNAc)_6$ only, is indicated [19]. Bars represent means ± standard error of at least three replicate experiments. Statistically significant differences when compared to treatment with 10 nM $(GlcNAc)_6$ were determined using the Dunnett test (two-sided; * $P < 0.05$).

PAMPs from the cell walls of the invading fungus [6,22]. Suspension-cultured plant cells react with medium alkalinization to treatment with nanomolar concentrations of chitin oligosaccharides [8]. We speculated that LysM effectors might affect chitin perception by the host [17,18] and tested the effect of Ecp6 treatment on PAMP-triggered immunity by measuring chitin-induced pH shifts in tomato and tobacco cell suspensions. Treatment of the cells with nanomolar chitin oligosaccharide [$(GlcNAc)_6$] concentrations resulted in medium alkalinization, whereas addition of equimolar amounts of Ecp6 indeed attenuated this response (Figure 2B). The pH shift attenuation occurred in a dose-dependent manner and, similarly, occurred with various chitin oligosaccharides of different lengths and in both tomato and tobacco suspensions (Figure 2B and Figure S3A). In contrast, even a 10-fold molar excess of Avr4 did not affect the chitin-induced pH shift (Figure 2B and Figure S3B). Similarly, a 10-fold excess of the *C. fulvum* effectors Avr9, Ecp1, and Ecp4, which, like Ecp6, are small cysteine-rich proteins that are abundantly secreted in the apoplast during colonization of the host plant but that do not bind chitin [15] (Figure S3C), did not affect the chitin-induced pH shift (Figure S3B). These data show that only Ecp6 is able to suppress chitin-triggered immunity. The control oligosaccharides

laminarihexaose (β-1,3-glucan), D-cellohexaose (β-1,4-glucan), chitosan hexamer $(GlcN)_6$, and chitosan did not induce a pH shift in tomato cell suspensions (Figure S3D). As expected, Ecp6 did not inhibit alkalinization induced by the bacterial PAMP flg22, the epitope of bacterial flagellin, which suggested that suppression of chitin-triggered immunity by Ecp6 occurs through specific binding of chitin oligosaccharide PAMPs (Figure S3E).

These findings were further substantiated in experiments to assess whether, besides cell suspensions, Ecp6 also perturbs chitin-induced host immunity in tomato and tobacco leaves. Treatment of leaf discs with $(GlcNAc)_6$ resulted in the production of reactive oxygen species (ROS), whereas this response was abolished in the presence of Ecp6 (Figure 3). Furthermore, inhibition of the induction of chitin-responsive genes in the presence of Ecp6 was recorded in tomato leaves (Figure S4). Similar to the alkalinization response in cell suspensions, Avr4 and Avr9 did not affect the chitin-induced ROS production, and Ecp6 could not inhibit

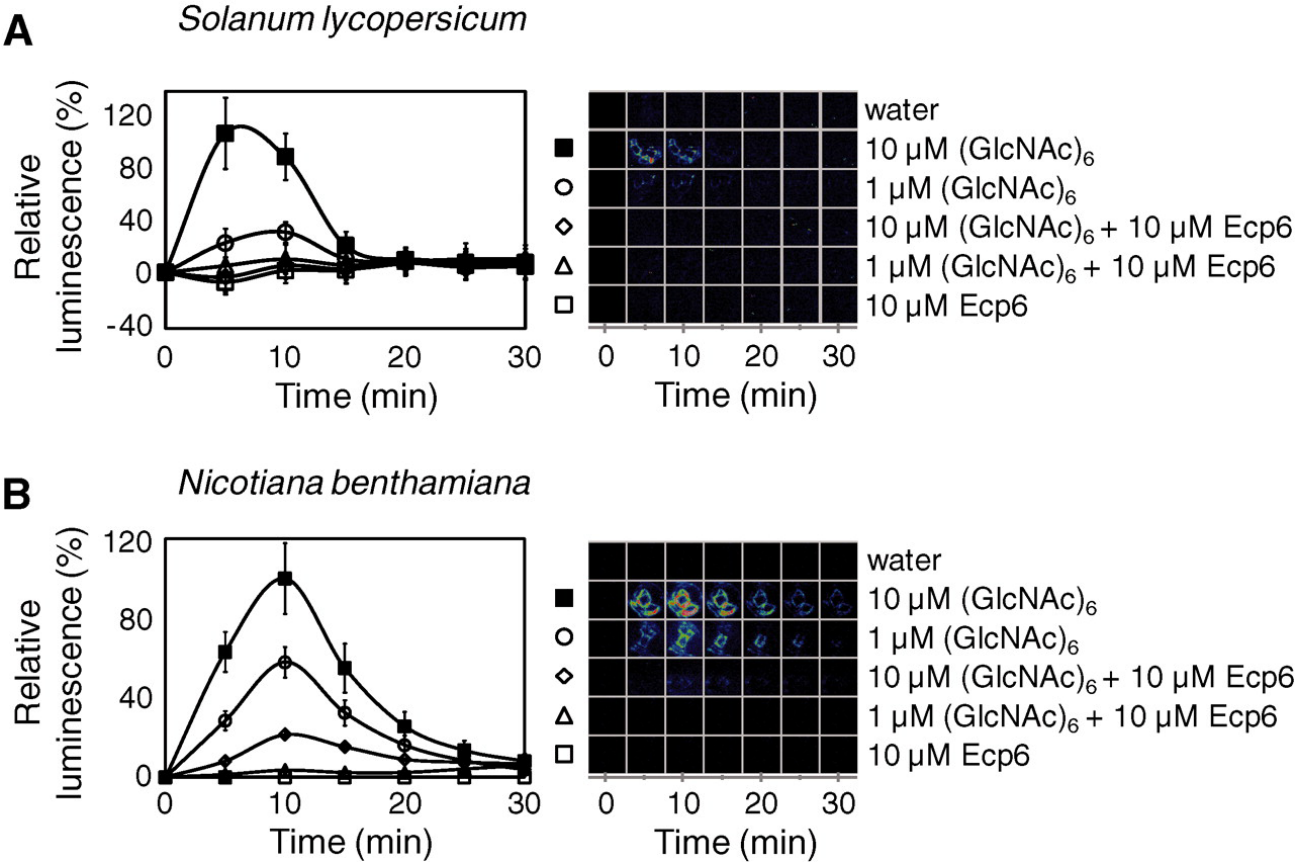


Figure 3. The chitin-induced oxidative burst in (A) tomato, *Solanum lycopersicum*, and (B) tobacco, *N. benthamiana*, leaf discs is inhibited by Ecp6. Production of ROS was determined using luminol-dependent chemiluminescence [19]. (Left) Integrated images of 5-min exposures were analysed with ImageJ, and the relative luminescence was calculated by normalization to water-treated leaf discs and plotted. (Right) Representative image sequence of a single experiment showing the ROS-dependent luminescence over time in leaf discs treated with one of the following: water,10 μM $(GlcNAc)_6$ (■), 1 μM $(GlcNAc)_6$(○), a mixture of 10 μM $(GlcNAc)_6$ and 10 μM Ecp6(◊), a mixture of 1 μM $(GlcNAc)_6$ and 10 μM Ecp6(Δ), and 10 μM Ecp6 (□). The figure is representative of three independent experiments.

the flg22-induced ROS burst (Figure S5).

Finally, we tested whether Ecp6 is able to compete directly for chitin binding with a plant PRR. It has previously been demonstrated that the rice PRR CEBiP directly binds chitin oligosaccharides [9]. As chitin binding was shown to this receptor, we performed competition assays in which a microsomal membrane

preparation from suspension-cultured rice cells containing CEBiP was treated with biotinylated $(GlcNAc)_8$ in the presence or absence of Ecp6 [23]. Although incubation of microsomal membranes with 0.4 µM biotinylated $(GlcNAc)_8$ resulted in labeling of the CEBiP receptor, incubation in the presence of a 100-fold molar excess of nonbiotinylated $(GlcNAc)_8$ prevented receptor labeling (Figure 4). Incubation of microsomal membranes with biotinylated $(GlcNAc)_8$ in the presence of an equimolar amount of Ecp6 almost completely

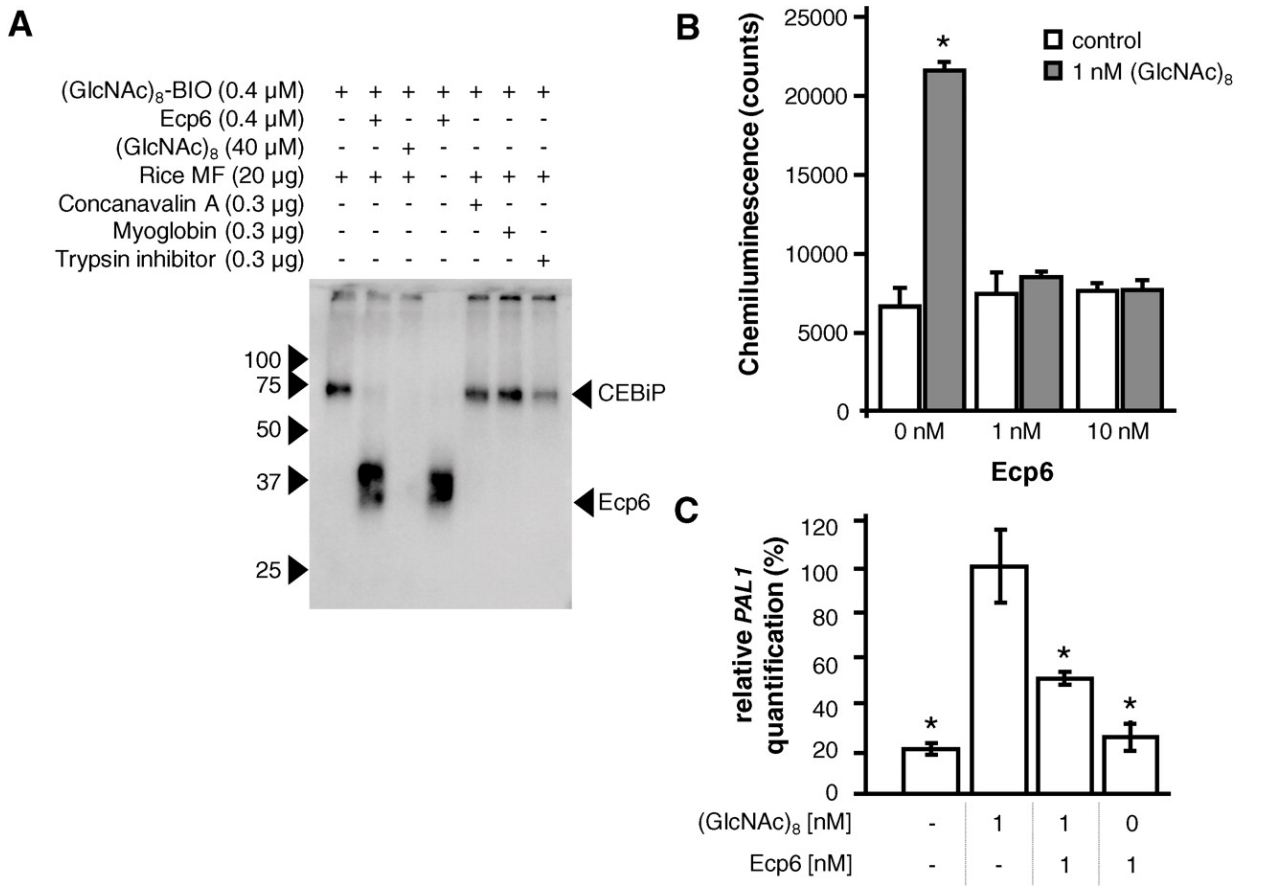


Figure 4. Ecp6 competes for chitin binding with the rice chitin receptor and inhibits chitin-induced defense responses in rice cells. (A) Western blot using a biotin antibody showing affinity labeling of a microsomal membrane preparation (Rice MF) from suspension-cultured rice cells containing the PRR CEBiP, with biotinylated $(GlcNAc)_8$ [$(GlcNAc)_8$-BIO], in the absence or presence of Ecp6, nonbiotinylated $(GlcNAc)_8$, and the control proteins concanavalin A, myoglobin, and trypsin inhibitor. The experiment was performed twice with similar results. (B) Ecp6 inhibits the chitin-induced oxidative burst in rice suspension cells. Production of ROS 20 min after induction with 1 nM $(GlcNAc)_8$ was determined as described previously [9] in the absence or presence of Ecp6 (1and 10 nM). The experiment was performed twice with similar results. (C) Ecp6 inhibits chitin-induced *PAL1* gene expression in rice suspension cells. The bars display the relative transcript level of the chitin-responsive gene *PAL1* normalized to the constitutively expressed *ubiquitin* gene, and the relative transcript level of the suspension cells treated with 1 µM $(GlcNAc)_8$ was set at 100%. The mean with standard error of two replicate experiments is shown, and asterisks indicate significant differences ($P < 0.05$) when compared with the 1 µM $(GlcNAc)_8$ treatment.

prevented receptor labeling while a signal at the height of Ecp6 was observed, which demonstrated that Ecp6 directly competes for chitin binding with the CEBiP chitin receptor (Figure 4).

In conclusion, our data show that the abundantly secreted *C. fulvum* LysM effector Ecp6 is a chitin-binding lectin that inhibits activation of chitin-triggered host immunity. Thus, at present, two chitin-binding *C. fulvum* effectors have been identified; Avr4, which carries an invertebrate chitin-binding domain with high affinity for long-chain chitin oligosaccharides and which protects fungal hyphae against lysis by plant chitinases, and Ecp6, which carries LysM domains with high affinity for various short-chain chitin

oligosaccharides and which prevents activation of chitin PAMP-triggered immunity. These distinct activities of both effectors corroborate the finding that Avr4 protects hyphae against hydrolysis by basic, vacuolar, endochitinases that are released by the host upon cellular collapse [15,16], but not necessarily against exochitinases that are present in the apoplast and that are able to release short-chitin oligosaccharides from the fungal cell wall. Besides differing from Avr4, the scavenger function of Ecp6 also differs from the role of the effector Avr2, which is secreted by *C. fulvum* to inhibit extracellular tomato cysteine proteases that are required for host basal defense [24–26]. *Ecp6* orthologs are present in many fungi, often occurring in multigene families [18]. This suggests that scavenging of chitin oligosaccharides to avoid perception by other organisms may be an important survival strategy of fungi.

Acknowledgments

We are indebted to Yaizu Suisankagaku Industrial Co. Ltd. for the supply of chitosan oligosaccharides and to the Consortium for Functional Glycomics (CFG) for performing the glycan array analysis. The authors thank P. de Wit, F. Govers, T. Bisseling, N. Talbot, and anonymous reviewers for critical evaluation of the manuscript and L. Verhage and L. Kaaij for technical assistance.

References

1. Janeway Jr, CA, Medzhitov R, **Innate immune recognition.** *Ann Rev Immu*, 2002, **20**:197-216.
2. Jones JDG ,Dangl JL, **The plant immune system.** *Nature*, 2006, **444**:323-329.
3. Chisholm ST, *et al*., **Host-microbe interactions: Shaping the evolution of the plant immune response.** *Cell*, 2006, **124**:803-814.
4. Nürnberger T, *et al*., **Innate immunity in plants and animals: Striking similarities and obvious differences.** *Immunol Rev*, 2004, **198**:249-266.
5. Gordon S, **Pattern recognition receptors: Doubling up for the innate immune response.** *Cell*, 2002, **111**:927-930.
6. Boller T, **Chemoperception of microbial signals in plant cells.** *Ann Rev Plant Physiol Plant Mol Biol*, 1995, **46**:189-214.
7. Lee CG, *et al*., **Chitin regulation of immune responses: an old molecule with new roles.** Curr Opin Immunol, 2008, **20**:684-689.
8. Felix G, Regenass M, Boller T, **Specific perception of subnanomolar concentrations of chitin fragments by tomato cells: Induction of extracellular alkalinization, changes in protein phosphorylation, and establishment of a refractory state.** *Plant J*, 1993, **4**:307-316.
9. Kaku H, *et al*., **Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor.** *Proc Natl Acad Sci USA*, 2006, **103**:11086-11091.
10. Miya A, *et al*., **CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*.** *Proc Natl Acad Sci USA*, 2007, **104**:19613-19618.
11. Wan J, *et al*., **A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in *Arabidopsis*.** *Plant Cell*, 2008, **20**:471-481.
12. Iizasa E, Mitsutomi M, Nagano Y, **Direct binding of a plant LysM receptor-like kinase, LysM RLK1/CERK1, to chitin *in vitro*.** *J Biol Chem*, 2010, **285**:2996-3004.
13. Buist G, *et al*., **LysM, a widely distributed protein motif for binding to (peptido)glycans.** *Mol Microbiol*, 2008, **68**:838-847.
14. Thomma BPHJ., *et al*., ***Cladosporium fulvum* (syn. *Passalora fulva*), a highly specialized plant pathogen as a model for**

functional studies on plant pathogenic *Mycosphaerellaceae*. *Mol Plant Pathol*, 2005, **6**:379-393.

15. van den Burg HA, *et al*., ***Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection.** *Mol Plant-Microbe Interact*, 2006, **19**:1420-1430.
16. van Esse HP, *et al*., **The chitin-binding *Cladosporium fulvum* effector protein Avr4 is a virulence factor.** *Mol Plant-Microbe Interact*, 2007, **20**:1092-1101.
17. Bolton MD, *et al*., **The novel Cladosporium fulvum lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species.** *Mol Microbiol*, 2008, **69**:119-136.
18. de Jonge R, Thomma BPHT, **Fungal LysM effectors: extinguishers of host immunity?** *Trends Microbiol*, 2009, **17**:151-157.
19. Supplementary Materials and Methods.
20. Blixt O, *et al*., **Printed covalent glycan array for ligand profiling of diverse glycan binding proteins.** *Proc Natl Acad Sci USA*, 2004, **101**:17033-17038.
21. van den Burg HA, *et al*., **Binding of the AVR4 elicitor of *Cladosporium fulvum* to chitotriose units is facilitated by positive allosteric protein-protein interactions: The chitin-binding site of AVR4 represents a novel binding site on the folding scaffold shared between the invertebrate and the plant chitin-binding domain.** *J Biol Chem*, 2004, **279**:16786-16796.
22. Schlumbaum A, *et al*., **Plant chitinases are potent inhibitors of fungal growth.** *Nature*, 1986, **324**:365-367.
23. Shinya T, *et al*., **Characterization of receptor proteins using affinity cross-linking with biotinylated ligands.** *Plant Cell* Physiol, 2010, **51**:262-270.
24. Rooney HCE, *et al*., ***Cladosporium* Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance.** *Science*, 2005, **308**:1783-1786.
25. van Esse HP, *et al*., **The *Cladosporium fulvum* virulence protein Avr2 inhibits host proteases required for basal defense.** *Plant Cell*, 2008, **20**:1948-1963.
26. Shabab M, *et al*., **Fungal effector protein AVR2 targets diversifying defense-related cys proteases of tomato.** *Plant Cell*, 2008, **20**:1169-1183.
27. Tjoelker LW, *et al*., **Structural and functional definition of the human chitinase chitin-binding domain.** *J Biol Chem*, 2000, **275**:514-520.
28. Wiseman T, Williston S, Brandts JF, Lin L-N, **Rapid measurement of binding constants and heats of binding using a new titration calorimeter.** *Anal Biochem*, 1989, **179**:131-137.
29. Joosten MHAJ, *et al*., **Purification and serological characterization of three basic 15-kilodalton pathogenesis-related proteins from tomato.** *Plant Physiol*, 1990, **94**:585-591.
30. Joosten MHAJ, *et al*., **The phytopathogenic fungus *Cladosporium fulvum* is not sensitive to the chitinase and ß-1,3-glucanase defence proteins of its host, tomato.** *Physiol Mol Plant Pathol*, 1994, **46**:45-59.
31. Felix G, *et al*., **Elicitor-induced ethylene biosynthesis in tomato cells: characterization and use as a bioassay for elicitor action.** *Plant Physiol*, 1991, **97**:19-25.
32. Felix G, Baureithel K, Boller T, **Desensitization of the perception system for chitin fragments in tomato cells.** *Plant physiol*, 1998, **117**:643-650.
33. Nishizawa Y, *et al.*, **Regulation of the chitinase gene expression in suspension-cultured rice cells by N-acetylchitooligosaccharides: differences in the signal transduction pathways leading to the activation of elicitor-responsive genes.** *Plant Mol Biol*, 1999, **39:**907-914.
34. Tsukada K, *et al*., **Rice receptor for chitin oligosaccharide elicitor does not couple to heterotrimeric G-protein: elicitor responses of suspension cultured rice cells from *Daikoku* dwarf (d1) mutants lacking a functional G-protein α-subunit.** *Physiol Plant*, 2002, **116**:373-382.

Supporting material

Materials and Methods

Production of recombinant Ecp6

Plasmid pPIC9-HFEcp6 for expression of affinity-tagged Ecp6 in the yeast *Pichia pastoris* was generated by 5' fusion of the cDNA sequence of mature Ecp6 to the His6-FLAG tag and subsequent ligation into pPIC9 (Invitrogen, Carlsbad, USA). Fermentation to produce recombinant Ecp6 was performed as described previously [24]. After removal of cells and concentration of the supernatant, His6-FLAG-tagged Ecp6 was purified using a Ni^{2+}-NTA Superflow column (Qiagen, Leusden, the Netherlands) according to the manufacturer's protocol. The eluted protein fractions were pooled and dialyzed against Milli-Q water. The final Ecp6 concentration was determined spectrophotometrically at 280 nm and confirmed using the Pierce BCA Protein Assay Kit (Thermo Scientific, USA) with bovine serum albumin (BSA) as a standard.

Polysaccharide affinity precipitation assay

The affinity of Ecp6 for various polysaccharides was determined by incubating 20 µg/ml of Ecp6 with 5 mg of chitin beads (New England Biolabs, Beverly, MA, USA), crab shell chitin, chitosan, xylan or cellulose (all from Sigma, St. Louis, USA) as described previously [27]. The incubations were performed in a total volume of 800 µl of water. Similarly, 5 µg of Avr9 was mixed with 3 mg of crab shell chitin, cellulose and chitosan in a total volume of 500 µl. After 3 h of gentle rocking at ambient temperature, the insoluble fraction was pelleted by centrifugation (5 min, 13,000 × g) and the supernatant was collected. The insoluble fraction was washed three times with water and subsequently boiled in 200 µl of 1% SDS solution. Presence of protein in supernatant and pellet was examined by Tricine SDS-polyacrylamide gel electrophoresis followed by standard silver or Coomassie staining.

Isothermal titration calorimetry

Chitin oligosaccharides ($(GlcNAc)_4$, $(GlcNAc)_5$ and $(GlcNAc)_6$ (purity >95%)) were purchased from Seikagaku Corporation (Tokyo, Japan), and $(GlcNAc)_8$ (purity >80%)) was supplied by Yaizu Suisankagaku Industrial (Shizuoka, Japan) and reacetylated before use. Isothermal titration calorimetry (ITC) experiments were performed at 303 Kelvin (28 °C) following standard procedures using a Microcal VP-ITC calorimeter [28]. The reaction cell (volume of 1.4482 ml) containing the protein sample was continuously stirred at 307 rpm while successive aliquots (7-9 µl) of ligand solution were added (final volume of all additions was ~250 µl). Ligand and protein were dissolved in demineralized water. The Ecp6 concentration in the cell was 22.5 µM and the chitin oligosaccharide concentration used was 800 µM for $(GlcNAc)_4$, $(GlcNAc)_5$, and $(GlcNAc)_6$, and 300 µM for $(GlcNAc)_8$. The heat-of-binding was measured after each injection. The integrated heat effects were analyzed using MicroCal Origin 7.0 (Origin Lab Corporation). The parameters that were varied to minimize the standard deviation of the fit to the experimental data were the binding constant (K_b), the enthalpy change (ΔH), and the number of binding sites per protein molecule (stoichiometry; n). The derived values for K_b, ΔH, and n at 303 Kelvin were used to calculate the changes in free energy (ΔG) and entropy (ΔS) according to the equation Eq 1: $-RT \ln K_b = \Delta G = \Delta H - T\Delta S$. In this equation, R is the ideal gas constant, T the absolute temperature expressed in Kelvin, K_b the equilibrium binding constant, ΔG the change in Gibbs free energy, ΔH the change in enthalpy, ΔT the change in temperature and ΔS the change in entropy.

In vitro fungal growth assays

Isolation of tomato chitinases was performed essentially as described [29,30]. A total protein extract was prepared from 500 g of fresh tomato leaves and soluble proteins were subjected to gel filtration with a Sephadex G-50 column (GE Healthcare, Chalfont St. Giles, UK), at a flow rate of 10 ml/h, and 14 fractions of 15 ml each were collected and dialyzed against demineralized water and subsequently freeze-dried. Each of the freeze-dried fractions was then dissolved in 2 ml of demineralized water and filter-sterilized. Subsequently, the fractions were screened for antifungal activity by challenging 50 µl of an overnight liquid culture of 100 conidia/ml of *Trichoderma viride* with 40 µl of the individual fractions.

Approximately 103 conidia of *T. viride* were incubated overnight at room temperature in 50 µl of potato

dextrose broth in 96-well microtiter plates. Subsequently, Avr4 or Ecp6 protein was added to the conidial suspensions at a final concentration of 10 or 100 µM. After a 2-h incubation period, 40 µl of extract containing tomato chitinases was added. Fungal growth was assessed microscopically after 24 h of incubation at 22°C.

Medium alkalinization experiments

Suspension cultured tomato cell line Msk8 was maintained as described [31] and used 3 to 5 days after subculture for alkalinization experiments [8]. To measure medium alkalinization, 2.5 ml aliquots of the suspension were placed in 12-well culture plates on a rotary shaker at 200 rpm and allowed to equilibrate for at least 2 hours. Upon addition of compounds, the pH of the medium was continuously measured using a small, combined-glass electrode (Mettler Toledo, Switzerland) and registered on a pen recorder. The $\Delta pH_{max,}$ which occurred within 3 to 5 min after application of chitin oligosaccharides, was derived from the recordings [8]. As previously noted by others [32], the maximal pH shift obtained after stimulation varied little within one experiment, when using one batch of cells, but varied significantly between different experiments, when using different batches of cells. To resolve this issue; in each experiment the ΔpH_{max} was normalized to a $(GlcNAc)_6$ control [10 nM]. Prior to addition, mixtures of protein (Ecp6 or Avr4) and chitin oligosaccharides were kept at room temperature for at least one hour while shaking gently. Laminarihexaose (β-1,3-glucan), D-cellohexaose (β-1,4-glucan) and chitosan hexamer $(GlcN)_6$ were obtained from Seikagaku Corporation (Tokyo, Japan).

Measurement of reactive oxygen species generation

Tomato Moneymaker Cf-0 leaf discs (± 0.16 cm2) were floated on water overnight in 24-well culture plates on a rotary shaker at 200 rpm. Reactive oxygen species (ROS) released by the leaf tissue were measured using a chemiluminescent assay. To this end, the water was replaced by 300 µl assay solution (68 µg/ml luminol (Sigma, St. Louis, MO, USA) and 40 µg/ml of horseradish peroxidase (Fluka, Buchs, Switzerland)). ROS was elicited with 1 to 10 µM of chitin oligosaccharides, 10 µM of laminarihexaose, D-cellohexaose or chitosan hexamer oligosaccharides and 10 nM of flg22, in the absence and presence of effector proteins. Prior to addition, all solutions were kept at room temperature for at least one hour while shaking gently. Elicitation in the absence of chitin oligosaccharides (water treatment) was included in all experiments for normalization. Luminescence was measured using a liquid-nitrogen cooled (± -90 °C Celsius) CCD camera system and single images of 5-minute exposures during 30 minutes in total were integrated and analyzed using Megascope and ImageJ, respectively. One hour after induction samples were collected and flash-frozen in liquid nitrogen to be used for subsequent RNA extraction.

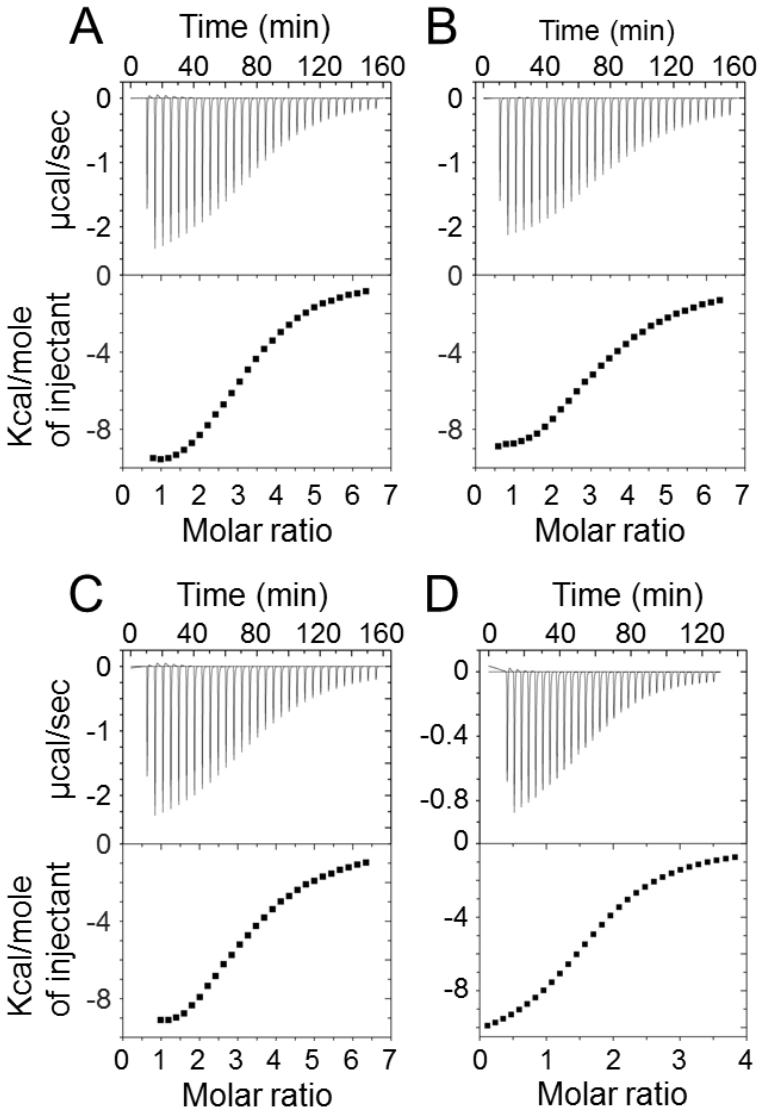
Real-time PCR of chitin-responsive host genes

Rice suspension cells pre-incubated for 15 hours at 25°C while rocking gently (700 rpm) were treated with 1 µM $(GlcNAc)_8$ in the absence or presence of 10 µM Ecp6, collected 30 minutes after treatment [9] and flash-frozen in liquid nitrogen. Total RNA was isolated from leaf material (tomato) or suspension cells (rice) using the RNeasy kit (Qiagen, Leusden, the Netherlands) according to the manufacturer's instructions. Total RNA was used for cDNA synthesis using an oligo(dT) primer (5'-VNTTTTTTTTTTTTTTTTT-3') and the SuperScript III reverse transcriptase kit (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions. The tomato homologs of the chitin-responsive Arabidopsis genes *At3g01830*, *At3g61190*, and *At5g51190* [11], encoding a calmodulin-like protein (SGN-U587927), a homolog of the programmed cell death inhibitor BAP2 (SGN-U582992), and the ethylene response factor 5 (SGN-U583503) were used as template for real-time PCR. Real-time PCR was conducted with primer pairs 5'-TGAGATAACGGTGGAGGAGG-3' and 5'-ACATTCCAAATGCTCCCATC-3' (*SGN-U587927*), 5'-GGTTTTCCAAAGTGGAACGA-3' and 5'-GCAAATAATCTTCGGGCAAA-3' (*SGN-U582992*), 5'-ACTTGAGAGAACGGAAGCCA-3' and 5'-ACCAAACTCGAGTCCCCTTT-3' (*SGN-U583503*), and 5'-CATCGGCAACGAGCGATT-3' and 5'-TGGTACCACCAGACATGACAATG-3' (*actin*) for tomato, and 5'-TGAATAACAGTGGAGTGTGGAG-3' and 5'-AACCTGCCACTCGTACCAAG-3' (*PAL1*) and 5'-AACCAGCTGAGGCCCAAGA-3' and 5'-ACGATTGATTTAACCAGTCCATGA-3' (*Ubq*) for rice [23]. Real-time PCR was conducted using an ABI7300 PCR machine (Applied Biosystems, Foster City, USA) in combination with the qPCR SensiMix kit (BioLine, London, UK). Real-time PCR conditions were as follows: an initial 95°C denaturation step for 10 min followed by denaturation for 15 s at 95°C and annealing/extension for 45 s at 60°C for 40 cycles. The reactions were analyzed on the 7300 System SDS

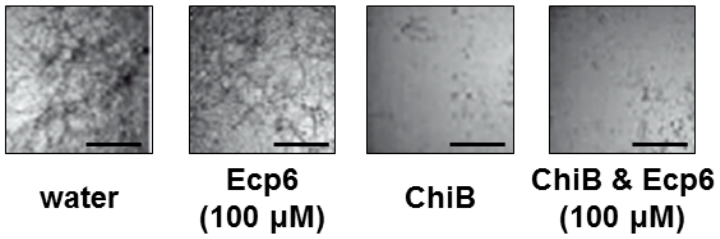
software (Applied Biosystems, Foster City, USA). To check for contamination with genomic DNA, real-time PCR was also performed on RNA without the addition of reverse transcriptase. Statistical analyses were performed in SPSS15.0 using one-way analysis of variance ($P < 0.05$) followed by the LSD and Dunnett t (two-sided) post hoc multiple comparisons.

Affinity labeling of rice membranes with biotinylated $(GlcNAc)_8$

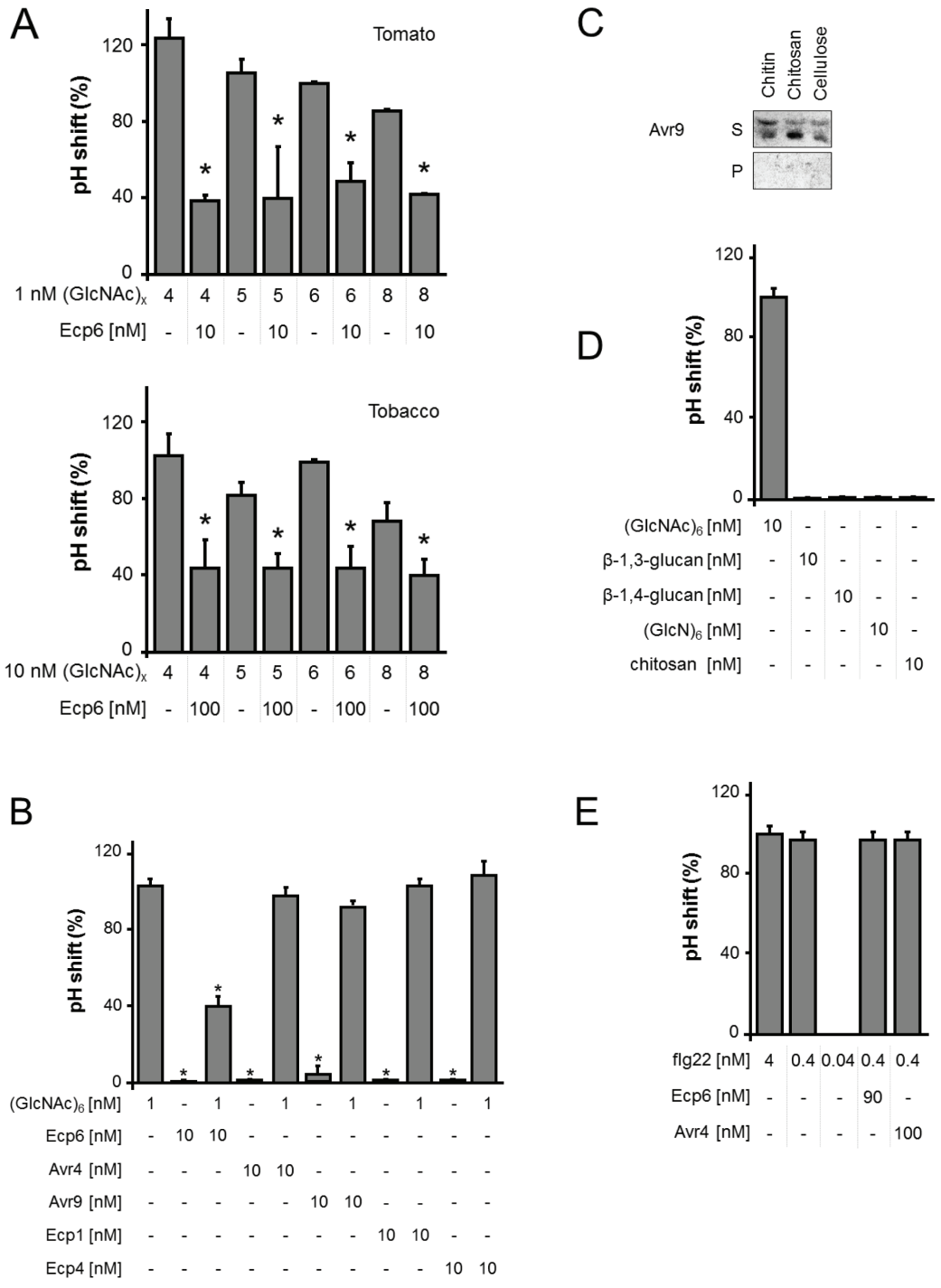
Affinity labeling with biotinylated $(GlcNAc)_8$, the conjugate of biocytin hydrazide and *N*-acetylchitooctaose, was performed as described previously [23]. Suspension-cultured rice cells of *Oryza sativa* L. cv. Nipponbare were maintained in a modified N-6 medium as described previously [34]. A microsomal membrane preparation from suspension-cultured rice cells was mixed with 0.4 μM of biotinylated $(GlcNAc)_8$ in the presence or absence of Ecp6, and adjusted to 30 μl with binding buffer. After incubation for 1h on ice, 3 μl of 3% EGS solution ((ethylene glycol bis[succinimidylsuccinate]); PIERCE, Rockford, IL, USA) was added to the mixture and kept for 30 min. The reaction was stopped by the addition of 1M Tris-HCl, mixed with SDS/PAGE sample buffer, boiled for 5 min, and used for SDS/PAGE. Western blotting was performed on Immun-Blot PVDF Membrane (Bio-Rad Laboratories, Hercules, CA, USA). Detection of biotinylated proteins was performed by using a rabbit antibody against biotin (BETHYL Laboratories, Inc., CA, USA) as a primary antibody and horseradish peroxidase-conjugated goat anti-rabbit IgG (CHEMICON International, Inc., CA, USA) as a secondary antibody. Biotinylated proteins were detected by the chemiluminescence with Immobilon Western Detection reagents (Millipore Corporation, Billerica, MA USA).



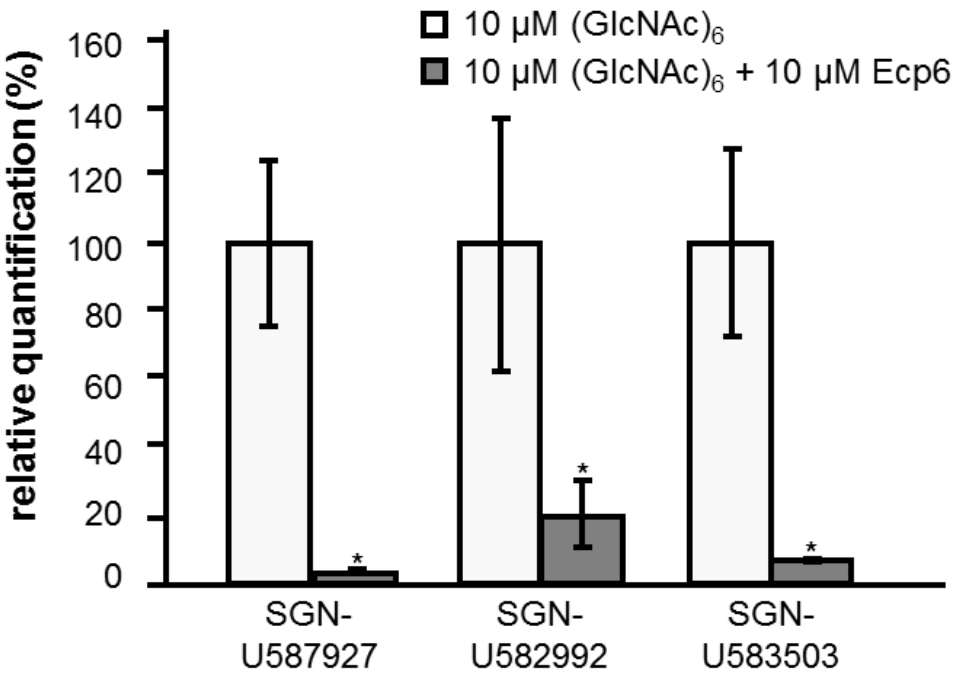
Supplemental Figure 1. Isothermal titration calorimetry (top panels) and the corresponding integrated heat (bottom panels) released by binding of different chitin oligosaccharides to Ecp6. Successive 7-9 µl aliquots of ligand (A: 800 µM of $(GlcNAc)_4$, B: 800 µM of $(GlcNAc)_5$, C: 800 µM of $(GlcNAc)_6$, D: 300 µM of $(GlcNAc)_8$), were added to Ecp6 and the heat-of-binding was measured after each injection. The concentration of Ecp6 in the sample cell was 22.5 µM (A, B and C) and 16.5 µM (D). Graphs shown are typical representatives of two independent replicate experiments.



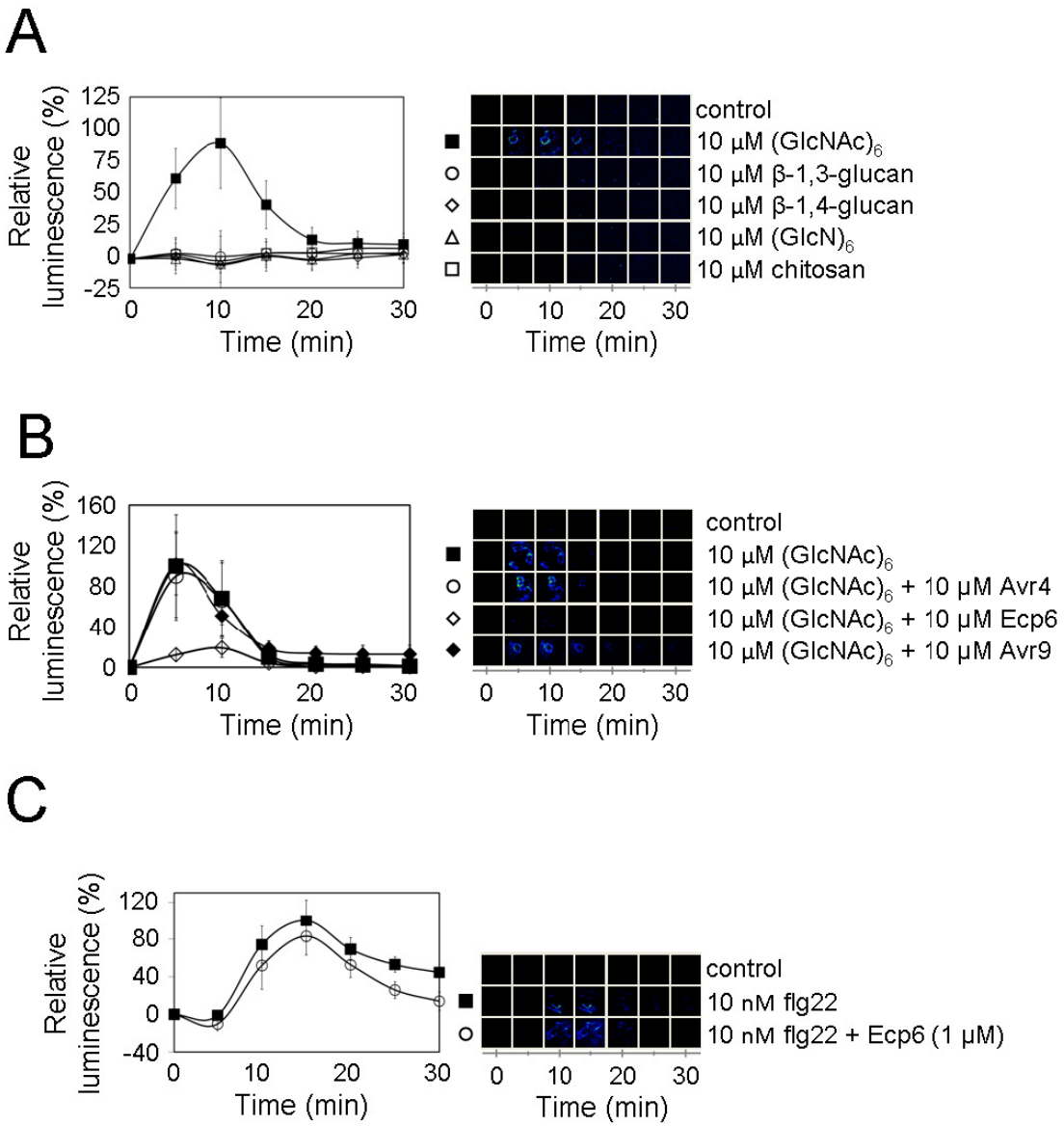
Supplemental Figure 2. Micrographs of *Trichoderma viride* taken 24 h after addition of water, 100 µM Ecp6 protein, crude extract of tomato leaves containing intracellular, basic chitinases (ChiB), and pretreatment with 100 µM Ecp6 followed by addition of tomato extract (Ecp6 & ChiB). The scale bars represent 50 µm. Pictures shown are typical representatives of two independent replicate experiments.



Supplemental Figure 3. Medium alkalinization of tomato (*Solanum lycopersicum*) and tobacco (*Nicotiana benthamiana*) suspension cells induced by chitin oligosaccharides, but not by flg22, is inhibited by Ecp6. (A) ΔpH_{max} determined after treatment of tomato (upper panel) and tobacco (lower panel) suspension cells with mixtures of GlcNAc polymers with varying degrees of polymerization (4, 5, 6 and 8) and Ecp6 after normalization to the response upon treatment with 1 nM (tomato) or 10 nM (tobacco) $(GlcNAc)_6$. The mean with standard error of at least 3 replicate experiments is shown. Statistically significant differences (*) when compared to treatment with 1 nM $(GlcNAc)_6$ were determined using the Dunnett test (2-sided; p-value < 0.05). (B) Ecp6, but not other *Cladosporium fulvum* effectors suppress medium alkalinization induced by chitin oligosaccharides. ΔpH_{max} determined after treatment of tomato suspension cells with 1 nM $(GlcNAc)_6$ and 10 nM of the *C. fulvum* effectors Ecp6, Avr4, Avr9, Ecp1 and Ecp4. The mean with standard error of at least 3 replicate experiments is shown. Statistically significant differences (*) when compared to treatment with 1 nM $(GlcNAc)_6$ were determined using the Dunnett test (2-sided; p-value < 0.05). (C) *C. fulvum* Avr9 does not bind chitin. Affinity precipitation assay [19] of Avr9 with insoluble chitin, chitosan and cellulose. Avr9 protein remaining in concentrated supernatant (S) and the insoluble polysaccharide pellet (P) after SDS-polyacrylamide gel electrophoresis and Coomassie staining is shown. (D) Medium alkalinization induced by control polysaccharides. ΔpH_{max} determined after treatment of tomato suspension cells with β-1,3-glucan, β-1,4-glucan, chitosan hexamers ($(GlcN)_6$), or chitosan when compared to 10 nM of $(GlcNAc)_6$. The mean with standard error of two replicate experiments is shown. (E) Ecp6 does not inhibit flg22-induced medium alkalinization. ΔpH_{max} determined after treatment of tomato suspension cells with mixtures of flg22 and Ecp6 or Avr4 after normalization to the response upon treatment with 4 nM flg22. The mean with standard error of two replicate experiments is shown.



Supplemental Figure 4. Ecp6 dampens expression of chitin-inducible genes in tomato leaf disks. Tomato leaf disks were treated either with 10 µM $(GlcNAc)_6$ in the absence (white bars) or presence (grey bars) of 10 µM Ecp6. One hour after treatment, the effect of Ecp6 was determined on three tomato homologs (*SGN-U587927* encoding a calmodulin-like protein, *SGN-U582992* encoding a homolog of the programmed cell death inhibitor BAP2, and *SGN-U583503* encoding the ethylene response factor 5) of chitin-responsive Arabidopsis genes [23] with real-time PCR. The bars display transcript levels of chitin-responsive genes relative to the constitutively expressed tomato *actin* gene, and the relative transcript levels of the leaf disks treated with 10 µM $(GlcNAc)_6$ were set at 100% for each individual gene. Error bars indicate the 95% confidence interval. Statistically significant differences between the treatments in absence and presence of 10 µM Ecp6 were determined with the Bonferroni Post Hoc (p-value < 0.05) and Dunnett t-test (2-sided; p-value < 0.05) and are indicated by asterisks.



Supplemental Figure 5. The chitin-induced, but not the flg22-induced, oxidative burst in tomato is inhibited by Ecp6 and not by other *Cladosporium fulvum* effectors. Production of reactive oxygen species was determined using luminol-dependent chemiluminescence (26). (A) Left: integrated images of 5 minute exposures were analyzed with ImageJ, and the relative luminescence calculated by normalization to water-treated leaf discs was plotted. Right: representative image sequence of a single experiment showing the reactive oxygen species dependent luminescence over time in leaf discs treated with: water, 10 µM $(GlcNAc)_6$ (■), 10 µM β-1,3-glucan (○),10 µM β-1,4-glucan (◊), 10 µM chitosan hexamers ($(GlcN)_6$) (Δ), and 10 µM chitosan (□). (B) Left: integrated images of 5 minute exposures were analyzed with ImageJ, and the relative luminescence calculated by normalization to water-treated leaf discs was plotted. Right: representative image sequence of a single experiment showing the reactive oxygen species dependent luminescence over time in leaf discs treated with: water, 10 µM $(GlcNAc)_6$ (■), and 10 µM $(GlcNAc)_6$ in the presence of 10 µM Avr4 (○), Ecp6 (◊), or Avr9 (♦). (C) Left: integrated images of 5 minute exposures were analyzed with ImageJ, and the relative luminescence calculated by normalization to water-treated leaf discs was plotted. Right: representative image sequence of a single experiment showing the reactive oxygen species-dependent luminescence over time in leaf discs treated with: water (control), 10 nM flg22 (■), and 10 µM flg22 in the presence of 1 µM Ecp6 (○). For each panel, a representative figure of three independent experiments is shown.

Supplementary Table 1. Glycan array analysis shows that Ecp6 is a chitin-binding protein. The printed array carries 406 glycans in replicates of 6. The table presents the glycan number, the structure or name, the average relative fluorescence unit (RFU) value from the 6 replicates, the standard deviation, the standard error of the mean and the coefficient of variation (%CV=100 X Std. Dev / Mean). The results for the chitin oligosaccharides on the array (probes 170-172) are highlighted.

Chart #	Masterlist Name	RFU	STDEV	SEM	%CV
1	Neu5Aca2-8Neu5Acb-Sp17	14	10	5	70
2	Neu5Aca2-8Neu5Aca2-8Neu5Acb-Sp8	8	8	4	101
3	Neu5Gcb2-6Galb1-4GlcNAc-Sp8	16	5	3	33
4	Galb1-3GlcNAcb1-2Mana1-3(Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	11	10	5	94
5	Gala-Sp8	17	15	7	87
6	Glca-Sp8	9	5	3	60
7	Mana-Sp8	16	9	5	59
8	GalNAca-Sp8	5	7	4	149
9	Fuca-Sp8	30	9	4	29
10	Fuca-Sp9	10	11	5	113
11	Rha-Sp8	3	8	4	287
12	Neu5Aca-Sp8	21	9	4	42
13	Neu5Aca-Sp11	10	13	6	131
14	Neu5Acb-Sp8	22	25	12	111
15	Galb-Sp8	9	12	6	127
16	Glcb-Sp8	31	5	2	16
17	Manb-Sp8	5	3	2	59
18	GalNAcb-Sp8	9	1	1	14
19	GlcNAcb-Sp0	4	5	2	131
20	GlcNAcb-Sp8	11	14	7	128
21	GlcN(Gc)b-Sp8	26	11	6	43
22	Galb1-4GlcNAcb1-3(Galb1-4GlcNAcb1-6)GalNAca-Sp8	14	13	7	98
23	GlcNAcb1-3(GlcNAcb1-4)(GlcNAcb1-6)GlcNAc-Sp8	24	16	8	64
24	[3OSO3][6OSO3]Galb1-4[6OSO3]GlcNAcb-Sp0	15	18	9	114
25	[3OSO3][6OSO3]Galb1-4GlcNAcb-Sp0	18	13	7	71
26	[3OSO3]Galb1-4Glcb-Sp8	10	5	3	52
27	[3OSO3]Galb1-4[6OSO3]Glcb-Sp0	33	8	4	25
28	[3OSO3]Galb1-4[6OSO3]Glcb-Sp8	7	5	3	72
29	[3OSO3]Galb1-3(Fuca1-4)GlcNAcb-Sp8	24	14	7	57
30	[3OSO3]Galb1-3GalNAca-Sp8	14	7	3	48
31	[3OSO3]Galb1-3GlcNAcb-Sp8	3	9	5	291
32	[3OSO3]Galb1-4(Fuca1-3)GlcNAcb-Sp8	7	7	3	90
33	[3OSO3]Galb1-4[6OSO3]GlcNAcb-Sp8	9	8	4	88
34	[3OSO3]Galb1-4GlcNAcb-Sp0	30	24	12	78
35	[3OSO3]Galb1-4GlcNAcb-Sp8	18	15	7	83
36	[3OSO3]Galb-Sp8	29	17	8	59
37	[4OSO3][6OSO3]Galb1-4GlcNAcb-Sp0	17	11	6	67
38	[4OSO3]Galb1-4GlcNAcb-Sp8	8	11	5	130
39	6-H2PO3Mana-Sp8	-1	4	2	-561
40	[6OSO3]Galb1-4Glcb-Sp0	18	12	6	65
41	[6OSO3]Galb1-4Glcb-Sp8	30	30	15	98
42	[6OSO3]Galb1-4GlcNAcb-Sp8	10	6	3	65
43	[6OSO3]Galb1-4[6OSO3]Glcb-Sp8	13	3	2	26
44	Neu5Aca2-3[6OSO3]Galb1-4GlcNAcb-Sp8	30	14	7	47
45	[6OSO3]GlcNAcb-Sp8	16	13	6	79

46	Neu5Ac(9Ac)a-Sp8	16	5	2	31
47	Neu5Ac(9Ac)a2-6Galb1-4GlcNAcb-Sp8	18	6	3	31
48	Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	16	13	7	81
49	GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	25	8	4	32
50	Galb1-4GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	20	4	2	19
51	Galb1-4GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	16	19	10	120
52	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6) Manb1-4GlcNAcb1-4GlcNAcb-Sp12	12	9	4	68
53	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6) Manb1-4GlcNAcb1-4GlcNAcb-Sp13	28	8	4	28
54	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6) Manb1-4GlcNAcb1-4GlcNAcb-Sp8	7	4	2	57
55	Fuca1-2Galb1-3GalNAcb1-3Gala-Sp9	40	16	8	39
56	Fuca1-2Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp9	6	21	11	334
57	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb-Sp8	26	10	5	38
58	Fuca1-2Galb1-3GalNAca-Sp8	10	16	8	153
59	Fuca1-2Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	23	8	4	37
60	Fuca1-2Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp9	15	5	3	36
61	Fuca1-2Galb1-3GlcNAcb1-3Galb1-4Glcb-Sp10	0	7	3	-6002
62	Fuca1-2Galb1-3GlcNAcb1-3Galb1-4Glcb-Sp8	13	15	8	113
63	Fuca1-2Galb1-3GlcNAcb-Sp0	7	17	8	254
64	Fuca1-2Galb1-3GlcNAcb-Sp8	26	11	6	43
65	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	28	7	3	25
66	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3) GlcNAcb-Sp0	35	18	9	52
67	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb-Sp0	14	5	2	35
68	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb-Sp8	29	23	11	80
69	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	23	10	5	44
70	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	19	6	3	30
71	Fuca1-2Galb1-4GlcNAcb-Sp0	17	5	3	32
72	Fuca1-2Galb1-4GlcNAcb-Sp8	17	7	3	38
73	Fuca1-2Galb1-4Glcb-Sp0	14	15	8	105
74	Fuca1-2Galb-Sp8	19	14	7	74
75	Fuca1-3GlcNAcb-Sp8	15	20	10	130
76	Fuca1-4GlcNAcb-Sp8	10	7	3	67
77	Fucb1-3GlcNAcb-Sp8	18	24	12	132
78	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb-Sp0	17	11	5	65
79	GalNAca1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp0	12	8	4	67
80	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	16	12	6	74
81	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	22	16	8	71
82	GalNAca1-3(Fuca1-2)Galb1-4Glcb-Sp0	11	13	6	112
83	GalNAca1-3(Fuca1-2)Galb-Sp8	20	13	6	63
84	GalNAca1-3GalNAcb-Sp8	23	8	4	34
85	GalNAca1-3Galb-Sp8	11	5	3	50
86	GalNAca1-4(Fuca1-2)Galb1-4GlcNAcb-Sp8	11	10	5	85
87	GalNAcb1-3GalNAca-Sp8	12	12	6	105
88	GalNAcb1-3(Fuca1-2)Galb-Sp8	22	16	8	74
89	GalNAcb1-3Gala1-4Galb1-4GlcNAcb-Sp0	9	9	5	107
90	GalNAcb1-4(Fuca1-3)GlcNAcb-Sp0	15	16	8	103

91	GalNAcb1-4GlcNAcb-Sp0	21	15	7	71
92	GalNAcb1-4GlcNAcb-Sp8	18	10	5	57
93	Gala1-2Galb-Sp8	9	12	6	135
94	Gala1-3(Fuca1-2)Galb1-3GlcNAcb-Sp0	19	12	6	66
95	Gala1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp0	55	15	7	27
96	Gala1-3(Fuca1-2)Galb1-4GlcNAc-Sp0	22	21	10	95
97	Gala1-3(Fuca1-2)Galb1-4Glcb-Sp0	12	8	4	64
98	Gala1-3(Fuca1-2)Galb-Sp8	11	18	9	161
99	Gala1-3(Gala1-4)Galb1-4GlcNAcb-Sp8	22	20	10	91
100	Gala1-3GalNAca-Sp8	8	5	3	61
101	Gala1-3GalNAcb-Sp8	23	17	8	73
102	Gala1-3Galb1-4(Fuca1-3)GlcNAcb-Sp8	7	8	4	102
103	Gala1-3Galb1-3GlcNAcb-Sp0	29	16	8	56
104	Gala1-3Galb1-4GlcNAcb-Sp8	30	13	7	45
105	Gala1-3Galb1-4Glcb-Sp0	28	22	11	77
106	Gala1-3Galb-Sp8	26	24	12	95
107	Gala1-4(Fuca1-2)Galb1-4GlcNAcb-Sp8	18	9	5	50
108	Gala1-4Galb1-4GlcNAcb-Sp0	6	10	5	158
109	Gala1-4Galb1-4GlcNAcb-Sp8	10	11	5	115
110	Gala1-4Galb1-4Glcb-Sp0	9	13	7	156
111	Gala1-4GlcNAcb-Sp8	8	7	4	95
112	Gala1-6Glcb-Sp8	29	4	2	14
113	Galb1-2Galb-Sp8	16	5	3	32
114	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	13	6	3	44
115	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4GlcNAcb-Sp0	17	3	1	16
116	Galb1-3(Fuca1-4)GlcNAcb-Sp0	24	11	5	46
117	Galb1-3(Fuca1-4)GlcNAc-Sp8	9	15	8	178
118	Galb1-3(Fuca1-4)GlcNAcb-Sp8	19	8	4	44
119	Galb1-3(Galb1-4GlcNAcb1-6)GalNAca-Sp8	22	24	12	110
120	Galb1-3(GlcNAcb1-6)GalNAca-Sp8	10	10	5	96
121	Galb1-3(Neu5Aca2-6)GalNAca-Sp8	20	8	4	41
122	Galb1-3(Neu5Acb2-6)GalNAca-Sp8	12	6	3	47
123	Galb1-3(Neu5Aca2-6)GlcNAcb1-4Galb1-4Glcb-Sp10	8	17	8	219
124	Galb1-3GalNAca-Sp8	23	10	5	44
125	Galb1-3GalNAcb-Sp8	16	14	7	84
126	Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	30	17	9	57
127	Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	31	17	8	54
128	Galb1-3GalNAcb1-4Galb1-4Glcb-Sp8	13	14	7	112
129	Galb1-3Galb-Sp8	19	12	6	65
130	Galb1-3GlcNAcb1-3Galb1-4GlcNAcb-Sp0	20	15	7	74
131	Galb1-3GlcNAcb1-3Galb1-4Glcb-Sp10	22	10	5	46
132	Galb1-3GlcNAcb-Sp0	14	2	1	12
133	Galb1-3GlcNAcb-Sp8	17	10	5	62
134	Galb1-4(Fuca1-3)GlcNAcb-Sp0	27	14	7	54
135	Galb1-4(Fuca1-3)GlcNAcb-Sp8	37	18	9	48
136	Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb-Sp0	27	12	6	45
137	Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb-Sp0	27	6	3	22
138	Galb1-4[6OSO3]Glcb-Sp0	14	7	4	50
139	Galb1-4[6OSO3]Glcb-Sp8	27	6	3	22

140	Galb1-4GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	20	7	3	35
141	Galb1-4GalNAcb1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	22	12	6	58
142	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-6) Manb1-4GlcNAcb1-4GlcNAcb-Sp12	17	14	7	83
143	Galb1-4GlcNAcb1-3GalNAca-Sp8	15	11	6	76
144	Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	46	9	4	19
145	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	17	6	3	33
146	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	19	6	3	32
147	Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	18	8	4	45
148	Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp8	17	8	4	45
149	Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp8	37	33	16	89
150	Galb1-4GlcNAcb1-6GalNAca-Sp8	12	17	9	142
151	Galb1-4GlcNAcb-Sp0	28	12	6	44
152	Galb1-4GlcNAcb-Sp8	14	10	5	70
153	Galb1-4Glcb-Sp0	21	20	10	96
154	Galb1-4Glcb-Sp8	16	7	4	46
155	GlcNAca1-3Galb1-4GlcNAcb-Sp8	24	25	13	104
156	GlcNAca1-6Galb1-4GlcNAcb-Sp8	8	3	2	41
157	GlcNAcb1-2Galb1-3GalNAca-Sp8	20	13	6	65
158	GlcNAcb1-3(GlcNAcb1-6)GalNAca-Sp8	9	19	10	217
159	GlcNAcb1-3(GlcNAcb1-6)Galb1-4GlcNAcb-Sp8	15	20	10	141
160	GlcNAcb1-3GalNAca-Sp8	17	7	3	40
161	GlcNAcb1-3Galb-Sp8	23	8	4	36
162	GlcNAcb1-3Galb1-3GalNAca-Sp8	28	11	6	40
163	GlcNAcb1-3Galb1-4GlcNAcb-Sp0	2	7	4	394
164	GlcNAcb1-3Galb1-4GlcNAcb-Sp8	15	7	3	43
165	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	23	17	8	71
166	GlcNAcb1-3Galb1-4Glcb-Sp0	12	5	3	46
167	GlcNAcb1-4-MDPLys	25	8	4	30
168	GlcNAcb1-4(GlcNAcb1-6)GalNAca-Sp8	7	10	5	144
169	GlcNAcb1-4Galb1-4GlcNAcb-Sp8	21	17	9	83
170	(GlcNAcb1-4)6b-Sp8	6290	622	311	10
171	(GlcNAcb1-4)5b-Sp8	7656	1254	627	16
172	GlcNAcb1-4GlcNAcb1-4GlcNAcb-Sp8	6093	394	197	6
173	GlcNAcb1-6(Galb1-3)GalNAca-Sp8	11	20	10	177
174	GlcNAcb1-6GalNAca-Sp8	20	14	7	70
175	GlcNAcb1-6Galb1-4GlcNAcb-Sp8	3	8	4	246
176	Glca1-4Glcb-Sp8	14	2	1	14
177	Glca1-4Glca-Sp8	3	14	7	476
178	Glca1-6Glca1-6Glcb-Sp8	25	21	10	82
179	Glcb1-4Glcb-Sp8	9	4	2	40
180	Glcb1-6Glcb-Sp8	11	14	7	123
181	G-ol-Sp8	25	25	13	103
182	GlcAa-Sp8	4	16	8	384
183	GlcAb-Sp8	4	11	6	258
184	GlcAb1-3Galb-Sp8	15	10	5	65
185	GlcAb1-6Galb-Sp8	22	15	7	68
186	KDNa2-3Galb1-3GlcNAcb-Sp0	11	11	6	102
187	KDNa2-3Galb1-4GlcNAcb-Sp0	26	12	6	46
188	Mana1-2Mana1-2Mana1-3Mana-Sp9	28	21	10	73

189	Mana1-2Mana1-3(Mana1-2Mana1-6)Mana-Sp9	19	13	6	69
190	Mana1-2Mana1-3Mana-Sp9	33	13	7	39
191	Mana1-6(Mana1-2Mana1-3)Mana1-6(Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	26	11	6	44
192	Mana1-2Mana1-6(Mana1-3)Mana1-6(Mana1-2Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	3	4	2	144
193	Mana1-2Mana1-2Mana1-3(Mana1-2Mana1-3(Mana1-2Mana1-6)Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	14	16	8	114
194	Mana1-3(Mana1-6)Mana-Sp9	18	11	6	65
195	Mana1-3(Mana1-2Mana1-2Mana1-6)Mana-Sp9	6	8	4	127
196	Mana1-6(Mana1-3)Mana1-6(Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	33	20	10	61
197	Mana1-6(Mana1-3)Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4 GlcNAcb-Sp12	4	12	6	297
198	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-6) Manb1-4GlcNAcb1-4GlcNAcb-Sp12	17	19	9	109
199	Manb1-4GlcNAcb-Sp0	17	4	2	21
200	Fuca1-3(Galb1-4)GlcNAcb1-2Mana1-3(Fuca1-3(Galb1-4)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	5	9	4	175
201	Neu5Aca2-3Galb1-3GalNAca-Sp8	12	3	2	28
202	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	19	3	2	17
203	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	9	16	8	174
204	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb-Sp0	6	4	2	63
205	Neu5Aca2-8Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	9	6	3	71
206	Neu5Aca2-8Neu5Aca2-8Neu5Aca-Sp8	10	11	5	110
207	Neu5Aca2-3(6-O-Su)Galb1-4(Fuca1-3)GlcNAcb-Sp8	25	3	1	11
208	Neu5Aca2-3(GalNAcb1-4)Galb1-4GlcNAcb-Sp0	20	11	5	54
209	Neu5Aca2-3(GalNAcb1-4)Galb1-4GlcNAcb-Sp8	6	3	1	46
210	Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	16	15	7	91
211	Neu5Aca2-3(Neu5Aca2-3Galb1-3GalNAcb1-4)Galb1-4Glcb-Sp0	18	5	3	30
212	Neu5Aca2-3(Neu5Aca2-6)GalNAca-Sp8	28	10	5	35
213	Neu5Aca2-3GalNAca-Sp8	30	26	13	87
214	Neu5Aca2-3GalNAcb1-4GlcNAcb-Sp0	37	7	3	18
215	Neu5Aca2-3Galb1-3[6OSO3]GlcNAc-Sp8	12	11	5	88
216	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb-Sp8	15	13	7	92
217	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	28	15	8	53
218	Neu5Aca2-3Galb1-3(Neu5Aca2-3Galb1-4)GlcNAcb-Sp8	14	3	2	22
219	Neu5Aca2-3Galb1-3[6OSO3]GalNAca-Sp8	19	12	6	64
220	Neu5Aca2-3Galb1-3(Neu5Aca2-6)GalNAca-Sp8	10	12	6	120
221	Neu5Aca2-3Galb-Sp8	22	11	5	50
222	Neu5Aca2-3Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	34	21	10	61
223	Neu5Aca2-3Galb1-3GlcNAcb1-3Galb1-4GlcNAcb-Sp0	10	12	6	120
224	Neu5Aca2-3Galb1-3GlcNAcb-Sp0	15	16	8	105
225	Neu5Aca2-3Galb1-3GlcNAcb-Sp8	8	9	5	114
226	Neu5Aca2-3Galb1-4[6OSO3]GlcNAcb-Sp8	6	18	9	282
227	Neu5Aca2-3Galb1-4(Fuca1-3)[6OSO3]GlcNAcb-Sp8	20	7	4	38
228	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3) GlcNAcb-Sp0	10	11	5	106
229	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	5	8	4	149
230	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp8	13	4	2	28
231	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb-Sp8	6	3	2	60
232	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4GlcNAcb-Sp8	20	12	6	63
233	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAc-Sp0	4	9	4	220

234	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	9	19	9	204
235	Neu5Aca2-3Galb1-4GlcNAcb-Sp0	5	14	7	265
236	Neu5Aca2-3Galb1-4GlcNAcb-Sp8	27	12	6	45
237	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	24	7	3	27
238	Neu5Aca2-3Galb1-4Glcb-Sp0	12	4	2	32
239	Neu5Aca2-3Galb1-4Glcb-Sp8	8	15	7	183
240	Neu5Aca2-6GalNAca-Sp8	6	8	4	128
241	Neu5Aca2-6GalNAcb1-4GlcNAcb-Sp0	30	27	13	89
242	Neu5Aca2-6Galb1-4[6OSO3]GlcNAcb-Sp8	14	8	4	58
243	Neu5Aca2-6Galb1-4GlcNAcb-Sp0	7	6	3	82
244	Neu5Aca2-6Galb1-4GlcNAcb-Sp8	26	18	9	67
245	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	16	10	5	60
246	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	26	15	8	59
247	Neu5Aca2-6Galb1-4Glcb-Sp0	10	4	2	41
248	Neu5Aca2-6Galb1-4Glcb-Sp8	13	16	8	126
249	Neu5Aca2-6Galb-Sp8	16	10	5	64
250	Neu5Aca2-8Neu5Aca-Sp8	13	16	8	126
251	Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb-Sp0	9	8	4	84
252	Neu5Acb2-6GalNAca-Sp8	-1	12	6	-1769
253	Neu5Acb2-6Galb1-4GlcNAcb-Sp8	15	14	7	89
254	Neu5Gca2-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	14	11	6	79
255	Neu5Gca2-3Galb1-3GlcNAcb-Sp0	7	2	1	33
256	Neu5Gca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	20	8	4	40
257	Neu5Gca2-3Galb1-4GlcNAcb-Sp0	-1	7	4	-1240
258	Neu5Gca2-3Galb1-4Glcb-Sp0	28	16	8	56
259	Neu5Gca2-6GalNAca-Sp0	3	11	5	395
260	Neu5Gca2-6Galb1-4GlcNAcb-Sp0	21	10	5	48
261	Neu5Gca-Sp8	18	15	8	85
262	[3OSO3]Galb1-4(Fuca1-3)[6OSO3]Glc-Sp0	18	12	6	64
263	[3OSO3]Galb1-4(Fuca1-3)Glc-Sp0	18	15	8	85
264	[3OSO3]Galb1-4(Fuca1-3)[6OSO3]GlcNAc-Sp8	24	6	3	25
265	[3OSO3]Galb1-4(Fuca1-3)GlcNAc-Sp0	12	8	4	70
266	Fuca1-2[6OSO3]Galb1-4GlcNAc-Sp0	14	13	7	97
267	Fuca1-2Galb1-4[6OSO3]GlcNAc-Sp8	16	13	6	80
268	Fuca1-2[6OSO3]Galb1-4[6OSO3]Glc-Sp0	17	14	7	85
269	Fuca1-2[6OSO3]Galb1-4Glc-Sp0	22	12	6	54
270	Fuca1-2Galb1-4[6OSO3]Glc-Sp0	31	11	5	36
271	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	17	9	5	54
272	Galb1-3(Galb1-4GlcNacb1-6)GalNAc-Sp14	15	12	6	80
273	Galb1-3(GlcNacb1-6)GalNAc-Sp14	-2	3	2	-171
274	Galb1-3(Neu5Aca2-3Galb1-4GlcNacb1-6)GalNAca-Sp14	-3	11	6	-430
275	Galb1-3GalNAca-Sp14	14	13	7	95
276	Galb1-3GlcNAcb1-3Galb1-3GlcNAcb-Sp0	19	9	5	49
277	Galb1-4(Fuca1-3)[6OSO3]GlcNAc-Sp0	7	4	2	56
278	Galb1-4(Fuca1-3)[6OSO3]Glc-Sp0	8	3	1	38
279	Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	12	9	4	72
280	Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	12	7	4	59
281	Neu5Aca2-3Galb1-3GlcNAcb1-3Galb1-3GlcNAcb-Sp0	12	4	2	32
282	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	8	5	2	62

283	[3OSO3]Galb1-4[6OSO3]GlcNAcb-Sp0	17	8	4	44
284	[3OSO3][4OSO3]Galb1-4GlcNAcb-Sp0	22	23	12	109
285	[6OSO3]Galb1-4[6OSO3]GlcNAcb-Sp0	22	3	2	16
286	6-H2PO3Glcb-Sp10	22	14	7	63
287	Gala1-3(Fuca1-2)Galb-Sp18	23	15	7	66
288	Gala1-3GalNAca-Sp16	22	7	3	30
289	Galb1-3GalNAca-Sp16	26	11	6	43
290	Galb1-3(Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-6)GalNAca-Sp14	31	10	5	32
291	Galb1-3Galb1-4GlcNAcb-Sp8	6	14	7	224
292	Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	-4	5	3	-147
293	Galb1-4GlcNAcb1-3(Galb1-4GlcNAcb1-6)Galb1-4GlcNAc-Sp0	5	2	1	36
294	Galb1-4GlcNAcb1-3(GlcNAcb1-6)Galb1-4GlcNAc-Sp0	11	7	4	68
295	Galb1-4GlcNAca1-6Galb1-4GlcNAcb-Sp0	3	3	1	85
296	Galb1-4GlcNAcb1-6Galb1-4GlcNAcb-Sp0	5	19	9	375
297	GalNAca1-3(Fuca1-2)Galb-Sp18	11	9	4	81
298	GalNAca-Sp15	3	7	3	201
299	GalNAcb1-3Galb-Sp8	-1	10	5	-1031
300	GlcAb1-3GlcNAcb-Sp8	5	6	3	124
301	GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	7	15	7	208
302	GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	15	7	3	45
303	GlcNAcb1-3Man-Sp10	12	8	4	65
304	GlcNAcb1-4GlcNAcb-Sp10	24	7	3	29
305	GlcNAcb1-4GlcNAcb-Sp12	10	8	4	79
306	HOOC(CH3)CH-3-O-GlcNAcb1-4GlcNAcb-Sp10	4	10	5	247
307	Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	19	16	8	87
308	Mana1-6Manb-Sp10	14	8	4	54
309	Mana1-6(Mana1-3)Mana1-6(Mana1-3)Manb-Sp10	7	5	3	71
310	Mana1-2Mana1-2Mana1-3(Mana1-2Mana1-6(Mana1-3)Mana1-6)Mana-Sp9	17	19	9	108
311	Mana1-2Mana1-2Mana1-3(Mana1-2Mana1-6(Mana1-2Mana1-3)Mana1-6)Mana-Sp9	11	16	8	147
312	Neu5Aca2-3Galb1-3(Neu5Aca2-3Galb1-4GlcNAcb1-6)GalNAca-Sp14	13	22	11	165
313	Neu5Aca2-3Galb1-3(Neu5Aca2-6)GalNAca-Sp14	27	10	5	37
314	Neu5Aca2-3Galb1-3GalNAca-Sp14	7	9	4	135
315	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	21	10	5	50
316	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	18	9	5	53
317	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	12	16	8	135
318	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-N(LT)AVL	21	4	2	20
319	Fuca1-2Galb1-3GalNAca-Sp14	25	15	8	61
320	Galb1-3(Neu5Aca2-6)GalNAca-Sp14	8	9	4	112
321	Galb1-4GlcNAcb1-3GalNAc-Sp14	21	2	1	9
322	Neu5Ac(9Ac)a2-3Galb1-4GlcNAcb-Sp0	11	6	3	56
323	Neu5Ac(9Ac)a2-3Galb1-3GlcNAcb-Sp0	23	19	9	81
324	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	13	17	9	133
325	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	29	16	8	56
326	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	26	15	7	57
327	Gala1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	8	11	5	133

328	GalNAcb1-3Gala1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	9	10	5	108
329	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	30	16	8	51
330	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	17	12	6	68
331	(Neu5Aca2-3-Galb1-3)(((Neu5Aca2-3-Galb1-4(Fuca1-3))GlcNAcb1-6)GalNAc-Sp14	16	6	3	36
332	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	5	13	7	287
333	GlcNAca1-4Galb1-4GlcNAcb-Sp0	21	7	4	33
334	GlcNAca1-4Galb1-3GlcNAcb-Sp0	8	12	6	149
335	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	6	15	7	265
336	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	66	21	10	31
337	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	14	11	5	77
338	GlcNAca1-4Galb1-3GalNAc-Sp14	11	3	2	33
339	Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp12	8	11	5	132
340	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp12	7	14	7	203
341	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6Manb1-4GlcNAcb1-4GlcNAc-Sp12	13	5	2	37
342	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3Manb1-4GlcNAcb1-4GlcNAc-Sp12	9	2	1	19
343	Galb1-4GlcNAcb1-2Mana1-3Manb1-4GlcNAcb1-4GlcNAc-Sp12	12	6	3	53
344	Galb1-4GlcNAcb1-2Mana1-6Manb1-4GlcNAcb1-4GlcNAc-Sp12	2	10	5	400
345	Galb1-4GlcNAcb1-2Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	10	14	7	132
346	GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	19	9	4	46
347	Galb1-4GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4(Fuca1-6) GlcNAcb-Sp22	8	2	1	24
348	Galb1-3GlcNAcb1-2Mana1-3(Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4(Fuca1-6) GlcNacb-Sp22	12	10	5	81
349	Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3(Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	44	14	7	31
350	[6OSO3]GlcNAcb1-3Gal b1-4GlcNAc-b-Sp0	11	9	4	84
351	KDNa2-3Galb1-4(Fuca1-3)GlcNAc-Sp0	27	5	3	19
352	KDNa2-6Galb1-4GlcNAc-Sp0	21	15	8	71
353	KDNa2-3Galb1-4Glc-Sp0	23	9	5	40
354	KDNa2-3Galb1-3GalNAca-Sp14	27	21	10	79
355	Fuca1-2Galb1-3GlcNAcb1-2Mana1-3(Fuca1-2Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	11	13	7	118
356	Fuca1-2Galb1-4GlcNAcb1-2Mana1-3(Fuca1-2Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	8	4	2	43
357	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3(Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAb-Sp20	17	6	3	37
358	Gala1-3Galb1-4GlcNAcb1-2Mana1-3(Gala1-3Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	11	11	5	99
359	Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	7	5	3	79
360	Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3(Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	11	5	2	44
361	Neu5Aca2-6GlcNAcb1-4GlcNAc-Sp21	19	8	4	40
362	Neu5Aca2-6GlcNAcb1-4GlcNAcb1-4GlcNAc-Sp21	44	30	15	69
363	Fuca1-2Galb1-3GlcNAcb1-3(Galb1-4(Fuca1-3)GlcNAcb1-6)Galb1-4Glc-Sp21	11	8	4	73
364	Galb1-4GlcNAcb1-2(Galb1-4GlcNAcb1-4)Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp21	16	17	9	106
365	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-3(GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	15	16	8	108
366	Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-3(Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	14	9	5	67
367	Gala1-3Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3(Gala1-3Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	13	14	7	109

368	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-3(GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	9	9	5	104
369	Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-3(Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	12	10	5	90
370	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3(Fuca1-2Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	20	9	5	46
371	NeuAca2-3Galb1-4GlcNAcb1-3GalNAc-Sp14	14	16	8	115
372	NeuAca2-6Galb1-4GlcNAcb1-3GalNAc-Sp14	11	14	7	124
373	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3GalNAca-Sp14	64	24	12	37
374	GalNAcb1-4GlcNAcb1-2Mana1-6(GalNAcb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp12	20	16	8	78
375	Galb1-3GalNAca1-3(Fuca1-2)Galb1-4Glc-Sp14	9	5	2	51
376	Galb1-3GalNAca1-3(Fuca1-2)Galb1-4GlcNAc-Sp14	17	3	2	19
377	GlcNAcb1-3GalNAca-Sp14	8	12	6	150
378	GlcNAcb1-6GalNAca-Sp14	22	15	8	69
379	Galb1-3GlcNacb1-3(Galb1-3GlcNacb1-3Galb1-4GlcNacb1-6)Galb1-4Glcb-Sp0	4	12	6	270
380	Galb1-3GlcNAcb1-3(Galb1-4(Fuca1-3)GlcNAcb1-6)Galb1-4Glc-Sp21	10	10	5	97
381	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb1-3(Galb1-4GlcNAcb1-6)Galb1-4Glc-Sp21	-4	3	2	-89
382	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb1-3(Galb1-4(Fuca1-3)GlcNAcb1-6)Galb1-4Glc-Sp21	11	11	5	98
383	Galb1-3GlcNAcb1-3(Galb1-3GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-6)Galb1-4Glc-Sp21	-4	10	5	-250
384	Galb1-4GlcNacb1-2(Galb1-4GlcNacb1-4)Mana1-3(Galb1-4GlcNacb1-2(Galb1-4GlcNacb1-6)Mana1-6)Manb1-4GlcNacb1-4GlcNacb-Sp21	5	4	2	91
385	GlcNacb1-2(GlcNacb1-4)Mana1-3(GlcNacb1-2Mana1-6)Manb1-4GlcNacb1-4GlcNac-Sp21	19	10	5	55
386	Fuca1-2Galb1-3GalNAca1-3(Fuca1-2)Galb1-4Glcb-Sp0	9	13	6	136
387	Fuca1-2Galb1-3GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	9	11	6	120
388	Galb1-3GlcNAcb1-3GalNAca-Sp14	15	16	8	110
389	Neu5Aca2-3(GalNAcb1-4)Galb1-4GlcNAcb1-3GalNAca-Sp14	13	24	12	181
390	GalNAca1-3(Fuca1-2)Galb1-3GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	13	13	7	104
391	Gala1-3Galb1-3GlcNAcb1-2Mana1-3(Gala1-3Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp19	37	11	5	29
392	Gala1-3Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3(Gala1-3Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp19	35	20	10	59
393	Neu5Aca2-3Galb1-3GlcNAcb1-2Mana1-3(Neu5Aca2-3Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp19	6	6	3	98
394	Galb1-4GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp12	13	11	5	80
395	GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp12	5	9	4	169
396	Neu5Aca2-3Galb1-3GlcNAcb1-3GalNaca-Sp14	-4	7	4	-207
397	Fuca1-2Galb1-4GlcNacb1-3GalNaca-Sp14	16	23	11	144
398	Galb1-4(Fuca1-3)GlcNacb1-3GalNaca-Sp14	16	10	5	65
399	GalNaca1-3GalNacb1-3Gala1-4Galb1-4GlcNacb-Sp0	14	5	3	38
400	Gala1-4Galb1-3GlcNacb1-2Mana1-3(Gala1-4Galb1-3GlcNacb1-2Mana1-6)Manb1-4GlcNacb1-4GlcNacb-Sp19	9	4	2	38
401	Gala1-4Galb1-4GlcNacb1-2Mana1-3(Gala1-4Galb1-4GlcNacb1-2Mana1-6)Manb1-4GlcNacb1-4GlcNacb-LVaNKT	23	18	9	77
402	Gala1-3Galb1-4GlcNacb1-3GalNaca-Sp14	7	14	7	182
403	Galb1-3GlcNAcb1-6Galb1-4GlcNAcb-Sp0	17	9	4	51
404	Galb1-3GlcNAca1-6Galb1-4GlcNAcb-Sp0	12	9	4	75
405	GalNAcb1-3Gala1-6Galb1-4Glcb-Sp8	5	11	6	221
406	GlcNAcb1-6(GlcNAcb1-3)GalNAca-Sp14	2	6	3	414

Chapter 4

Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing

de Jonge R*, van Esse HP*, Maruthachalam K, Bolton MD, Santhanam P, Keykha Saber M, Zhang Z, Usami T, Lievens B, Subbarao KV and Thomma BPHJ (2012) **Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing.** *Proc Natl Acad Sci USA* **109**:5110-5115.
*These authors contributed equally to this research

Abstract

Fungal plant pathogens secrete effector molecules to establish disease on their hosts, while plants in turn utilize immune receptors to try and intercept these effectors. The tomato immune receptor Ve1 governs resistance to race 1 strains of the soil-borne vascular wilt fungi *Verticillium dahliae* and *V. albo-atrum*, but the corresponding *Verticillium* effector remained unknown thus far. By high-throughput population genome sequencing, a single 50 Kb sequence stretch was identified that only occurs in race 1 strains, while subsequent transcriptome sequencing of *Verticillium*-infected *Nicotiana benthamiana* plants revealed only a single highly expressed ORF in this region, designated *Ave1* (for *Avirulence on Ve1 tomato*). Functional analyses confirmed that Ave1 activates Ve1-mediated resistance and demonstrated that Ave1 markedly contributes to fungal virulence, not only on tomato but also on *Arabidopsis*. Interestingly, Ave1 is homologous to a widespread family of plant natriuretic peptides (PNPs). Besides plants, homologous proteins were only found in the bacterial plant pathogen *Xanthomonas axonopodis* and the plant pathogenic fungi *Colletotrichum higginsianum*, *Cercospora beticola* and *Fusarium oxysporum* f. sp. *lycopersici*. The distribution of Ave1 homologs, coincident with the presence of *Ave1* within a flexible genomic region, strongly suggests that *Verticillium* acquired *Ave1* from plants through horizontal gene transfer. Remarkably, by transient expression we show that also the Ave1 homologs from *F. oxysporum* and *C. beticola* can activate *Ve1*-mediated resistance. In line with this observation, *Ve1* was found to mediate resistance towards *F. oxysporum* in tomato, showing that this immune receptor is involved in resistance against multiple fungal pathogens.

Introduction

Throughout evolution, microbial pathogenicity towards plant hosts independently emerged on multiple occasions in diverse taxa harboring plant-associated microbes, including bacteria, oomycetes and fungi [1]. At the same time, plant genomes evolved to encode immune receptors that sense the presence of various types of microbial invaders by detection of the presence of microbial molecules or their plant-manipulating activities [2-4]. Cell surface receptors, referred to as pattern recognition receptors (PRR), detect conserved microbial molecules, referred to as microbe-associated molecular patterns (MAMPs), to activate MAMP-triggered immunity (MTI). Successful plant pathogens overcome MTI by the use of secreted effectors, many of which have molecular targets inside host cells, which perturb host defenses in a pro-active manner [4,5]. In turn, plants evolved to intercept the activity of particular pathogen effectors through novel receptors that are generally referred to as resistance proteins. While some of these have been characterized as cell surface receptors, most of them are cytoplasmic proteins of the nucleotide binding leucine-rich repeat (NB-LRR) type that again activate inducible host defenses, referred to as effector-triggered immunity (ETI) [4,6]. Nevertheless, the delineation between MAMPs and effectors, as well as between MTI and ETI, is blurred and rather a continuum [3].

The acquisition of particular effector genes in microbial genomes has resulted in emergence of

pathogenicity, or in host range expansion (7-9). Novel effectors can be acquired in various ways, including gene duplication and subsequent diversification. Expansion of effector families is especially striking in plant pathogenic oomycete species that harbor large repertoires of RXLR and Crinkler effectors [9-11]. Substantial expansion of effector gene families has also been observed in the genomes of the fungal plant pathogens *Ustilago maydis* and *Blumeria graminis* [12-14]. Interestingly, effector genes are frequently found in regions that are enriched for transposable elements that may provide a mechanism for amplification and diversification of effectors in pathogen genomes [9, 14,15]. Novel effectors can also be acquired through horizontal gene transfer (HGT), which involves the transmission of genetic material across species boundaries. The extent to which HGT contributes to genome evolution in eukaryotes is not clear, but multiple reports have proposed that HGT occurred regularly among eukaryotic plant pathogens [8,16-18]. Moreover, recent evidence for frequent HGT events between fungi and oomycetes suggests that HGT facilitated the evolution of plant parasitism in oomycetes [17].

Verticillium dahliae is an asexual soil-borne, xylem-invading, fungal plant pathogen that is responsible for vascular wilt diseases in over 200 dicotyledonous plant species, including economically important crops such as tomato [19,20]. A typical infection starts by penetration of the root, after which the fungus enters the xylem and starts to produce conidia which are carried with the water flow to distal plant parts [19]. In only a few plant species, monogenic sources of resistance towards Verticillium wilt have been described, including the *Ve* locus from tomato that controls race 1 *V. dahliae* and *V. albo-atrum* strains [21-24]. The resistance is mediated by the *Ve1* gene that encodes a predicted receptor-like protein-type cell surface receptor [23,25]. Strains that are not contained by the *Ve* locus are assigned to race 2 and are generally less aggressive on tomato plants that lack *Ve1* when compared with race 1 strains [26,27]. This suggests that Ve1 recognizes a virulence factor in race 1 strains that is absent in race 2 strains.

Various methods have been used to identify pathogen effectors that activate host immune receptors. Whereas in sexually propagating fungi genetic mapping can be used, in asexual fungi most approaches are based on functional screens for a hypersensitive response (HR); tissue necrosis as culmination of a strong immune reaction [28-30]. So far, attempts to identify the *V. dahliae* effector that triggers Ve1-mediated resistance have been unsuccessful. In this study, we performed a novel comparative population genomics approach, by applying high-throughput population genome sequencing, to identify the *V. dahliae* effector that activates tomato Ve1.

Results

Comparative population genomics identifies *Verticillium* effector Ave1

Recently, the genome of the *V. dahliae* race 2 strain VdLs.17 was sequenced using Sanger technology, and determined to be ~34 Mb with ~10,500 predicted genes [31]. In this study, we determined the genome sequences of ten *V. dahliae* strains; four of which belonged to race 1 and six to race 2. For each strain, ~11 million paired-end Illumina reads, representing a predicted 30X genome coverage based on the VdLs.17 reference genome sequence, were *de novo* assembled into draft genomes of ~34 Mb (Table 1).

The completeness of the genomes was assessed by the core eukaryotic genes mapping approach [32]. We subsequently aligned race 1 scaffold and contig sequences with all race 2 sequences, including the VdLs.17 reference genome, and all unaligned race 1 sequences were retained. This revealed a small number of race 1 scaffolds that were larger than 1 Kb and that did not align to race 2 sequences. Further comparisons between the race 1-specific sequences revealed a single 50 Kb region that was shared by all race 1 strains (Figure 1) and that contains 68 predicted ORFs (>180 nucleotides), including 10 that encode putative secreted effectors.

In order to validate the bioinformatic ORF prediction in the 50 Kb race 1-specific region, deep RNA sequencing was performed on a time course of *Nicotiana benthamiana* plants infected by race 1 strain JR2 (Table S1). For each sample, ~25 million paired-end reads were mapped onto the JR2 genome. While over 8,000 *V. dahliae* genes were expressed, reads mapped only to a single locus in the 50 Kb race 1-specific region that was called *Ave1*, for *Avirulence on Ve1* tomato. RACE PCR experiments confirmed that the *Ave1* gene model spans 582 bp and comprises two exons that are interrupted by an intron in the 5'UTR (Figure 1; Figure S1*A*). *Ave1* encodes a predicted 134 amino acid secreted (D>0.8) protein, and based on the RNA-Seq reads it was determined that *Ave1* expression is induced during host colonization, a characteristic of typical effector proteins (Figure S1*B*) [5,33].

Verticillium Ave1 is a virulence factor that activates Ve1-mediated resistance

We subsequently performed functional analyses to prove that Ave1 is the *V. dahliae* effector that is recognized by tomato Ve1. Firstly, heterologous expression of *Ave1* using Potato Virus X resulted in a hypersensitive response (HR) only on *Ve1* tomato (Figure S2). This recognition was confirmed by *Agrobacterium tumefaciens*-mediated transient expression assays in *Nicotiana tabacum*, showing that co-expression of *Ave1* with *Ve1*, but not with *Ve2*, resulted in HR (Figure 2*A*). We subsequently performed genetic deletion and complementation experiments in *V. dahliae* to confirm the role of Ave1 in activating disease resistance. As expected, targeted deletion of *Ave1* in race 1 *V. dahliae* strain JR2 resulted in gain of virulence on *Ve1* tomato (Figure 3*A*), while subsequent complementation of the deletion strains using a genomic fragment

Table 1. Assembly statistics of *V. dahliae* genome sequences

Strain	Origin	Race	Assembly size (Mb)	Scaffold N50 (Kb)	# of scaffolds	%CEGMA partial (complete)*
JR2	Tomato	1	35,1	59,4	4753	95,2 (87,5)
St14.01	Pistachio	1	34,7	65,9	3684	95,6 (89,9)
CBS381.66	Tomato	1	34,5	48,3	4411	94,0 (87,1)
VdLs16	Lettuce	1	34,9	68,0	3469	96,4 (92,3)
DVD-3	Potato	2	34,1	43,9	9318	91,5 (81,0)
DVD-31	Tomato	2	34,0	36,9	4513	93,6 (86,3)
DVD-161	Potato	2	34,1	42,4	4078	94,0 (86,7)
DVD-S26	Soil	2	35,3	47,1	5361	94,4 (87,9)
DVD-S29	Soil	2	34,1	49,2	3712	92,7 (86,3)
DVD-S94	Soil	2	35,0	57,4	4073	95,6 (89,1)
VdLs.17	Lettuce	2	33,8	1.270	52	97,6 (94,8)

*percentage of partial and complete gene models determined by the Core Eukaryotic Genes Mapping Approach (CEGMA)

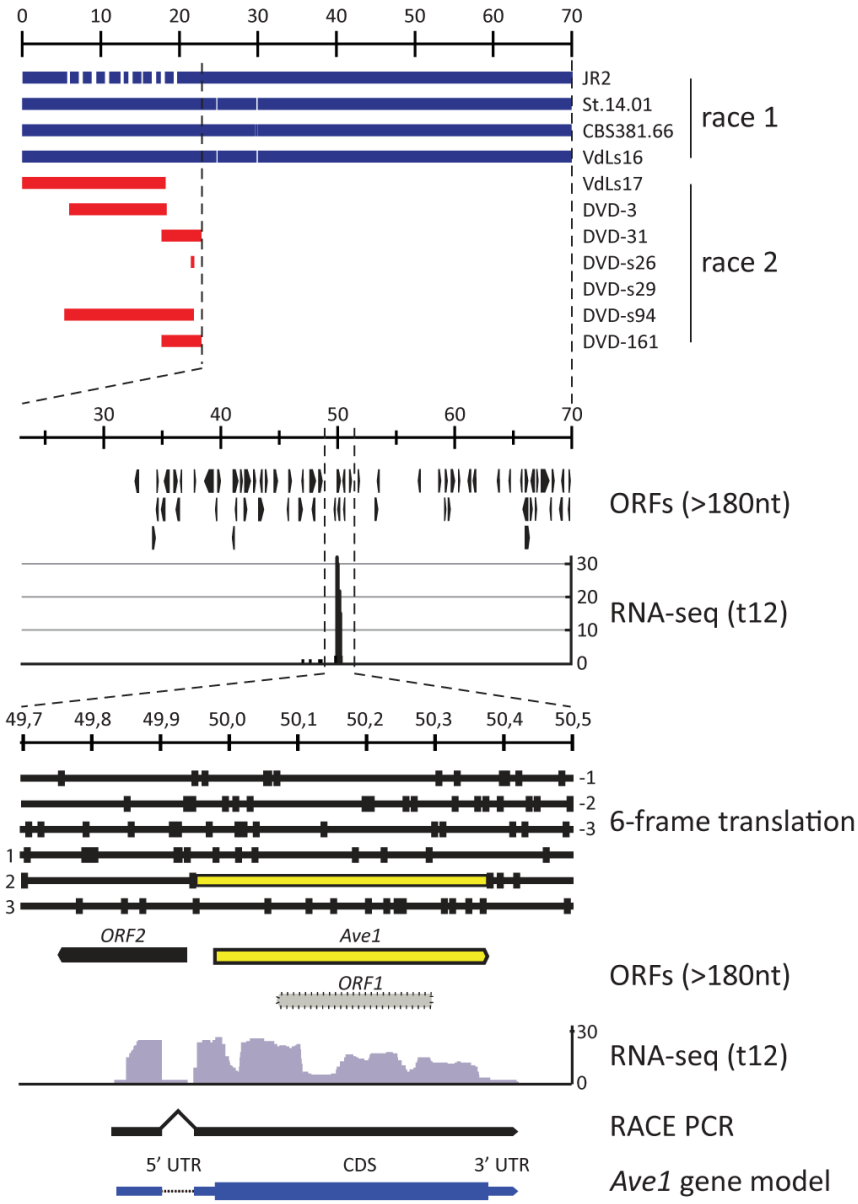


Figure 1. *Verticillium* comparative population genomics and transcriptome sequencing identifies race 1-specific effector *Ave1*. Alignment of race 1 (blue) and race 2 (red) contigs outlining a 50 Kb race 1-specific region comprising 68 ORFs >180 nucleotides. RNA sequencing reads mapped only to the Ave1 locus, identifying two exons split by a 55 bp 5' UTR intron. The Ave1 gene model was confirmed by RACE PCR.

including 1.5 Kb up- and downstream of the *Ave1* coding sequence (*pAve1::Ave1*) restored avirulence on *Ve1* tomato (Figure S3*A*). In addition, complementation of the *V. dahliae* race 2 strains VdLs.17 and DvdS26 with *pAve1::Ave1* resulted in loss of virulence on these plants (Figure 3*B*). Collectively, these experiments provide solid evidence for a role of Ave1 as elicitor of disease resistance mediated by the Ve1 immune receptor in tomato.

According to the paradigm that plant immune receptors intercept pathogen virulence factors, it was expected that Ave1 acts as a virulence factor on tomato plants lacking *Ve1*. To test this hypothesis, *Ave1* deletion strains were inoculated on *ve1* tomato plants, showing that *Ave1* deletion strains displayed markedly reduced aggressiveness on tomato plants lacking *Ve1* (Figure 3*A*, Figure S3*A*). When compared with the wild-type fungus, inoculation with *Ave1* deletion strains resulted in reduced stunting and fungal colonization (Figure S3*B*). Conversely, complementation of race 2 strains and *Ave1* deletion strains with

pAve1::Ave1 resulted in significantly increased virulence on tomato plants lacking *Ve1* (Figure 3*B*, Figure S3*A*).

We have recently shown that *Ve1* remains fully functional after interfamily transfer to the Brassicaceous model plant *Arabidopsis thaliana*, as *Ve1*-transgenic Arabidopsis is resistant to race 1 but not to race 2 strains of *V. dahliae* and *V. albo-atrum* [24]. To confirm that Ave1 activates Ve1-mediated resistance in Arabidopsis, we inoculated the *Ave1* deletion strains along with the corresponding wild type race 1 *V. dahliae* on wild type and *Ve1*-transgenic Arabidopsis plants (Figure S4*A*). Whereas *Ve1*-expressing Arabidopsis were resistant to the wild-type race 1 strain, the resistance was broken upon targeted deletion of *Ave1*. As expected, resistance was restored by complementation of the *Ave1* deletion strains with *pAve1::Ave1*. Our results confirm that in Arabidopsis, similar to tomato, Ve1-mediated race 1 resistance is activated by Ave1 (Figure S4*A*). Interestingly, complementation with *pAve1::Ave1* enhanced the virulence of a race 2 strain on Arabidopsis plants, demonstrated by a significant increase in fungal colonization, suggesting that Ave1 acts as a virulence factor also on Arabidopsis (Figure S4*B*).

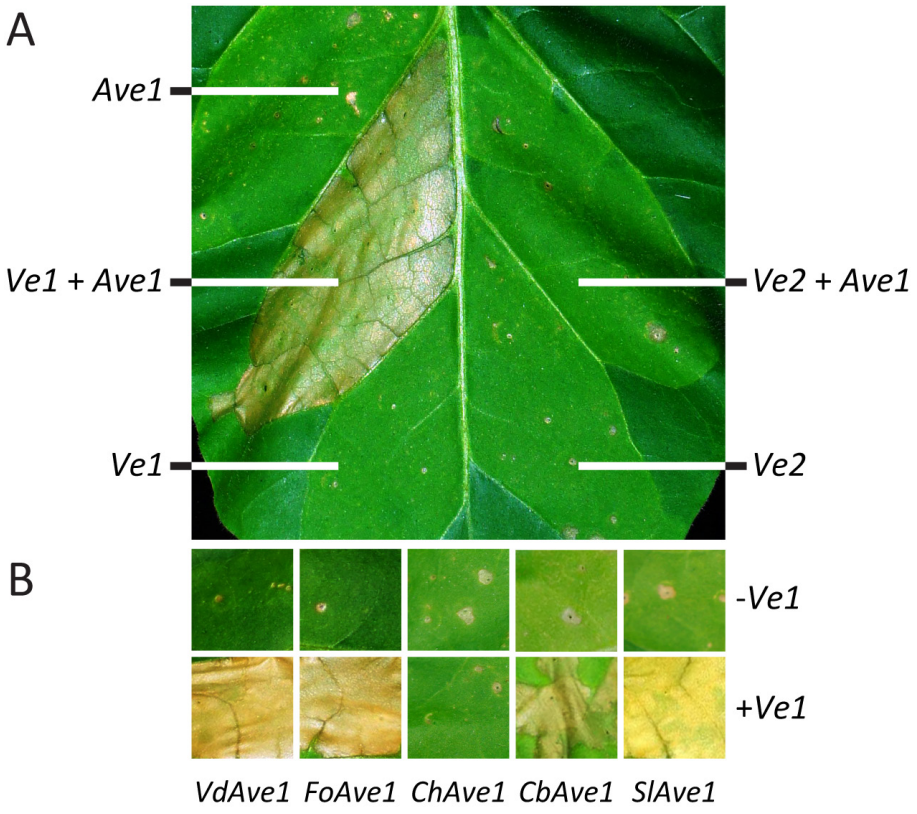


Figure 2. Co-expression of *Ave1* and *Ve1* in *Nicotiana tabacum* activates a hypersensitive response. A) *Verticillium dahliae Ave1* was transiently co-expressed with tomato *Ve1* and *Ve2* in *N. tabacum*. As a negative control, *Ve1*, *Ve2* and *Ave1* were expressed separately. B) *Ave1* homologs of *V. dahliae* (VdAve1), *Fusarium oxysporum* f. sp. *lycopersici* (*FoAve1*), *Colletotrichum higginsianum* (*ChAve1*), *Cercospora beticola* (*CbAve1*) and tomato (*SlAve1*) were co-expressed with tomato *Ve1* in *N. tabacum*. As a negative control, *Ave1* homologs were expressed separately. Leaves were photographed at 5 days after infiltration to visualize cell death resulting from recognition by Ve1.

Absence of *Ave1* allelic variation in a collection of *Verticillium* strains

To analyze *Ave1* diversity, we sequenced 85 alleles from *Verticillium* strains isolated from various host plants and different geographical locations (Table S2). Intriguingly, not a single nucleotide polymorphism (SNP) was found in the 85 *Ave1* alleles tested. Interestingly, an *Ave1* allele was amplified from the sequenced *V.*

albo-atrum strain VaMs102 [31] that, based on blast analysis, was thought not to contain *Ave1*. Likely, *Ave1* is lacking in the genome assembly as a consequence of the low coverage of sequencing [31]. The finding that also *V. albo-atrum Ave1* alleles are identical is remarkable as *V. dahliae* and *V. albo-atrum* share only 92% nucleotide sequence identity, with only 0.3% identical genes [31]. As expected, *Ave1* alleles were not identified in any of the 19 *Verticillium* race 2 strains analyzed nor in the 32 *V. dahliae* and 3 *V. albo-atrum* strains that are not pathogenic on tomato (Table S2).

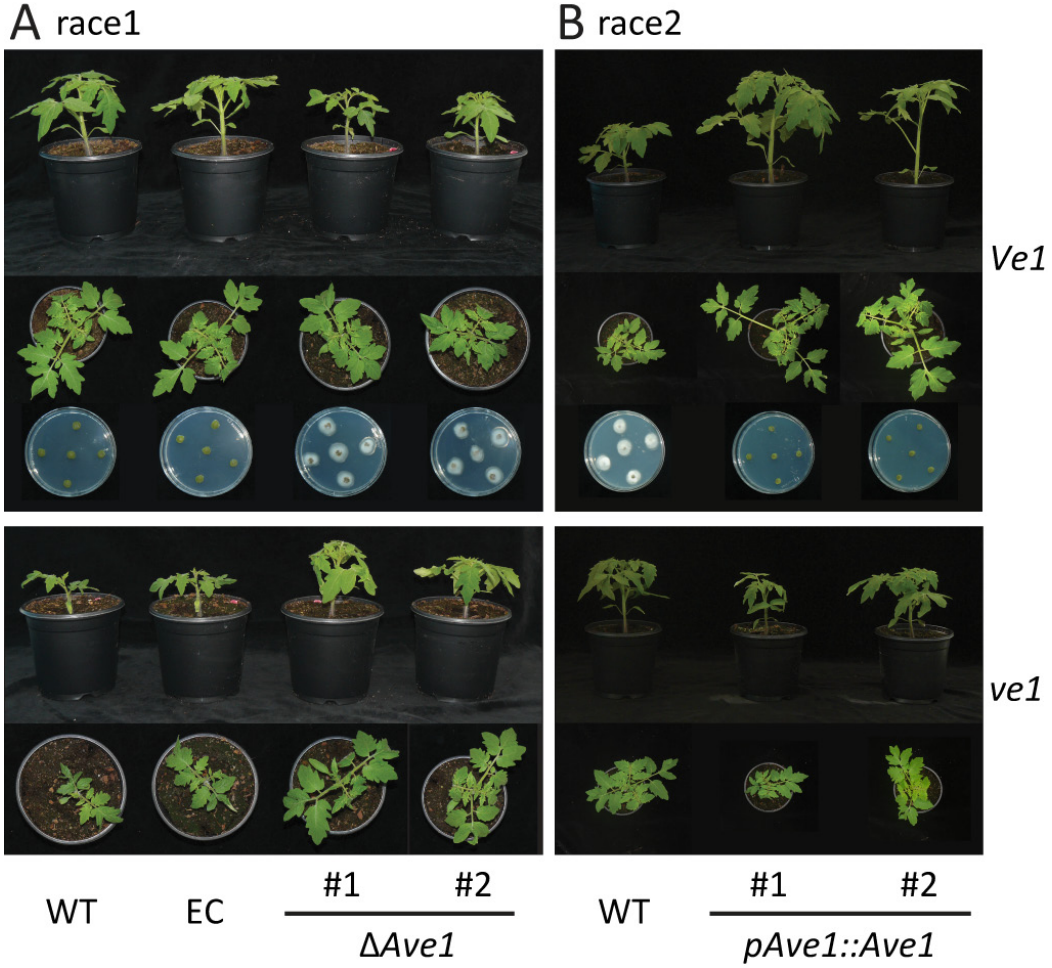


Figure 3. *Verticillium dahliae* Ave1 activates the tomato immune receptor Ve1 and enhances virulence on susceptible tomato. A) Top: *Ave1* deletion strains escape recognition by *Ve1* tomato when compared with wild-type (WT) and ectopic transformant (EC) evidenced by stunted *Ve1* plants at 14 dpi and fungal outgrowth upon plating of stem sections. Bottom: *Ave1* deletion strains show compromised virulence on tomato lacking *Ve1*. Pictures for two independent *Ave1* deletion strains are shown. B) Top: *Ave1* expression in race 2 *V. dahliae* (*pAVe1::Ave1*) results in recognition by *Ve1* tomato at 14 dpi. Bottom: Race 2 p*Ave1::Ave1* transformants show enhanced virulence on tomato lacking *Ve1*. Pictures for two independent *Ave1*-transgenic strains are shown.

Gene distribution strongly suggests that *Ave1* was horizontally acquired from plants

Interestingly, not a single fungal homolog of *V. dahliae* Ave1 was identified in BLASTp analysis. Remarkably, however, over 200 Ave1 homologs were identified in plants. In addition, an Ave1 homolog has previously been identified as the virulence factor XacPNP in the plant pathogenic bacterium *Xanthomonas axonopodis* pv. *citri*, causal agent of citrus canker [34]. Further in-depth analysis with tBLASTn revealed an unannotated *Ave1* homolog in the genome of the tomato pathogenic, xylem-invading fungus *Fusarium oxysporum* f. sp. *lycopersici*, designated *FoAve1*, and two homologs in the genomes of the fungal pathogens *Colletotrichum*

higginsianum (Broad Institute) and *Cercospora beticola* (M. Bolton, unpublished data), designated *ChAve1* and *CbAve1*, respectively.

To assess the evolutionary relationships between the various Ave1 homologs, *V. dahliae* Ave1 (VdAve1) was aligned with FoAve1, ChAve1, CbAve1, XacPNP and the 50 most homologous plant proteins (Figure S5). Phylogenetic analysis applying maximum likelihood (ML) indicated that VdAve1 shares common ancestry with ChAve1, CbAve1 and five closely related plant proteins from the taxonomically diverse species grape (*Vitis vinifera*), castor bean (*Ricinus communis*) and tomato (*Solanum lycopersicum*), as well as

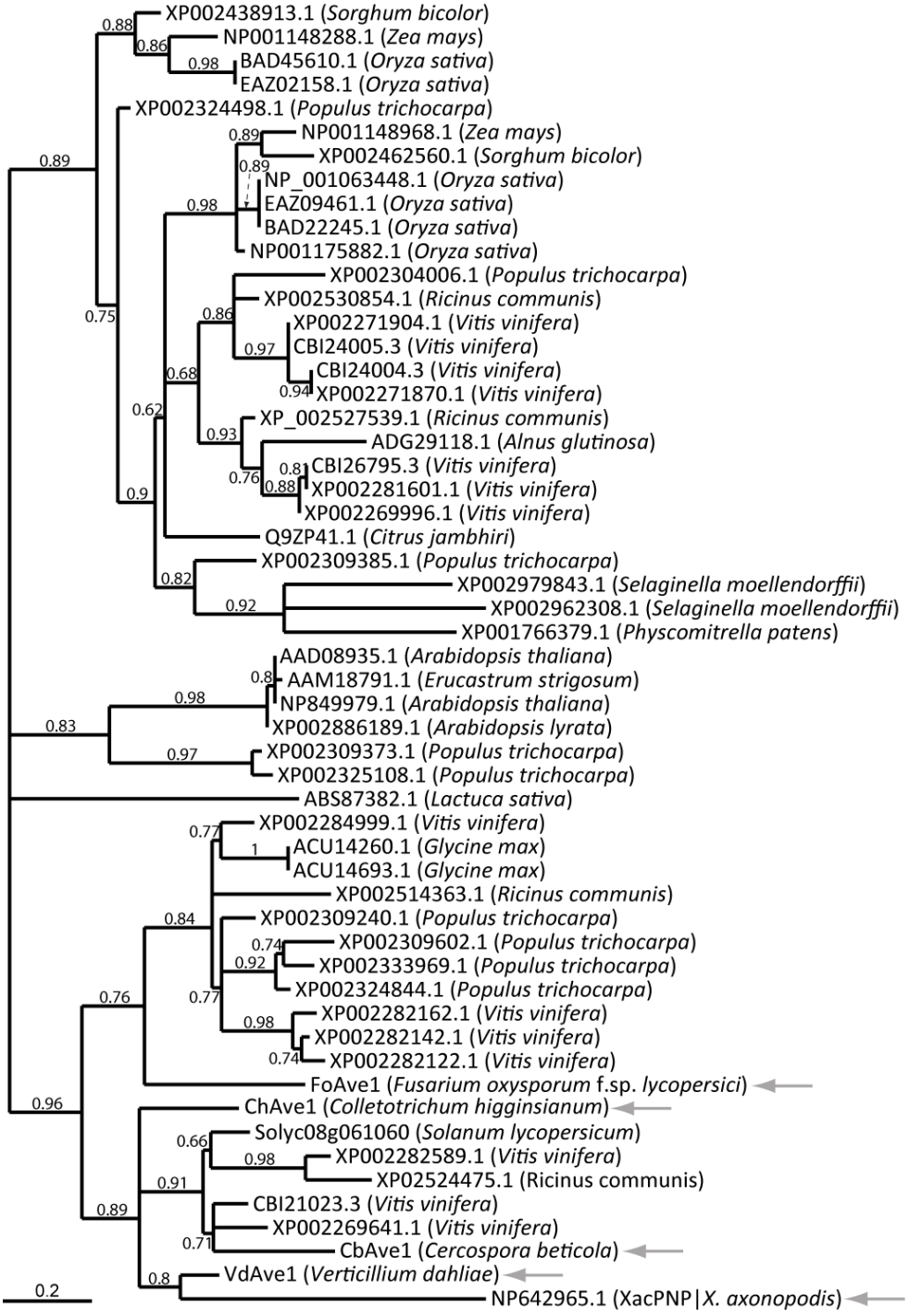


Figure 4. Evolutionary relationship of Ave1 homologs from *Verticillium dahliae* (VdAve1), *Colletotrichum higginsianum* (ChAve1), *Cercospora beticola* (CbAve1), *Fusarium oxysporum* f. sp. *lycopersici* (FoAve1), *Xanthomonas axonopodis* (XacPNP) (indicated by arrows) and 50 related plant-derived proteins, determined using maximum likelihood (ML) analyses. Branch lengths are proportional to phylogenetic distances, and the result of the approximate likelihood ratio test is given at the nodes to indicate branch support.

with XacPNP, although the latter protein is significantly divergent (Figure 4). FoAve1 clusters in a distinct clade that contains 11 proteins from poplar (*Populus trichocarpa*), soybean (*Glycine max*), grape and castor bean (Figure 4). All proteins share four cysteine residues that are likely involved in disulphide bridges that contribute to protein stability upon secretion (Figure S5).

It has previously been suggested that *X. axonopodis* pv. *citri* acquired XacPNP from plants by horizontal gene transfer [34]. The abundance of Ave1 orthologs in plants, combined with the absence of orthologous sequences in fungi other than *F. oxysporum* f. sp. *lycopersici*, *C. higginsianum* and *C. beticola* similarly suggests that *Verticillium* horizontally acquired *Ave1* from plants. Robust phylogenetic analysis reveals evolutionary relationships between Ave1 homologs that contradict species phylogeny, which is generally considered as evidence for horizontal gene transfer (HGT; [17,18]) (Figure 4; Figure S5). Additional evidence for HGT can be found in the genomic context of *Ave1*. The recent genome comparison between *V. dahliae* strain VdLs.17 and the highly homologous *V. albo-atrum* strain VaMs102 revealed 4 lineage-specific regions (LS1-LS4) that are absent in VaMs102 [31]. These regions are highly enriched in transposable elements, supporting their plasticity [31]. Interestingly, the race 1-specific region harboring *Ave1* is physically associated with LS3 (Figure S6*A*) and is characterized by a *Ty1-copia* retro-transposon immediately adjacent to *Ave1* and variability in GC content (Figure S6*A*).

FoAve1 is located on chromosome 14 of the *F. oxysporum* f. sp. *lycopersici* genome; a lineage-specific chromosome that is proposed to be responsible for pathogenicity towards tomato [35]. Various transposable elements flank *FoAve1* (Figure S6*B*).

FoAve1 is restricted to *F. oxysporum* f. sp. *lycopersici*

F. oxysporum as a species includes morphologically indistinguishable pathogenic as well as non-pathogenic strains. In spite of the broad host range of the species, individual strains typically infect only a single or a few plant species and are assigned to *formae speciales* based on host specificity. To investigate whether Ave1 is restricted to the *formae specialis lycopersici*, we assessed the presence of *FoAve1* in other *formae speciales* of *F. oxysporum*. However, *FoAve1* was exclusively detected in tomato pathogenic *F. oxysporum* f. sp. *lycopersici* strains (Table S3). We subsequently assessed the allelic variation of *FoAve1*. In 72 *F. oxysporum* f. sp. *lycopersici* strains tested, *FoAve1* was identified and determined to be identical.

Homologs of *V. dahliae* Ave1 are recognized by Ve1

We have previously argued that the Ve1 receptor shares traits with MAMP receptors such as CEBiP, CERK1, FLS2 and EFR [3,24]. The identification of Ave1 homologs in a number of fungal pathogens allowed testing this hypothesis. Intriguingly, *A. tumefaciens*-mediated co-expression of *Ve1* with *FoAve1* and CbAve1, but not with ChAve1 in *N. tabacum* induced HR, demonstrating that tomato Ve1 recognizes Ave1 homologs from four distinct fungal pathogenic species; *V. dahliae*, *V. albo-atrum*, *F. oxysporum* and *C. beticola* (Figure 2*B*). Remarkably, co-expression of *Ve1* with the *Ave1* homolog from tomato, *SlAve1*, in *N. tabacum* also induced HR (Figure 2*B*).

The finding that co-expression of *Ve1* with the *Ave1* homolog from the tomato pathogen *F. oxysporum* f. sp. *lycopersici*, *FoAve1*, induces HR allowed to test whether *Ve1* confers resistance also to this pathogen. Therefore, we inoculated non-transgenic MoneyMaker (LA2706) tomato plants that lack resistance against *V. dahliae* and *F. oxysporum*, and *Ve1* transgenes [21] with *F. oxysporum* f. sp. *lycopersici* (Figure 5). A clear disease reduction was observed on *Ve1* plants, demonstrating that *Ve1* confers resistance to *F. oxysporum* f. sp. *lycopersici*.

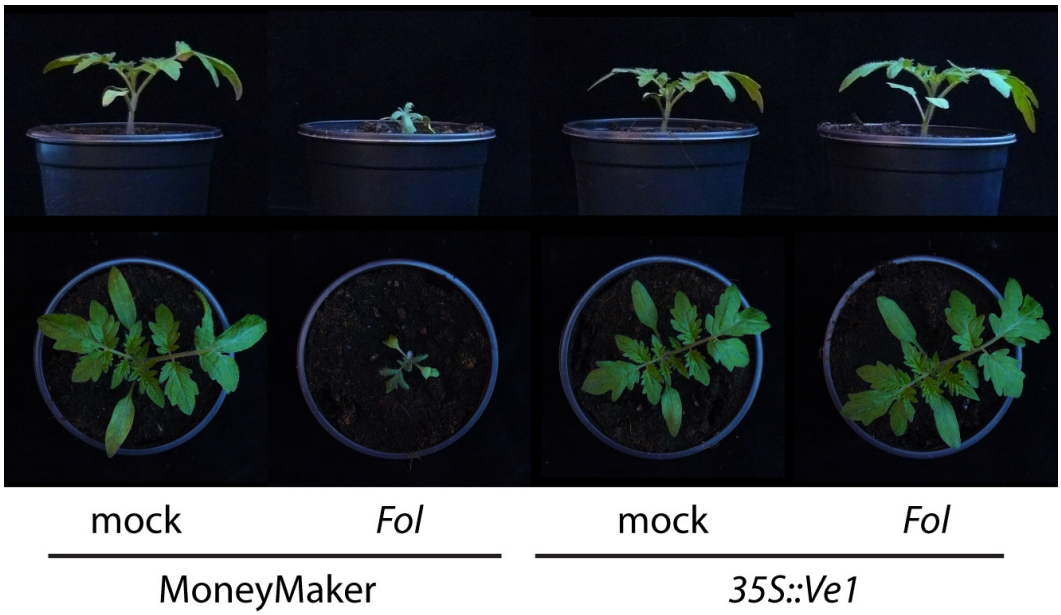


Figure 5. Immune receptor Ve1 controls infection of *Fusarium oxysporum* f. sp. *lycopersici* in tomato. Side view (upper row) and top view (lower row) of non-transgenic (MoneyMaker) and *Ve1*-transgenic (*35S::Ve1*) tomato plants at 13 days post mock-inoculation (mock) or inoculation with *F. oxysporum* f. sp. *lycopersici* (*Fol*).

Discussion

In tomato, resistance against race 1 strains of the vascular fungi *V. dahliae* and *V. albo-atrum* is mediated by the cell surface receptor-like protein Ve1 [21]. Unfortunately, traditional approaches employed in the past to identify the *Verticillium* effector that activates Ve1-mediated resistance in tomato, including the biochemical characterization of protein fractions that induce necrosis in resistant plants [28-30] and heterologous *in planta* expression of pathogen cDNA libraries [37], were unsuccessful. In this study, we employed a novel approach to identify the avirulence protein that corresponds to Ve1, making use of high-throughput sequencing. To this end, we sequenced the genomes of multiple *Verticillium* race 1 and race 2 strains. Comparative analyses revealed only a single 50 Kb sequence stretch that was specifically present in race 1 strains, containing only a single open reading frame that was highly expressed *in planta*. Functional analysis of this locus, named *Ave1*, confirmed that it encodes the effector that is recognized by Ve1. Thus, our study shows that population genomics can be used as a powerful tool for the identification of novel avirulence components in complex fungal genomes.

Pathogen effectors are typically lineage-specific, meaning that generally no homologs occur in other species, and often not even in all strains of the same species [5]. Thus, it was expected that homologs could not be found in other fungal species. Surprisingly, BlastP analyses identified many Ave1 homologs in plants,

several of which are annotated as expansin-like proteins that share a conserved family-45 endoglucanase (EG45-like) domain with cell wall-loosening expansins [38]. Other Ave1 homologs are characterized as plant natriuretic peptides (PNPs; [39]). Natriuretic peptides (NPs) were originally identified in vertebrates where they have been implicated in the maintenance of osmotic and cardiovascular homeostasis [40]. In plants, PNPs are mobile signaling molecules that are secreted in the apoplast, particularly under conditions of biotic and abiotic stress, and that play an important role in the regulation of water and ion homeostasis and consequently can affect many downstream processes, including photosynthesis [39,41]. Our analyses have shown that Ave1 acts as a potent virulence factor of *Verticillium*, not only in tomato plants that lack the Ve1 resistance protein, but also in Arabidopsis. Possibly, modulation of water and ion homeostasis by Ave1 increases the sap stream in the xylem, leading to accelerated host colonization. In addition to the many plant homologs, a homolog was identified in the citrus canker pathogen, *X. axonopodis* pv. *citri*, that was previously characterized as a bacterial virulence factor [34,42]. XacPNP is thought to mimic PNPs by manipulating the physiology of the host, including water homeostasis, stomatal opening and photosynthesis to promote bacterial proliferation [42]. The presence of numerous Ave1 orthologs in plants, absence of orthologs in fungi other than *F. oxysporum* f. sp. *lycopersici*, *C. higginsianum* and *C. beticola*, and the association of *Ave1* with a flexible genomic region containing various transposable elements (TEs) in the genome strongly suggest that *Verticillium* acquired *Ave1* from plants through horizontal gene transfer (HGT). Despite similar ancestry of VdAve1, ChAve1, CbAve1 and XacPNP, as suggested by phylogenetic analysis, direct transfer between *Verticillium*, *C. higginsianum*, *C. beticola* and *X. axonopodis* is unlikely as these plant pathogens infect different hosts, occupy distinct niches within these hosts, and are thus unlikely to encounter each other.

Eventually, in depth analyses revealed unannotated *Ave1* homologs in the genomes of the plant pathogenic fungi, *F. oxysporum* f. sp. *lycopersici*, *C. higginsianum* and *C. beticola*. *Verticillium*, *F. oxysporum* and *C. higginsianum* belong to the class of Sordariomycetes whereas *C. beticola* belongs to the Dothideomycetes. Both classes comprise many other plant pathogens with sequenced genomes such as the Sordariomycetes *F. graminearum*, *F. solani*, *F. verticillioides, C. graminicola* and *Magnaporthe oryzae* and the Dothideomycetes *Mycosphaerella graminicola* and *Leptosphaeria maculans* (15, 35, 43-46). *Ave1* is not found in these close relatives, nor was it detected in *F. oxysporum formae speciales* other than *lycopersici*. Interestingly, our phylogenetic analysis identified distinct origins for *VdAve1*, *ChAve1* and *CbAve1* on the one hand, and *FoAve1* on the other hand, suggesting independent HGT events. Independence of the HGT events is further supported by different TEs flanking the *Ave1* loci in *Verticillium* and *Fusarium* and the absence of Ave1 homologs in closely related fungi. Recently, a large phylogenomic analysis involving 6 plants species and 46 fungal species identified 4 plant-to-fungus HGTs, suggesting that genetic exchange between plants and fungi occurs more often [16].

We have previously noted that Ve1 has traits of a pattern recognition receptor (PRR) that acts in MAMP-triggered immunity [3,24]. This was based on the observation that Ve1 resistance affects two fungal species, *V. dahliae* and *V. albo-atrum*, involvement of the PRR co-receptor BAK1/SERK3 in Ve1 signaling,

the relatively weak nature of Ve1-mediated resistance, the existence of Ve1 homologs in various plant families, and the transferability of *Ve1* across plant families. Our present evidence showing that FoAve1 and CbAve1 are also recognized by Ve1, and that *Ve1*-expressing tomato is resistant to *F. oxysporum*, further substantiates its role as a PRR, and further adds to the notion that PRRs and R proteins cannot strictly be separated and should be considered as a continuum [3].

The *Ave1* gene is fully conserved in all race 1 *Verticillium* strains that were tested, suggesting that identical alleles are required for maximum virulence. Deletion of *Ave1* from the genome imposes a significant virulence penalty, as *Ave1* acts as a virulence factor not only on tomato, but also on Arabidopsis. The absence of *Ave1* in race 2 *Verticillium* strains explains earlier observations that race 1 *Verticillium* strains are more aggressive on *ve1* tomato than race 2 strains [26,27]. *FoAve1*, like *Ave1,* is fully conserved in all *F. oxysporum* f. sp. *lycopersici* tested. This suggests that FoAve1 is crucial for the virulence of *F. oxysporum* f. sp. *lycopersici*.

Materials and Methods

Verticillium genomics

Verticillium dahliae genomic DNA was isolated from conidia that were harvested from 10-day-old cultures grown on potato dextrose agar. Library preparation (500 bp inserts) and Illumina sequencing (100 bp paired-end reads) was performed at the Beijing Genome Institute (BGI, Hong Kong). Draft genome assemblies and the VdLs.17 reference assembly [31] were compared all versus all by MUMMER3 [47] to identify race 1-specific sequences as described in the *SI Material and Methods*.

For deep transcriptome sequencing, three-week-old *N. benthamiana* plants were inoculated with strain JR2 as previously described [21], harvested at 4, 8, 12 and 16 days post inoculation and flash frozen in liquid nitrogen. Total RNA was extracted using the RNeasy Mini Kit (Qiagen). cDNA synthesis, library preparation (200 bp inserts) and Illumina sequencing (90 bp paired-end reads) was performed at BGI and the obtained reads were mapped on the draft JR2 genome using Tophat [48] as described in the *SI Materials and Methods*.

Ave1 functional analysis

For heterologous expression, we cloned *VdAve1* in the binary pSfinx vector [29] and performed *Agrobacterium tumefaciens*-mediated transformation on tomato plants [47]. For constitutive expression, *Ave1* homologs were cloned in the modified, Gateway compatible, pBIN variant pSol2092, and *Ve1* and *Ve2* were used in pEarleyGate100 and pMOG800 [21,49]. *A. tumefaciens*-mediated transformation of *N. benthamiana* was performed as described previously [50]. *Ave1* knock-outs in *V. dahliae* were generated by cloning of the *Ave1* flanking sequences in pRF-HU2 [51]. For genomic complementation *Ave1* and flanking sequences were cloned in pRW1P [52]. Also see *SI Materials and Methods*.

Ave1 protein sequence analysis

Ave1 homologs were identified in public databases by Blast analyses (SI Materials and Methods, Table S4), and phylogenetic analyses were conducted as described in the *SI Materials and Methods*.

Ave1 allelic variation

To determine the allelic variation, the coding sequence of *Ave1* from 85 race 1 *V. dahliae* strains and two *V. albo-atrum* strains (Table S2) and of *FoAve1* from 72 *F. oxysporum* f. sp. *lycopersici* strains (Table S3) was amplified and sequenced using primers *VdAve1F* and *VdAve1R* (Table S5) and primers *FoAve1-F* and *FoAve1-R* (Table S5) respectively.

Supporting Information

Supplementary Materials and Methods, Supplementary Figures S1, S2, S3, S4, S5, S6 and Supplementary Tables S1, S2, S3, S4, S5

Acknowledgments

We thank Patrick Smit for providing the pSol2092 plasmid.

References

1. Morris CE, *et al*., **Expanding the paradigms of plant pathogen life history and evolution of parasitic fitness beyond agricultural boundaries.** *PLoS Pathog*, 2009, **5**:e1000693.
2. Boller T, Felix G, **A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors.** *Annu Rev Plant Biol*, 2009, **60**:379-402.
3. Thomma BPHJ, Nürnberger T, Joosten MHAJ, **Of PAMPs and effectors: the blurred PTI-ETI dichotomy.** *Plant Cell*, 2011, **23**:4-15.
4. Dodds PN, Rathjen JP, **Plant immunity: towards an integrated view of plant–pathogen interactions.** *Nat Rev Genet*, 2010, **11**:539-548
5. de Jonge R, Bolton MD, Thomma BPHJ, **How filamentous pathogens co-opt plants: the ins and outs of fungal effectors.** *Curr Opin Plant Biol*, 2011, **14**:400-406.
6. Jones JDG, Dangle JL, **The plant immune system.** *Nature*, 2006, 444:323-329.
7. Richards TA, *et al*., **Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms.** *Curr Biol*, 2006, **16**:1857-1864.
8. Friesen T, *et al.,* **Emergence of a new disease as a result of interspecific virulence gene transfer.** *Nat Genet*, 2006, **38**:953-956.
9. Raffaele S, *et al*., **Genome evolution following host jumps in the Irish potato famine pathogen lineage.** *Science*, 2010 **330**:1540-1543.
10. Tyler BM, *et al*., ***Phytophthora*** **genome sequences uncover evolutionary origins and mechanisms of pathogenesis.** *Science*, 2006, **313**:1261-1266.
11. Haas B, *et al*., **Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*.** *Nature*, 2009, **461**:393-398.
12. Kämper J, *et al*., **Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*.** *Nature*, 2006, **444**:97-

101.

13. Ridout CJ, *et al*., **Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasite fitness and defeat of plant resistance.** *Plant Cell*, 2006, **18**:2402-2414
14. Sacristán S, *et al.,* **Coevolution between a family of parasite virulence effectors and a class of LINE-1 retrotransposons.** *PLoS One*, 2009, **4**:e7463.
15. Rouxel T, *et al*. **Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations.** *Nat Commun*, 2011, **2**:202.
16. Richards TA, *et al*., **Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi.** *Plant Cell*, 2009, **21**:1897-1911.
17. Richards TA, *et al*., **Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes.** *Proc Natl Acad Sci USA*, 2011, **108**:15258-15263.
18. Richards TA, **Genome evolution: horizontal movements in the fungi.** *Curr Biol*, 2011, **21**:R112-R114.
19. Fradin EF, Thomma BPHJ, **Physiology and molecular aspects of *Verticillium* wilt diseases caused by *V. dahliae* and *V. albo-atrum*.** *Mol Plant Pathol*, 2006, **7**:71-86.
20. Klosterman SJ, Atallah ZK, Vallad GE, Subbaroa KV, **Diversity, pathogenicity, and management of *Verticillium* species.** *Annu Rev Phytopathol*, 2009, **47**:39-62.
21. Fradin EF, *et al*., **Genetic dissection of *Verticillium* wilt resistance mediated by tomato Ve1.** *Plant Physiol*, 2009, **150**:320-332.
22. Schaible L, Cannon OS, Waddoups V, **Inheritance of resistance to *Verticillium* wilt in a tomato cross.** *Phytopathol*, 1951, **41**:986-990.
23. Kawchuk LM, *et al*., **Tomato *Ve* disease resistance genes encode cell surface-like receptors.** *Proc Natl Acad Sci USA*, 2001, **98**:6511-6515.
24. Fradin EF, *et al*., **Interfamily transfer of tomato Ve1 mediates Verticillium resistance in *Arabidopsis*.** *Plant Physiol*, 2011, **156**:2255-2265.
25. Wang G, *et al*., **The diverse roles of extracellular leucine-rich repeat-containing receptor-like proteins in plants.** *Crit Rev Plant Sci*, 2010, **29**:285-299.
26. Amen J, Shoemaker PB, **Histopathology of resistant and susceptible tomato cultivars inoculated with *Verticillium dahliae* races 1 and 2.** *Phytopathology*, 1985, **75**:1361-1362.
27. Paternotte SJ, van Kesteren HA, **A new aggressive strain of *Verticillium albo-atrum* in Verticillium-resistant cultivars of tomato in the Netherlands.** *Neth J Plant Pathol*, 1993, **99**:169-172.
28. Wevelsiep L, Kogel KH, Knogge W, **Purification and characterization of peptides from *Rhynchosporium secalis* inducing necrosis in barley.** *Physiol Mol Plant Pathol*, 1991, **39**:417-482.
29. Takken FL, *et al*., **A functional cloning strategy, based on a binary PVX-expression vector, to isolate HR-inducing cDNAs of plant pathogens.** *Plant J*, 2000, **24**:275-283.
30. Joosten MHAJ, Cozijnsen TJ, de Wit PJGM, **Host resistance to a fungal tomato pathogen lost by a single base-pair change in an avirulence gene.** *Nature*, 1994, **367**:384-386.
31. Klosterman SJ, *et al*., **Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens.** *PLoS Pathog*, 2011, **7**:e1002137.
32. Parra G, Bradnam K, Korf I, **CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes.** *Bioinformatics*, 2007, **23**:1061–1067.
33. Bolton MD, *et al*., **The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species.** *Mol Microbiol*, 2008, **69**:119-136.
34. Nembaware V, Seoighe C, Sayed M, Gehring C, **A plant natriuretic peptide-like gene in the bacterial pathogen**

Xanthomonas axonopodis **may induce hyper-hydration in the plant host: a hypothesis of molecular mimicry.** *BMC Evol Biol*, 2004, **24**:4-10.

35. Ma LJ, *et al*., **Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*.** *Nature*, 2010, **464**:367-373.
36. Lievens B, Houterman PM, Rep M, **Effector gene screening allows unambiguous identification of *Fusarium oxysporum* f. sp. *lycopersici* races and discrimination from other *formae speciales*.** *FEMS Microbiol Lett*, 2009, **300**:201-215.
37. Luderer R, Takken FL, de Wit PJGM, Joosten MHAJ, ***Cladosporium fulvum*** **overcomes Cf-2-mediated resistance by producing truncated AVR2 elicitor proteins.** *Mol Microbiol*, 2002, **45**:875-884.
38. Ludidi NN, Heazlewood JL, Seoighe C, Irving HR, Gehring CA, **Expansin-like molecules: novel functions derived from common domains.** *J Mol Evol*, 2002, **54**:587-594.
39. Gehring CA, Irving HR, **Natriuretic peptides--a class of heterologous molecules in plants.** *Int J Biochem Cell Biol*, 2003, **35**:1318-1322.
40. Toop T, Donald JA, **Comparative aspects of natriuretic peptide physiology in non-mammalian vertebrates: a review.** *J Comp Physiol B*, 2004, **174**:189-204.
41. Ruzvidzo O, Donaldson L, Valentine A, Gehring CA, **The *Arabidopsis thaliana* natriuretic peptide AtPNP-A is a systemic regulator of leaf dark respiration and signals via the phloem.** *J Plant Physiol*, 2011, **168**:1710-1714.
42. Göttig N, *et al*., ***Xanthomonas axonopodis*** **pv. *citri* uses a plant natriuretic peptide-like protein to modify host homeostasis.** *Proc Natl Acad Sci USA*, 2008, **105**:18631-18636.
43. Dean RA, *et al*., **The genome sequence of the rice blast fungus *Magnaporthe grisea*.** *Nature*, 2005, **434**:980-986.
44. Cuomo CA, *et al*., **The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization.** *Science*, 2007, **317**:1400-1402.
45. Coleman JJ, *et al*., **The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion.** *PLoS Genet*, 2009, **5**:e1000618.
46. Goodwin SB, *et al*., **Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis.** *PLoS Genet*, 2011, **7**:e1002070.
47. Kurtz S, *et al*., **Versatile and open software for comparing large genomes.** *Genome Biol*, 2004, 5:R12.
48. Trapnell C, Pachter L, Salzberg SL, **TopHat: discovering splice junctions with RNA-Seq.** *Bioinformatics*, 2009, **25**:1105-1111.
49. Earley KW, *et al*., **Gateway-compatible vectors for plant functional genomics and proteomics.** *Plant J*, 2006, **45**:616-629.
50. van der Hoorn RAL, Laurent F, Roth R, de Wit PJGM, **Agroinfiltration is a versatile tool that facilitates comparative analyses of Avr9/Cf-9-induced and Avr4/Cf-4-induced necrosis.** *Mol Plant Microbe Interact*, 2000, **13**:439-446.
51. Frandsen RJ, Andersson JA, Kristensen MB, Giese H, **Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi.** *BMC Mol Biol*, 2008, **9**:70.
52. Houterman PM, Cornelissen BJ, Rep M, **Suppression of plant resistance gene-based immunity by a fungal effector.** *PLoS Pathog*, 2008, **4**:e1000061.

Supporting Information

Supplementary Materials and Methods

Verticillium comparative genomics

For each strain, ~ 1 Gb of paired-end sequence reads were checked for quality and filtered accordingly using the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html; version 0.0.13). Filtered reads were assembled with SOAPdenovo63mer (http://soap.genomics.org.cn/soapdenovo.html; version 1.05) using an optimized K-mer setting of 35. Draft genome assemblies of all strains and the VdLs.17 reference genome assembly [1] were aligned with the program Nucmer, applying default options, available in MUMMER3 [2]. Using a set of custom Perl scripts and MS Excel, race 1-specific sequences were identified.

Deep transcriptome sequencing and mapping

For each time point ~ 2 Gb of reads were mapped on the draft genome of race 1 *V. dahliae* isolate JR2 using Tophat version; 1.4.0 [3]. Similarly, ~ 1 Gb of reads, obtained from RNA isolated from *in vitro* cultured JR2 in Czapek Dox medium, were mapped on the draft genome of JR2. Transcript assembly and transcript abundance estimations were performed with Cufflinks version 0.9.3 [4].

Ave1 functional analysis

RACE PCR to validate the *Ave1* gene model was performed with the GeneRacerTM kit (Invitrogen, USA) as described by the manufacturer using primers RACE-AVE1-FWD, RACE-AVE1-REV and RACE-NESTED-F (Table S5). RACE PCR to detect ORF1 transcripts was unsuccessful (Figure 1). For heterologous expression, a potato virus X-mediated expression vector was used [5]. To this end, the *Ave1* coding sequence was amplified from genomic DNA of *V. dahliae* strain JR2 with primers PVX-Ave1-F and PVX-Ave1-R that introduced a 5′ *Cla*I site and a 3′ *Not*I restriction site (Table S5) and cloned into the binary pSfinx vector [5]. *Agrobacterium tumefaciens* transformation and inoculations on resistant (*Ve1*/*Ve1*) and susceptible (*ve1*/*ve1*) tomato plants were performed as described previously [6].

To generate constructs for constitutive expression driven by the CaMV 35S promoter, the coding sequences of *Ave1*homologs were cloned into Gateway destination vector pSol2092, a Gateway-compatible pBIN derivative (kind gift of Dr. Patrick Smit). *VdAve1* and *FoAve1* were amplified from cDNA using the primer pairs AT-Ave1-F / AT-Ave1-R, and AT-FoAve1-F / AT-FoAve1-R, respectively (Table S5). *ChAve1*, *CbAve1* and *SlAve1* were obtained by gene synthesis (Eurofins/MWG, Germany). The tomato *Ve1* and *Ve2* gene were cloned into Gateway destination vector pEarleyGate100 [7,8]. Also, pMOG800_Ve1 and pMOG800_Ve2 were used [7]. The constructs were transformed into *A. tumefaciens* strain GV3101 and infiltrated into *Nicotiana tabacum* cv. Petite Havana SR1 as described previously [9]. Briefly, an overnight culture of *A. tumefaciens* was harvested at OD600 of 0.8 to 1 by centrifugation and resuspended to a final OD of 2. *A. tumefaciens* cultures containing constructs to express *Ave1* and *Ve1* were mixed in a 1:1 ratio and infiltrated into five- to six-week-old tobacco leaves. At five days post infiltration, leaves were examined for necrosis.

Ave1 knockouts were generated by amplifying the sequences flanking the *Ave1* coding sequence using the primers KO-Ave1-F1 and KO-Ave1-R1 (Table S5), and the primers KO-Ave1-F2 and KO-Ave1-R2 (Table S5), and cloned into the vector pRF-HU2 as described [10]. Genomic complementation of race 2 strains and of *Ave1* knockouts in race 1 was performed using a genomic construct consisting of the *Ave1* ORF plus 1.5 Kb up- and downstream sequence (*pAve1::Ave1*) cloned into the vector pRW1P [11] using primers CO-Ave1-F and CO-Ave1-R (Table S5). Based on the 5′ RACE PCR, a primer pair (5UTR-AVE1-FWD / PVX-Ave1-R) was designed targeting the complete mRNA sequence to verify correct splicing and expression of the *Ave1* gene in transgenic strains (Figure S1*C*).

V. dahliae transformation was performed as described previously [7]. *V. dahliae* and *Fusarium oxysporum* inoculations on MoneyMaker (*ve1*/*ve1*), Motelle (*Ve1*/*Ve1*) and transgenic MoneyMaker (*35s::Ve1*) plants to assess the impact on virulence and avirulence in absence and presence of *Ve1*, respectively, were performed as described previously [7,12]. Plants were regularly inspected during a three-week interval. Two weeks after *V. dahliae* inoculation on Motelle plants, stem sections immediately above the cotyledons were taken, surface sterilized, sliced, transferred onto potato dextrose agar supplemented with chloramphenicol (100 µg/ml) and incubated at 22°C. Outgrowth of *V. dahliae* from stem slices indicates loss of recognition by the host.

V. dahliae inoculations on wild type *Arabidopsis thaliana* (Col-0) and *Ve1*-transgenic Arabidopsis (*p35S::Ve1*; were performed as described previously [13]. Plants were regularly inspected during a four-week interval. Three weeks post inoculation, flowering parts were removed and photographs were taken. For quantification per *V. dahliae* transformant all aboveground tissues of five plants were flash-frozen in liquid nitrogen. The samples were ground to a powder, of which an aliquot of approximately 100 mg was used for DNA isolation [14]. Real-time PCR was conducted with primers AtRub-F3 and AtRub-R3 for Arabidopsis *RuBisCo* and primers VdELF-1a-F and VdELF-1a-R for *V. dahliae elongation factor 1-alpha* (Table S5). Real-time PCR was conducted using an ABI7300 PCR machine (Applied Biosystems, Foster City, USA) in combination with the qPCR SensiMix kit (BioLine, London, UK). Real-time PCR conditions were as follows: an initial 95°C denaturation step for 10 min followed by denaturation for 15 s at 95°C and annealing for 30s at 60°C and extension at 72°C for 40 cycles.

Protein analyses

BLASTp was used to detect homologs of Ave1 in the non-redundant (nr) database hosted at NCBI. In addition, tBLASTn was used at the MycoCosm (http://genome.jgi-psf.org/programs/fungi/index.jsf) database and the Fungal Genome Initiative of the Broad institute (http://www.broadinstitute.org/annotation/genome/FGI_Blast.1/Blast.html) to detect homology in 105 fungal nuclear genomes (Table S4). Signal peptides were determined with SignalP3.0 (http://www.cbs.dtu.dk/services/SignalP/) using the Neural Network method [15]. Amino acid sequence alignments were generated using MAFFT [16], incorporating the FFT-NS-I strategy which optimizes between accuracy and speed.

Phylogenetic analyses

Phylogenetic analyses were conducted online using the Phylogeny.fr web-service (http://www.phylogeny.fr/version2_cgi/index.cgi; [17]), which incorporates Gblocks [18] and PhyML [19], and locally using the software package MEGA5 [20]. For alignment curation by Gblocks we used the least stringent settings, allowing for smaller blocks, gap positions within the blocks and less stringent flanking positions. Maximum likelihood phylogenies were calculated using PhyML [19], applying the SH-like approximate likelihood-ratio test (aLRT; [21]) for branch support. The appropriate substitution model (JTT) was determined in MEGA5 by applying the model estimator [20]. This model was then used in PhyML to estimate substitution probabilities along the branches.

Data deposition

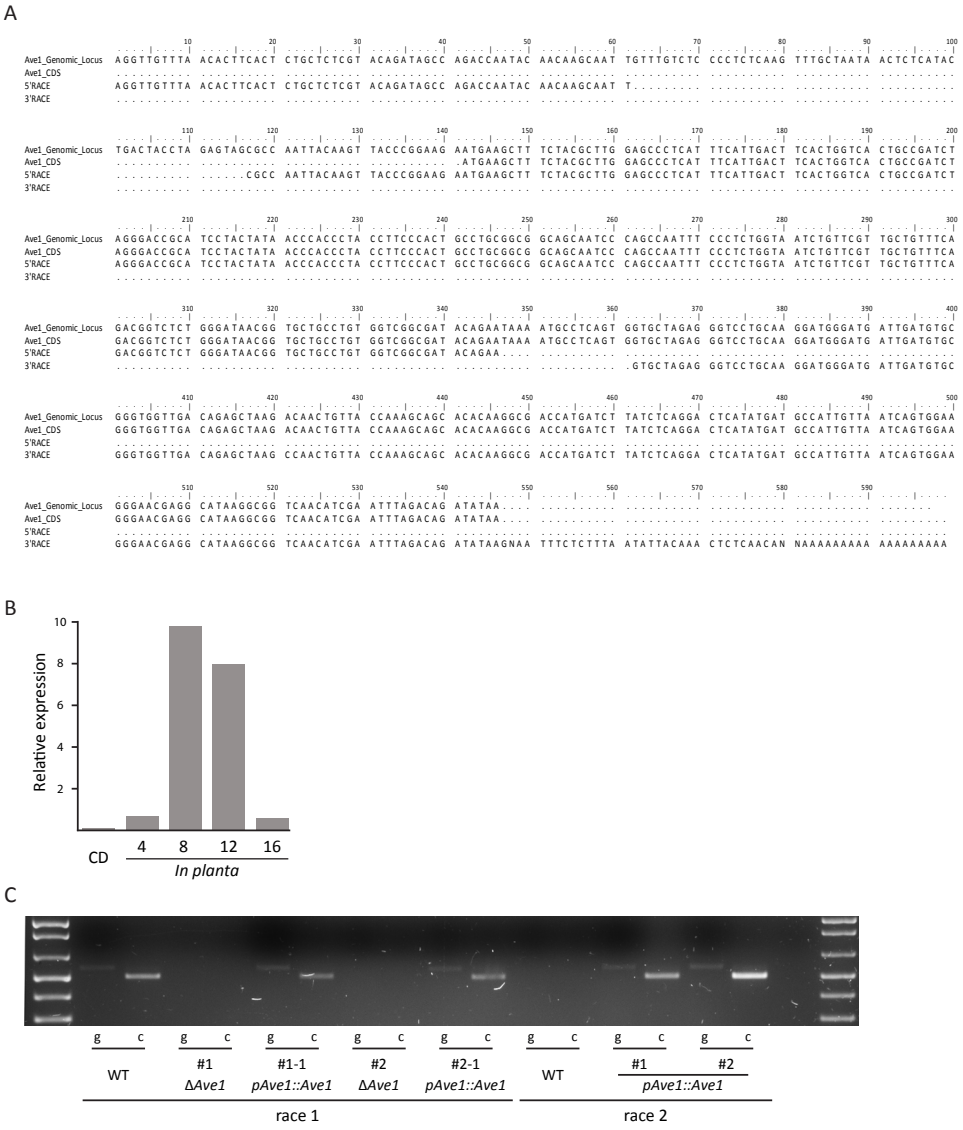
The sequences reported in this paper have been deposited in GenBank under accession numbers JN616379 (*VdAve1*), JQ283440 (*FoAve1*), JQ283439 (*ChAve1*), JQ583777 (*CbAve1*), JQ625338- JQ625341 (race 1 scaffold sequences from *V. dahliae* strains Ls16, CBS381.66, JR2 and St14.01 respectively).

Supplementary References

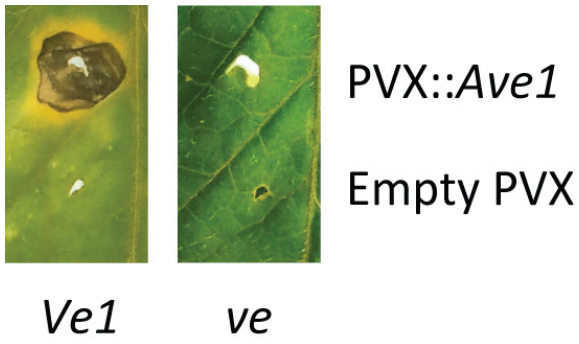
1. Klosterman S, *et al*., **Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens.** *PLoS Pathog*, 2011, **7**:e1002137.
2. Kurtz S, *et al*., **Versatile and open software for comparing large genomes.** *Genome Biol*, 2004, **5**:R12.
3. Trapnell C, Pachter L, Salzberg SL, **TopHat: discovering splice junctions with RNA-Seq.** *Bioinformatics*, 2009, **25**:1105-1111.
4. Trapnell, C *et al*., **Transcripts assembly and quantification by RNA-Seq reveals unannotated transcripts and isoforms switching during cell differentiation.** *Nat Biotech*, 2010, **28**:511-515.
5. Takken FL, *et al*., **A functional cloning strategy, based on a binary PVX-expression vector, to isolate HR-inducing cDNAs of plant pathogens.** *Plant J*, 2000, **24**:275-283.
6. van Esse HP, et al., **Affinity-tags are removed from *Cladosporium fulvum* effector proteins expressed in the tomato leaf apoplast.** *J Exp Bot*, 2006, **57**:599-608.
7. Fradin EF, *et al.,* **Genetic dissection of *Verticillium* wilt resistance mediated by tomato Ve1.** *Plant Physiol*, 2009, **150**:320-332.
8. Earley KW, et al., **Gateway-compatible vectors for plant functional genomics and proteomics.** *Plant J*, 2006, **45**:616-629.
9. van der Hoorn RAL, Laurent F, Roth R, de Wit PJGM, **Agroinfiltration is a versatile tool that facilitates comparative analyses**

of Avr9/Cf-9-induced and Avr4/Cf-4-induced necrosis. *Mol Plant-Microbe Interact*, 2000, **13**:439-446.

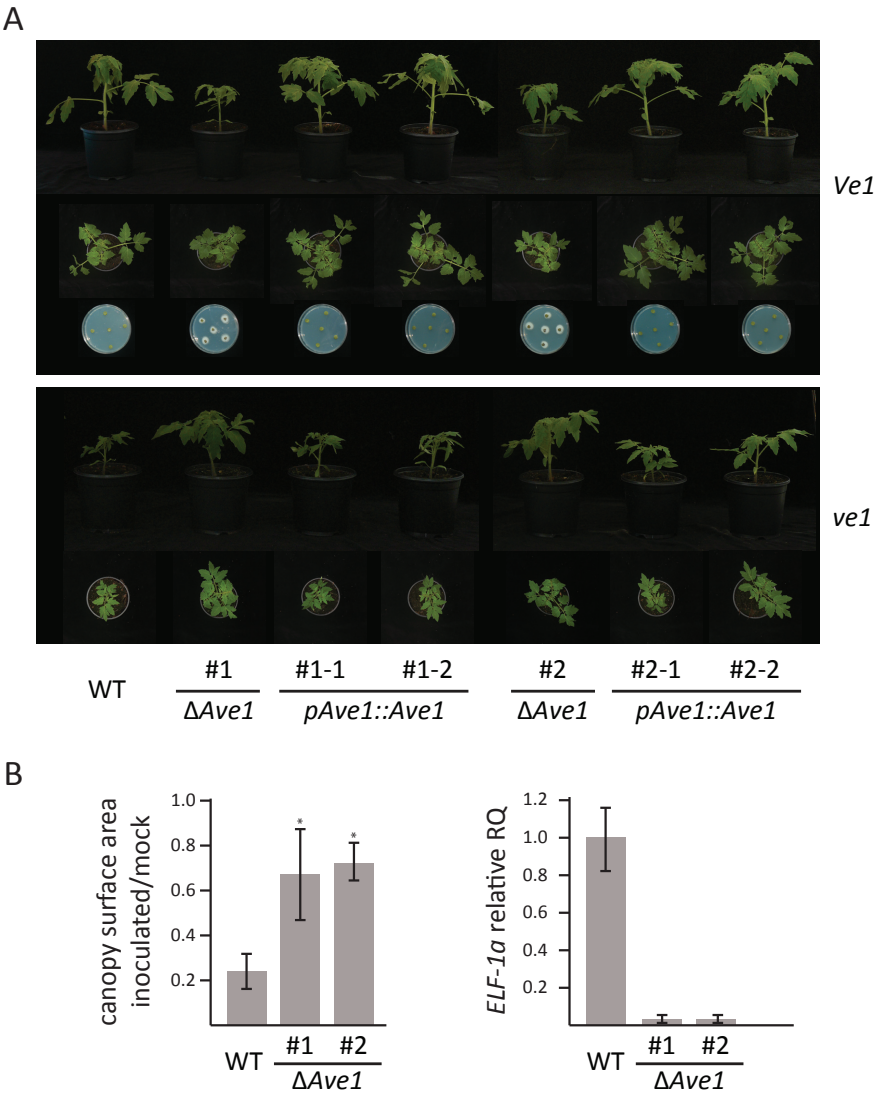
10. Frandsen RJ, Andersson JA, Kristensen MB, Giese H, **Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi.** *BMC Mol Biol*, 2008, **9**:70.
11. Houterman PM, Cornelissen BJ, Rep M, **Suppression of plant resistance gene-based immunity by a fungal effector.** *PLoS Pathog*, 2008, **4**:e1000061.
12. van Esse HP, Bolton MD, Stergiopoulos I, de Wit PJGM, Thomma BPHJ, **The chitin-binding *Cladosporium fulvum* effector protein Avr4 is a virulence factor.** *Mol Plant-Microbe Interact*, 2007, **20**:1092-1101.
13. Fradin EF, *et al.,* **Interfamily transfer of tomato Ve1 mediates Verticillium resistance in Arabidopsis.** *Plant Physiol*, 2011, **156**:2255-2265.
14. Fulton TM, Chunwongse J, Tanksley SD, **Microprep protocol for extraction of DNA from tomato and other herbaceous plants.** *Plant Mol Biol Rep*, 1995, **13**:207-209.
15. Bendtsen JD, Nielsen H, von Heijne G, Brunak S, **Improved prediction of signal peptides: SignalP3.0.** *J Mol Biol*, 2004, **340**:783-795.
16. Katoh K, Toh H, **Parallelization of the MAFFT multiple sequence alignment program.** *Bioinformatics*, 2010, **26**:1899-1900.
17. Dereeper A, *et al*., **Phylogeny.fr: robust phylogenetic analysis for the non-specialist.** *Nucleic Acids Res*, 2008, **36**:W465-469.
18. Castresana J, **Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis.** *Mol Biol Evol*, 2000, **17**:540-552.
19. Guindon S, *et al*., **New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0.** *Syst Biol*, 2010, **59**:307-321.
20. Tamura K, *et al*., **MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Mol Biol Evol*, 2011, **28**:2731-2739.
21. Guindon S, Gascuel O, **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Syst Biol*, 2003 **52**:696-704.



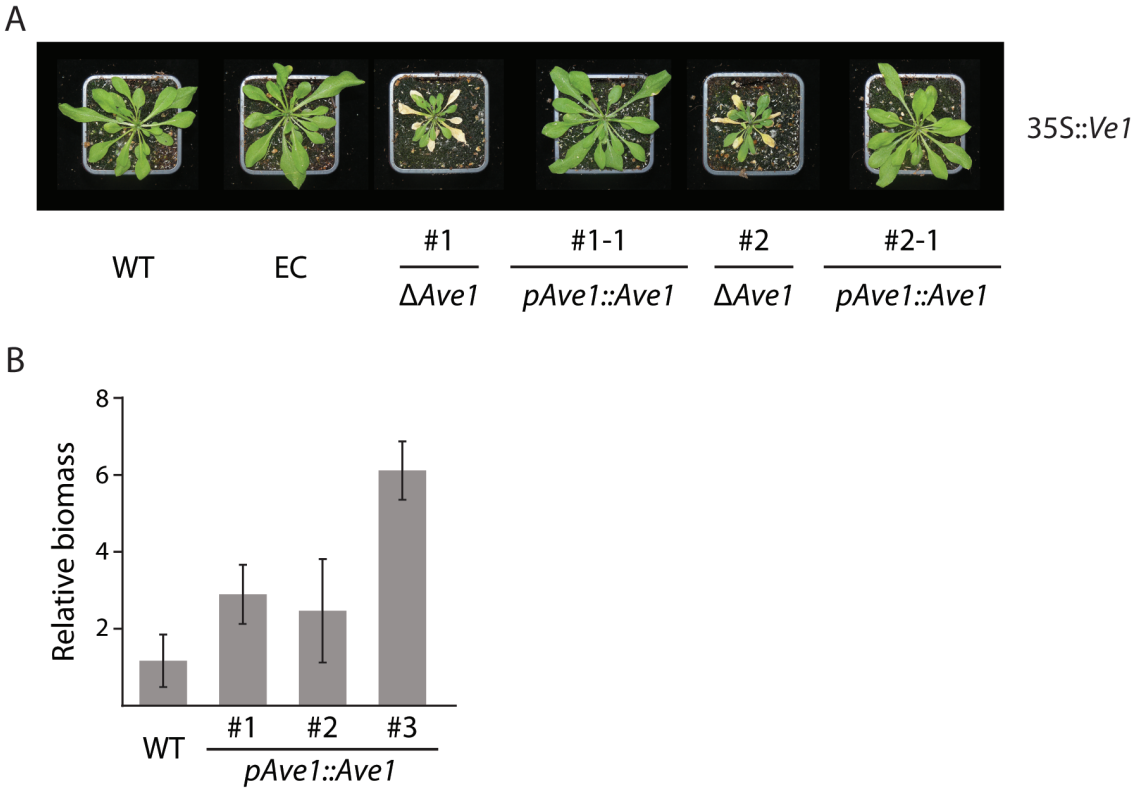
Supplementary Figure 1. *Ave1* transcription in *Verticillium dahliae*. A) Sequence alignment of the genomic Ave1 locus, the *Ave1* coding sequence (CDS) and the mRNA sequences obtained in 5'and 3'RACE experiments. The sequence gap observed in the alignment of the 5' RACE sequence indicates the intron in the 5'UTR. B) *Ave1* expression in Czapek Dox culture medium (CD) and during the interaction of *V. dahliae* with *Nicotiana benthamiana* at 4, 8, 12 and 16 days post inoculation. RNA sequencing reads were used to determine the relative expression of *Ave1*. C) Amplification of *Ave1* from genomic DNA (g) and cDNA (c) in wild-type race 1 *V. dahliae* (WT), two independent *Ave1* deletion strains (Δ*Ave1*), and the same *Ave1* deletion strains upon complementation with a genomic *Ave1* fragment (*pAve1::Ave1*) that were inoculated on susceptible tomato plants. Furthermore, amplification of *Ave1* is shown from genomic DNA (g) and cDNA (c) in wild-type race 2 *V. dahliae* (WT) and two transformants that are complemented with a genomic *Ave1* fragment (*pAve1::Ave1*) that were inoculated on susceptible tomato plants. For amplification, a primer pair spanning the intron in the 5'UTR was used.



Supplementary Figure 2. Heterologous expression of *Ave1* results in a hypersensitive response in *Ve1* tomato. Potato Virus X based expression of *Ave1* results in an expanding hypersensitive response in *Ve1* (left), but not in *ve1* tomato (right).

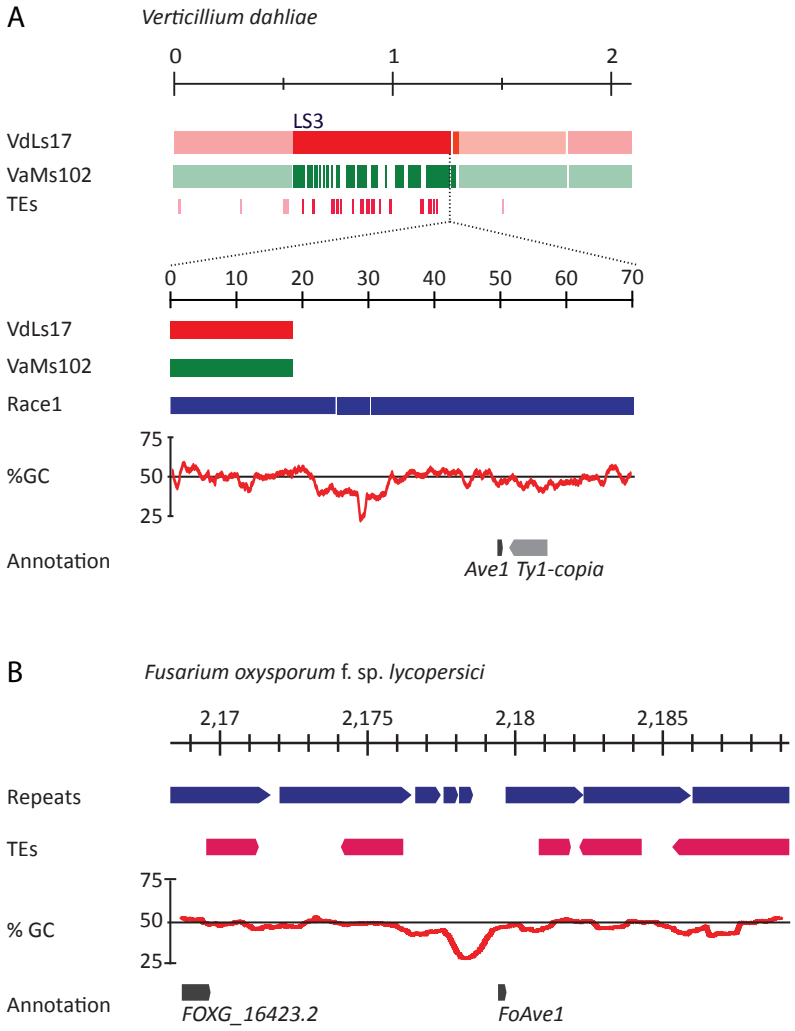


Supplementary Figure 3. *Ave1* expression complements race 1 *Ave1* deletion strains and contributes to virulence. A) Top: *Ave1* expression (*pAve1::Ave1*) in *Ave1* deletion strains (Δ*Ave1*) results in recognition by *Ve1* tomato as evidenced by healthy plants (photographs 14 dpi) and the absence of fungal outgrowth after plating of stem sections. Bottom: *Ave1* expression (*pAve1::Ave1*) in *Ave1* deletion strains (Δ*Ave1*) restores virulence on *ve1* tomato. Pictures for two independent *Ave1* expressing strains per deletion strain are shown. B) *Ave1* deletion strains are impaired in virulence, evidenced by the reduction in stunting when compared with mock-inoculated plants at 14 dpi (left) and decrease in fungal biomass at 14 dpi measured with real-time PCR (right).



Supplementary Figure 4. *Verticillium dahliae* Ave1 triggers resistance in *Ve1*-transgenic Arabidopsis and enhances virulence on susceptible plants. A) *Ave1* deletion strains (Δ*Ave1*) escape recognition by Ve1 when compared to wild-type (WT) and ectopic transformant (EC). *Ave1* expression (*pAve1::Ave1*) in *Ave1* deletion strains (Δ*Ave1*) restores resistance. Pictures for two independent *Ave1* deletion strains and one complemented *Ave1* expressing strain per deletion strain at 21 dpi are shown. B) *Ave1* expression (*pAve1::Ave1*) enhances the virulence of *V. dahliae* race 2 on susceptible Arabidopsis, evidenced by increased fungal biomass accumulation measured with real-time PCR when compared to wild-type (WT). The results of a representative experiment are shown and the error bars represent standard deviations.

Supplementary Figure 5. Amino acid sequence alignment of VdAve1, FoAve1, ChAve1, CbAve1, XacPNP and 50 homologous plant proteins. The positions of four conserved cysteine residues are indicated with asterisks on top of the alignment.



Supplementary Figure 6. *Verticillium dahliae* and *Fusarium oxysporum Ave1* are located in flexible genomic regions. A) Alignment of a 2 Mb section of chromosome 4 of the reference genomes of *V. dahliae* (VdLs.17) and *V. albo-atrum* (VaMs102) illustrating the genomic context of *Ave1* in race 1 strains. The alignment depicts lineage-specific region 3 (LS3) of VdLs.17 enriched in transposable elements (TEs; 1) that partially overlaps with the race 1-specific region containing *Ave1*. The GC content and a *Ty1-copia* retrotransposon flanking *Ave1* in race 1 strains are shown. B) *FoAve1* is located in a flexible region of the *Fusarium oxysporum* f. sp. *lycopersici* genome. Subsection (2,170,000 – 2,185,000) of chromosome 14 of the *F. oxysporum* f. sp. *lycopersici* genome illustrating the genomic context of *FoAve1*. The GC content and the presence of various transposable elements and overlapping repeat sequences are shown.

Supplementary Table 1. Mapping statistics of the RNA sequencing of *Verticillium dahliae*-infected *Nicotiana benthamiana* and *V. dahliae* cultured in Czapek Dox Broth.

RNA-Seq of *Verticillium dahliae* strain JR2				
Sample	**Number of reads**	**Number of mapped reads**	**Properly paired (%)**	**Singletons (%)**
4 dpi*	26,200,004	13,642	52,03	42,47
8 dpi*	26,177,780	10,333	50,54	40,37
12 dpi*	26,177,780	30,422	68,10	28,84
16 dpi*	26,911,112	247,201	68,91	26,30
Czapek Dox	12,955,422	10,327,776	76,85	14,74

* Samples of *Nicotiana benthamiana* collected at 4, 8, 12 and 16 days post inoculation with *Verticillium dahliae*.

Supplementary Table 2. *Verticillium* strains analyzed for *Ave1* allelic variation.

Species	Isolate	Original host	Location	Year	Pathogenicity on tomato	Race	*Ave1*
Verticillium dahliae	TV103	tomato	Japan (Tokyo)	1971	+	1	+
	U22	*Aralia cordata*	Japan (Gunma)	1986	+	1	+
	Shio	tomato	Japan (Tokyo)		+	1	+
	TO2	tomato	Japan (Gunma)	1984	+	1	+
	TK23	tomato	Japan (Kanagawa)	1992	+	1	+
	ATCC_201177	tomato	Canada (Ontario)		+	1	+
	Vdp4	sweet pepper	Japan (Nagano)	1991	+	1	+
	Kgm	tomato	Japan		+	1	+
	Gto2	tomato	Japan (Gunma)	2001	+	1	+
	TR-1	tomato	Japan		+	1	+
	M-1	melon	Japan (Gunma)		+	1	+
	09095-1-2B	lettuce	Japan (Ibaraki)	2009	+	1	+
	09095-4B	lettuce	Japan (Ibaraki)	2009	+	1	+
	09096-2B	lettuce	Japan (Ibaraki)	2009	+	1	+
	09096-3B	lettuce	Japan (Ibaraki)	2009	+	1	+
	JR2	tomato	Canada (Ontario)		+	1	+
	ST 14.01	pistachio	USA (California)		+	1	+
	CBS 381.66	tomato	Canada (Quebec)		+	1	+
	VdLs16	lettuce	USA (California)	1996	+	1	+
	5410	tomato	Australia	2002	+	1	+
	VdLs1	lettuce	USA (CA)		+	1	+
	Gh1003	cotton	USA (CA)		+	1	+
	Cf36	pepper	USA (CA)		+	1	+
	Ar136	horseradish	USA (IL)		+	1	+
	Cf38	pepper	USA (CA)		+	1	+
	Cf45	pepper	USA (CA)		+	1	+
	VdLs183	lettuce	USA (CA)		+	1	+
	Cs225	artichoke	USA (CA)		+	1	+
	Ca271	pepper	USA (CA)		+	1	+
	Fca414	strawberry	USA (CA)		+	1	+
	Cv896	watermelon	USA (CA)		+	1	+
	VdLs897	lettuce	USA (CA)		+	1	+
	So934	spinach	Denmark		+	1	+
	Le1087	tomato	USA (CA)		+	1	+
	Dvd-S100	tomato	Canada		+	1	+
	Dvd-S21	tomato	Canada		+	1	+
	Dvd-S90	tomato	Canada		+	1	+
	Dvd-T5	tomato	Canada		+	1	+

Fca2219	strawberry	USA (CA)		?		+
Fca2220	strawberry	USA (CA)		?		+
Fca2221	strawberry	USA (CA)		?		+
Fca2222	strawberry	USA (CA)		?		+
Fca2223	strawberry	USA (CA)		?		+
Fca2224	strawberry	USA (CA)		?		+
Fca2225	strawberry	USA (CA)		?		+
Fca2226	strawberry	USA (CA)		?		+
Fca2227	strawberry	USA (CA)		?		+
Fca2228	strawberry	USA (CA)		?		+
Fca2229	strawberry	USA (CA)		?		+
Fca2230	strawberry	USA (CA)		?		+
Fca2231	strawberry	USA (CA)		?		+
Fca2232	strawberry	USA (CA)		?		+
Fca2233	strawberry	USA (CA)		?		+
Fca2234	strawberry	USA (CA)		?		+
Fca2235	strawberry	USA (CA)		?		+
Fca2236	strawberry	USA (CA)		?		+
Fca2237	strawberry	USA (CA)		?		+
Fca2238	strawberry	USA (CA)		?		+
Fca2239	strawberry	USA (CA)		?		+
Fca2240	strawberry	USA (CA)		?		+
Fca2242	strawberry	USA (CA)		?		+
Gh1002	cotton	USA (CA)		?		+
Gh2332	cotton	China		?		+
VdLs1954	lettuce	USA (CA)		?		+
VdLs1961	lettuce	USA (CA)		?		+
VdLs1962	lettuce	USA (CA)		?		+
VdLs1968	lettuce	USA (CA)		?		+
Oe2154	olive	Italy		?		+
Oe2155	olive	Italy		?		+
Oe2157	olive	Italy		?		+
Oe2158	olive	Italy		?		+
Oe2159	olive	Italy		?		+
Oe2160	olive	Italy		?		+
Oe2162	olive	Italy		?		+
Oe2163	olive	Italy		?		+
Oe2165	olive	Italy		?		+
Oe2167	olive	Italy		?		+
Ca2169	pepper	Italy		?		+
Sm2171	eggplant	Italy		?		+
Oe2174	olive	Italy		?		+
Oe2176	olive	Italy		?		+
So1221	spinach	USA (Washington)		?		+
So2538	spinach	Denmark		?		+
VdLs.17	lettuce	USA (California)		+	2	-
DVD-3	potato	Canada (Essex Co.)	1993	+	2	-
DVD-31	tomato	Canada (Essex Co.)	1993	+	2	-
DVD-161	potato	Canada (ON) (Simcoe Co.)	1993	+	2	-
DVD-s26	soil	Canada (Essex Co.)	1994	+	2	-
DVD-s29	soil	Canada (Essex Co.)	1994	+	2	-
DVD-s94	soil	Canada (Kent Co.)	1996	+	2	-
TO20	tomato	Japan (Gunma)	1991	+	2	-
TO21	tomato	Japan (Gunma)	1991	+	2	-
TO22	tomato	Japan (Gunma)	1991	+	2	-
TO23	tomato	Japan (Gunma)	1991	+	2	-
TO24	tomato	Japan (Gunma)	1991	+	2	-
TO26	tomato	Japan (Gunma)	1991	+	2	-
TK15	tomato	Japan (Kanagawa)	1992	+	2	-
Ud1-2-1	*Aralia cordata*	Japan (Gunma)	2006	+	2	-
Ud1-4-1	*Aralia cordata*	Japan (Gunma)	2006	+	2	-

	Gok1	okra	Japan (Gunma)	2001	+	2	-
	Gca1	cabbage	Japan (Gunma)	2001	+	2	-
	CA39	cabbage	Japan (Gunma)	199?	+	2	-
	St.100	soil	Belgium		-		-
	U48	*Aralia cordata*	Japan (Gunma)	1987	-		-
	Cns	eggplant	Japan (Nagano)	1980	-		-
	22210	eggplant	Japan (Tokushima)	1972	-		-
	Vdp3	sweet pepper	Japan (Nagano)	1991	-		-
	U2	*Aralia cordata*	Japan (Gunma)		-		-
	U20	*Aralia cordata*	Japan (Gunma)	1986	-		-
	P2-1	sweet pepper	Japan (Hokkaido)	1992	-		-
	P2-2	sweet pepper	Japan (Hokkaido)	1992	-		-
	P8-1	sweet pepper	Japan (Hokkaido)	1998	-		-
	P9-1	sweet pepper	Japan (Hokkaido)	1999	-		-
	P9-2	sweet pepper	Japan (Iwate)	1999	-		-
	Gud1	*Aralia cordata*	Japan (Gunma)	2001	-		-
	Gfk1	*Petasites japonicus*	Japan (Gunma)	2001	-		-
	CA43	cabbage	Japan (Gunma)	199?	-		-
	84023	eggplant	Japan (Nagano)	1979	-		-
	Chr208	chrysanthemum	Japan (Tokyo)	1970	-		-
	Ibh	Chinese cabbage	Japan (Ibaraki)		-		-
	Y3-1	eggplant	Japan (Yamagata)	1972	-		-
	CA26	cabbage	Japan (Gunma)	199?	-		-
	84034	eggplant	Japan (Mie)	1981	-		-
	Ara406	*Aralia cordata*	Japan (Gunma)	1977	-		-
	09100-1	lettuce	Japan (Ibaraki)	2009	-		-
	NBRC_9435	eggplant	Japan (Aichi)		-		-
	NBRC_9470	eggplant	Japan (Oosaka)		-		-
	Ns-1	eggplant	Japan (Ibaraki)	1997	-		-
	Tns	*Solanum mammosum*	Japan (Chiba)	1995	-		-
	Hns	*Solanum pseudocapsicum*	Japan (Chiba)	1995	-		-
	Po-I	Iceland poppy	Japan (Chiba)	2001	-		-
	NBRC_6119		Brazil		-		-
	WA	potato	USA (Washington)		-		-
	ChiA	cotton	China (Uighur)	1999	-		-
Verticillium longisporum	CA9	cabbage	Japan (Gunma)	199?	-		-
	CA10	cabbage	Japan (Gunma)	199?	-		-
	CA58	cabbage	Japan (Gunma)	199?	-		-
	Dk-1	Japanese radish	Japan (Chiba)	1999	-		-
	84013	Chinese cabbage	Japan		-		-
Verticillium albo-atrum	HP	potato	Japan (Hokkaido)		-		+
	NBRC_31023		UK		-		-
	MAFF_235137	alfalfa	Japan (Hokkaido)	1981	-		-
	MAFF_235138	alfalfa	Japan (Hokkaido)	1980	-		-
	VaaMS102	alfalfa	USA (Pennsylvania)	1986	-		+
Verticillium tricorpus	NBRC_31025		UK		-		-
	CE98Vt1	potato	Japan	1998	-		-
	CE20VtLe3	tomato	Japan (Chiba)	2000	-		-
	MAFF_712235	delphinium	Japan (Miyagi)	1999	-		-
	MIH001	lettuce	Japan (Hyogo)	2002	-		-
	eLTS1	lettuce	Japan (Chiba)	2000	-		-
	Ls.432	lettuce	USA (California)	2001	-		-
Gibellulopsis nigrescens	Vn78	melon	Japan (Tokyo)		-		-
	CE20VnAc1	anemone	Japan (Chiba)	2000	-		-
Musicillium theobromae	MAFF_239114	banana	Japan (Tokyo)	1999	-		-
Plectosporium tabacinum	CorA	coriander	Japan (Chiba)	2011	+		-

Supplementary Table 3. *Fusarium* strains analyzed for *FoAve1* allelic variation.

Species	Forma specialis	Isolate	Host	Pathogenicity on tomato	*FoAve1*
Fusarium oxysporum	asparagi	NRRL 28973	asparagus	-	-
	asparagi	NRRL 28362	asparagus	-	-
	asparagi	NRRL 28379	asparagus	-	-
	conglutinans	81-4	cabbage	-	-
	cubense	NRRL 25603	banana	-	-
	cubense	NRRL 25609	banana	-	-
	cubense	NRRL 26029	banana	-	-
	cucumerinum	ATCC 16416	cucumber	-	-
	cucumerinum	ATCC 201950	cucumber	-	-
	cucumerinum	ATCC 36330	cucumber	-	-
	cucumerinum	Afu-50(B)	cucumber	-	-
	cucumerinum	Afu-52	cucumber	-	-
	cucumerinum	FOCU-CM1C	cucumber	-	-
	cucumerinum	FOCU-707E	cucumber	-	-
	cucumerinum	FOCU-22P	cucumber	-	-
	cucumerinum	FOCU-45K	cucumber	-	-
	cucumerinum	Afu-57(B)	cucumber	-	-
	cucumerinum	FOCU-26E	cucumber	-	-
	cucumerinum	FOCU-33N	cucumber	-	-
	cucumerinum	FOCU-39E	cucumber	-	-
	cucumerinum	FOCU-48F	cucumber	-	-
	cucumerinum	FOCU-16F	cucumber	-	-
	cucumerinum	FOCU-17W	cucumber	-	-
	cucumerinum	Cu:4-1 Koma 4	cucumber	-	-
	cucumerinum	NETH 11179	cucumber	-	-
	cucumerinum	ATCC 36332	cucumber	-	-
	cucumerinum	0019	cucumber	-	-
	cucumerinum	0018	cucumber	-	-
	cucumerinum	Cu: 5-0 Koma 5	cucumber	-	-
	cucumerinum	9906-3	cucumber	-	-
	cucumerinum	9909-2	cucumber	-	-
	cucumerinum	9904-1	cucumber	-	-
	cucumerinum	9901-2	cucumber	-	-
	cucumerinum	9903-1	cucumber	-	-
	cucumerinum	0016	cucumber	-	-
	cucumerinum	0020	cucumber	-	-
	cucumerinum	ATCC 42352	cucumber	-	-
	cucumerinum	ATCC 42357	cucumber	-	-
	cucumerinum	NRRL 26437	cucumber	-	-
	cucumerinum	00/0092/1	cucumber	-	-
	cucumerinum	Foc-1	cucumber	-	-
	cucumerinum	Foc-2	cucumber	-	-
	cucumerinum	10196	cucumber	-	-
	cucumerinum	9909-3	cucumber	-	-
	cucumerinum	1030554	cucumber	-	-
	cucumerinum	305117	cucumber	-	-
	cucumerinum	727508	cucumber	-	-
	cucumerinum	744004	cucumber	-	-
	cucumerinum	744005	cucumber	-	-
	cucumerinum	0017	cucumber	-	-
	cucumerinum	DSM 62313	cucumber	-	-
	cucumerinum	9903-2	cucumber	-	-
	cucumerinum	9906-2	cucumber	-	-
	cucumerinum	Tf 213 - Afu Shiomi 2	cucumber	-	-
	cucumerinum	NETH 10782(B)	cucumber	-	-
	cucumerinum	305116	cucumber	-	-
	dianthi	NRRL 26147	carnation	-	-
	dianthi	NRRL 26960	carnation	-	-

	gladioli	NRRL 28914	gladiolus	-	-
	gladioli	NRRL 26993	gladiolus	-	-
	gladioli	NRRL 26990	gladiolus	-	-
	lilii	NRRL 28395	lily	-	-
	lilii	NRRL 26955	lily	-	-
	luffae	FOL-167	*Luffa cylindrica*	-	-
	luffae	FOL-114	*Luffa cylindrica*	-	-
	lycopersici	FOL-HH6M	tomato	+	+
	lycopersici	FOL-24L	tomato	+	+
	lycopersici	WCS861/E240	tomato	+	+
	lycopersici	IPO1530/B1	tomato	+	+
	lycopersici	FOL-93H	tomato	+	+
	lycopersici	281	tomato	+	+
	lycopersici	4287	tomato	+	+
	lycopersici	WCS862/E241	tomato	+	+
	lycopersici	Fol-70	tomato	+	+
	lycopersici	BFOL-53	tomato	+	+
	lycopersici	IP03	tomato	+	+
	lycopersici	4887	tomato	+	+
	lycopersici	D1	tomato	+	+
	lycopersici	LSU-3	tomato	+	+
	lycopersici	Fol1	tomato	+	+
	lycopersici	Fol-650B	tomato	+	+
	lycopersici	218	tomato	+	+
	lycopersici	FRC-0-1078	tomato	+	+
	lycopersici	18947	tomato	+	+
	lycopersici	Fol-1295T	tomato	+	+
	lycopersici	548	tomato	+	+
	lycopersici	Fol036	tomato	+	+
	lycopersici	MX395	tomato	+	+
	lycopersici	CA92/95	tomato	+	+
	lycopersici	14844	tomato	+	+
	lycopersici	5397	tomato	+	+
	lycopersici	FOL-295A	tomato	+	+
	lycopersici	E179	tomato	+	+
	lycopersici	FOL-R5-6 / E172	tomato	+	+
	lycopersici	D2	tomato	+	+
	lycopersici	OSU-451	tomato	+	+
	lycopersici	E175	tomato	+	+
	lycopersici	BFOL-51	tomato	+	-
	lycopersici	FOL-lyc07038	tomato	+	+
	lycopersici	FRC-O-1113N	tomato	+	+
	lycopersici	FRC-O-1113A	tomato	+	+
	lycopersici	FOL-MM59	tomato	+	+
	lycopersici	LSU-7	tomato	+	+
	lycopersici	FOL-MM25	tomato	+	+
	lycopersici	FOL-MN25	tomato	+	+
	lycopersici	C24/B2	tomato	+	+
	lycopersici	E79	tomato	+	+
	lycopersici	48112	tomato	+	+
	lycopersici	WCS801/E329	tomato	+	+
	lycopersici	E181	tomato	+	+
	lycopersici	626	tomato	+	+
	lycopersici	CBS 412.90	tomato	+	+
	lycopersici	CBS 645.78	tomato	+	+
	lycopersici	DSM 62059	tomato	+	+
	lycopersici	MD-S2	tomato	+	+
	lycopersici	MD-L3	tomato	+	+
	lycopersici	CBS 413.90	tomato	+	-
	lycopersici	CBS 414.90	tomato	+	+
	lycopersici	CBS 646.78	tomato	+	+
	lycopersici	Fol-W841	tomato	+	+

lycopersici	EY-101	tomato	+	+
lycopersici	ATCC 417	tomato	+	+
lycopersici	EY-102	tomato	+	+
lycopersici	ATCC 605	tomato	+	+
lycopersici	00/60309/1	tomato	+	+
lycopersici	MUCL 19445	tomato	+	+
lycopersici	CBS 758.68	tomato	+	+
lycopersici	CBS 165.85	tomato	+	+
lycopersici	DSM 62338	tomato	+	+
lycopersici	NRRL 26034	tomato	+	+
lycopersici	NRRL 26037	tomato	+	+
lycopersici	NRRL 26200	tomato	+	+
lycopersici	NRRL 26202	tomato	+	+
lycopersici	NRRL 26203	tomato	+	+
lycopersici	NRRL 26383	tomato	+	+
lycopersici	Bt.01	tomato	+	+
lycopersici	Fol045	tomato	+	-
melonis	NRRL 26046	melon	-	-
melonis	CBS 423.90	melon	-	-
melonis	CBS 420.90	melon	-	-
niveum	CBS 187.60	watermelon	-	-
niveum	CBS 418.90	watermelon	-	-
niveum	CBS 419.90	watermelon	-	-
opuntarium	NRRL 28368	opuntia	-	-
opuntarium	NRRL 28279	opuntia	-	-
opuntarium	NRRL 28363	opuntia	-	-
radicis-cucumerinum	29	cucumber	-	-
radicis-cucumerinum	33	cucumber	-	-
radicis-cucumerinum	Afu-68(A)	cucumber	-	-
radicis-cucumerinum	Afu-72	cucumber	-	-
radicis-cucumerinum	Afu-58	cucumber	-	-
radicis-cucumerinum	Afu-4(A)	cucumber	-	-
radicis-cucumerinum	Afu-68(A)	cucumber	-	-
radicis-cucumerinum	00/0092/2	cucumber	-	-
radicis-cucumerinum	8	cucumber	-	-
radicis-cucumerinum	14	cucumber	-	-
radicis-cucumerinum	16	cucumber	-	-
radicis-cucumerinum	20	cucumber	-	-
radicis-cucumerinum	21A	cucumber	-	-
radicis-cucumerinum	22	cucumber	-	-
radicis-cucumerinum	24	cucumber	-	-
radicis-cucumerinum	28	cucumber	-	-
radicis-cucumerinum	30	cucumber	-	-
radicis-cucumerinum	31	cucumber	-	-
radicis-cucumerinum	32	cucumber	-	-
radicis-cucumerinum	34	cucumber	-	-
radicis-cucumerinum	35	cucumber	-	-
radicis-cucumerinum	36	cucumber	-	-
radicis-cucumerinum	38	cucumber	-	-
radicis-cucumerinum	60B	cucumber	-	-
radicis-cucumerinum	Afu-33	cucumber	-	-
radicis-cucumerinum	Afu-29(B)	cucumber	-	-
radicis-cucumerinum	Afu-44(B)	cucumber	-	-
radicis-cucumerinum	Afu-11(A)	cucumber	-	-
radicis-cucumerinum	Afu-3	cucumber	-	-
radicis-cucumerinum	AK-2	cucumber	-	-
radicis-lycopersici	DP83	tomato	+	-
radicis-lycopersici	HRS-SB153R	tomato	+	-
radicis-lycopersici	FRC-O-1090	tomato	+	-
radicis-lycopersici	ATCC 52429	tomato	+	-
radicis-lycopersici	FORL-19R	tomato	+	-
radicis-lycopersici	FORL-C709	tomato	+	-

radicis-lycopersici	FORL-C405B	tomato	+	-
radicis-lycopersici	FORL-C1018F	tomato	+	-
radicis-lycopersici	FORL-C809L	tomato	+	-
radicis-lycopersici	S7	tomato	+	-
radicis-lycopersici	DP95	tomato	+	-
radicis-lycopersici	FORL-C710B	tomato	+	-
radicis-lycopersici	FORL-C696A	tomato	+	-
radicis-lycopersici	FORL-C1058P	tomato	+	-
radicis-lycopersici	FORL-GAR3	tomato	+	-
radicis-lycopersici	FORL-C58M	tomato	+	-
radicis-lycopersici	ATCC 60095	tomato	+	-
radicis-lycopersici	FORL-89-1511	tomato	+	-
radicis-lycopersici	FU-87-1	tomato	+	-
radicis-lycopersici	J-36	tomato	+	-
radicis-lycopersici	FORL-OSU374	tomato	+	-
radicis-lycopersici	FORL-DJV78	tomato	+	-
radicis-lycopersici	FORL-CRH673	tomato	+	-
radicis-lycopersici	FORL-C434	tomato	+	-
radicis-lycopersici	FORL-Pt473C	tomato	+	-
radicis-lycopersici	FORL-Pt473D	tomato	+	-
radicis-lycopersici	HRS-SB082Q	tomato	+	-
radicis-lycopersici	01157	tomato	+	-
radicis-lycopersici	FORL-VK9B	tomato	+	-
radicis-lycopersici	FORL-C1327A	tomato	+	-
radicis-lycopersici	FORL-C734B	tomato	+	-
radicis-lycopersici	DP61	tomato	+	-
radicis-lycopersici	FORL-C815A	tomato	+	-
radicis-lycopersici	DP44	tomato	+	-
radicis-lycopersici	DP37	tomato	+	-
radicis-lycopersici	FORL-C202	tomato	+	-
radicis-lycopersici	FORL-UK3Q	tomato	+	-
radicis-lycopersici	01150	tomato	+	-
radicis-lycopersici	FORL-FL418	tomato	+	-
radicis-lycopersici	MUCL 39794	tomato	+	-
radicis-lycopersici	MUCL 39793	tomato	+	-
radicis-lycopersici	MUCL 39792	tomato	+	-
radicis-lycopersici	MUCL 39791	tomato	+	-
radicis-lycopersici	MUCL 39790	tomato	+	-
radicis-lycopersici	DP282	tomato	+	-
radicis-lycopersici	FORL-C623	tomato	+	-
radicis-lycopersici	FORL-C624A	tomato	+	-
radicis-lycopersici	FORL-C622A	tomato	+	-
radicis-lycopersici	PB9	tomato	+	-
radicis-lycopersici	MUCL 39800	tomato	+	-
radicis-lycopersici	MUCL 39798	tomato	+	-
radicis-lycopersici	CBS 101587	tomato	+	-
radicis-lycopersici	CBS 873.95	tomato	+	-
radicis-lycopersici	CBS 872.95	tomato	+	-
radicis-lycopersici	MUCL 39788	tomato	+	-
radicis-lycopersici	MUCL 39789	tomato	+	-
radicis-lycopersici	MUCL 39799	tomato	+	-
radicis-lycopersici	MUCL 39795	tomato	+	-
radicis-lycopersici	MUCL 38936	tomato	+	-
radicis-lycopersici	MUCL 39797	tomato	+	-
radicis-lycopersici	MUCL 39796	tomato	+	-
radicis-lycopersici	CBS 874.95	tomato	+	-
radicis-lycopersici	FORL-FRC-O-1097K	tomato	+	-
radicis-lycopersici	FORL-C651	tomato	+	-
radicis-lycopersici	01090/B	tomato	+	-
radicis-lycopersici	FORL-C838H	tomato	+	-
radicis-lycopersici	43	tomato	+	-
radicis-lycopersici	46	tomato	+	-

	radicis-lycopersici	NRRL 26033	tomato	+	-
	radicis-lycopersici	NRRL 26379	tomato	+	-
	radicis-lycopersici	NRRL 26381	tomato	+	-
	radicis-lycopersici	41	tomato	+	-
	spinaciae	NRRL 26874	spinach	-	-
	spinaciae	NRRL 26875	spinach	-	-
	spinaciae	NRRL 26876	spinach	-	-
	tulipae	NRRL 28974	tulip	-	-
	tulipae	NRRL 22556	tulip	-	-
	tulipae	NRRL 26954	tulip	-	-
Fusarium solani		CABI 17960	*Solanum tuberosum*	-	-
		CBS 165.87	*Solanum tuberosum*	-	-
Fusarium javanicum		CBS 616.66	*Cucurbita viciifolia*	-	-

Supplementary Table 4. Publically available fungal genomes queried by tBLASTn analyses.

Species	Database	Version
Acremonium alcalophilum	JGI	2
Agaricus bisporus var *bisporus*	JGI	2
Agaricus bisporus var. *burnettii*	JGI	1
Alternaria brassicicola	JGI	1
Aspergillus aculeatus	JGI	1.1
Aspergillus carbonarius	JGI	3
Aspergillus flavus	Broad	2
Aspergillus fumigatus	Broad	1
Aspergillus nidulans	JGI	AspGD
Aspergillus niger	JGI	3
Aspergillus terreus	Broad	1
Auricularia delicata	JGI	1
Batrachochytrium dendrobatidis	Broad	1
Baudoinia compniacensis	JGI	1
Bjerkandera adusta	JGI	1
Botrytis cinerea	Broad	1
Candida albicans WO1	Broad	1
Candida caseinolytic	JGI	1
Candida tenuis	JGI	1
Candida tropicalis	Broad	3
Ceriporiopsis subvermispora B	JGI	1
Chaetomium globosum	JGI	1
*Coccidiodes immitis*RS	Broad	1
Coccidiodes posadasii CPA0066	Broad	1
Cochliobolus heterostrophus	JGI	1
Cochliobolus sativus	JGI	1
Colletotrichum graminicola	JGI	1
Colletotrichum higginsianum	MPIZ	1
Coniophora puteana	JGI	1
Coprinopsis cinerea	JGI	1
Cryphonectria parasitica	JGI	2
Cryptococcus neoformans var. *grubii*	JGI	1

Dacryopinax sp.	JGI	1
Dichomitus squalens	JGI	1
Dothistroma septosporum	JGI	1
Fomitiporia mediterranea	JGI	1
Fomitopsis pinicola	JGI	1
Fusarium graminearum	JGI	1
Fusarium oxysporum	JGI	1
Ganoderma sp.	JGI	1
Gloeophyllum trabeum	JGI	1
Hansenula polymorpha	JGI	2
Heterobasidion annosum	JGI	2
Histoplasma capsulatum NAm1	Broad	1
Hysterium pulicare	JGI	1
Laccaria bicolor	JGI	2
Leptosphaeria maculans	JGI	1
Magnaporthe oryzae	Broad	6
Malassezia globosa	JGI	1
Melampsora laricis-populina	JGI	1
Microsporum canis	Broad	1
Microsporum gypseum	Broad	1
Mycosphaerella fijiensis	JGI	2
Mycosphaerella graminicola	JGI	2
Nectria haematococca	JGI	2
Neurospora crassa	JGI	1
Neurospora discreta FGSC 8579 mat A	JGI	1
Neurospora tetrasperma FGSC 2508 mat A	JGI	2
Neurospora tetrasperma FGSC 2509 mat a	JGI	1
*Paracoccidioides brasiliensis*Pb01	Broad	1
Phanerochaete carnosa	JGI	1
Phanerochaete chrysosporium	JGI	2
Phlebia brevispora	JGI	1
Phlebiopsis gigantea	JGI	1
Pichia membranifaciens	JGI	1
Pichia stipitis	JGI	2
Pleurotus ostreatus PC15	JGI	2
Pleurotus ostreatus PC9	JGI	1
Postia placenta	JGI	1
Puccinia graminis	JGI	1
Punctularia strigosozonata	JGI	1
Pyrenophora teres f. teres	JGI	1
Pyrenophora tritici-repentis	JGI	1
Rhizopus oryzae	JGI	3
Rhodotorula graminis strain	JGI	1.1
Rhytidhysteron rufulum	JGI	1
Saccharomyces cerevisiae	JGI	1
Schizophyllum commune	JGI	1
Sclerotinia sclerotiorum	Broad	2
Septoria musiva	JGI	1
Serpula lacrymans	JGI	2

Serpula lacrymans	JGI	1
Setosphaeria turcica	JGI	1
Spathaspora passalidarum	JGI	2
Sporobolomyces roseus	JGI	1
Sporotrichum thermophile	JGI	2
Stagonospora nodorum	JGI	2
Stereum hirsutum	JGI	1
Thielavia terrestris	JGI	2
Trametes versicolor	JGI	1
Tremella mesenterica Fries	JGI	1
Trichoderma atroviride	JGI	2
Trichoderma reesei	JGI	1
Trichoderma reesei	JGI	2
Trichoderma viren	JGI	2
Trichophyton equinum	Broad	1
Trichophyton rubrum	Broad	2
Trichophyton tonsurans	Broad	1
Ucinocarpis reesii	Broad	2
Ustilago maydis	JGI	1
Verticillium albo-atrum	Broad	1
Verticillium dahliae	Broad	1
Wallemia sebi	JGI	1
Wickerhamomyces anomalus	JGI	1
Wolfiporia cocos	JGI	1

Supplementary Table 5. Primers used in this study.

ID	Sequence	Target/Purpose
PVX-Ave1-F	CACCGAATTCATCGATATGAAGCTTTCTACGCTTGGAG	PVX
PVX-Ave1-R	CACCAAGCTTGAATTCGCGGCCGCTTATATCTGTCTAAATTCGATGTTGACC	PVX
AT-Ave1-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAAGCTTTCTACGCTT	ATTA
AT-Ave1-R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTATATCTGTCTAAATTC	ATTA
AT-FoAve1-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAAACTACTCGCACTA	ATTA
AT-FoAve1-R	GGGGACCACTTTGTACAAGAAAGCTGGGTTCATCTTTGTACAAAATCGATATTT	ATTA
KO-Ave1-F1	GGTCTTAAUCCAGTAGTTCGCAACTTCCAA	Knock-out (left)
KO-Ave1-R1	GGCATTAAUAAAGGATGGCGCGAAGAC	Knock-out (left)
KO-Ave1-F2	GGACTTAAUTCAACACCTTAAATCCCCCTA	Knock-out (right)
KO-Ave1-R2	GGGTTTAAUGACCTCGTAGGAGGACGCTAC	Knock-out (right)
CO-Ave1-F	GAATTCTTAATTAAAGCCATGTCCACATGTGGTTC	Genomic complementation
CO-Ave1-R	GAATTCTTAATTAATCCCAGACCTCGTAGGAGG	Genomic complementation
VdAve1F	AAGGGGTCTTGCTAGGATGG	Ave1 allelic variation
VdAve1R	TGAAACACTTGTCCTCTTGCT	Ave1 allelic variation
FoAve1-F	TCCCTTTTCACGCTCCTACT	FoAve1 allelic variation
FoAve1-R	GACAGATGCAGATTGCTGGA	FoAve1 allelic variation
VdELF-1a-F	CCATTGATATCGCACTGTGG	Biomass qPCR
VdELF-1a-R	TGGAGATACCAGCCTCGAAC	Biomass qPCR
AtRub-F3	GCAAGTGTTGGGTTCAAAGCTGGTG	Biomass qPCR
AtRub-R3	CCAGGTTGAGGAGTTACTCGGAATGCTG	Biomass qPCR
RACE-AVE1-FWD	CAGCAATCCCAGCCAATTTCCCTCTG	Race PCR
RACE-AVE1-REV	CTTGCAGGACCCTCTAGCACCACTG	Race PCR
5UTR-AVE1-FWD	CTTCACTCTGCTCTCGTACAG	Intron verification
RACE-NESTED-F	GTTCGTTGCTGTTTCAGACGGTCTCTG	Race PCR

Chapter 5

Chromosome plasticity drives asexual genome evolution; birth of pathogen effector genes

de Jonge R, Bolton MD, Kombrink A, Yadeta KA, van den Berg GCM and Thomma BPHJ

Abstract

Pathogen populations that undergo regular sexual reproduction are thought to pose a great risk to agriculture because they can recombine alleles that contribute to virulence in the face of dynamic environmental conditions. However, whereas strictly asexual microorganisms are often considered as evolutionary dead ends, they comprise many devastating plant pathogens. Here, we investigated variation in the asexual plant pathogenic fungus *Verticillium dahliae*, and discovered that extensive chromosomal rearrangement establishes highly dynamic 'plastic' regions in the genome to generate variation. Such plastic regions occur at the flanks of chromosomal breakpoints, enabling rapid development of novel effector genes that mediate disease development and enable pathogen adaptation.

Introduction

Sexual recombination drives genetic diversity in eukaryotic genomes, and fosters adaptation to new environments [1-3]. However, sexual reproduction comes at a cost because two compatible individuals need to locate each other to generate offspring, and a decrease in fitness due to break up of co-adapted combinations of interacting alleles may occur [3-5]. Although asexual and sexual reproduction occurs within a single species in many fungi and involves transitions between haploid and diploid stages, about 20% of all fungal phyla reproduce strictly asexually [5]. Asexual organisms are thought to be less flexible than sexual ones, relying solely on random mutations to adapt to changing environments, and are considered as evolutionary dead ends [6, 7]. Here, we examined genetic diversity in a population of the asexual fungus *Verticillium dahliae*, a soil-borne broad host-range plant pathogen that invades the water-conducting xylem vessels of susceptible plant species to cause vascular wilt disease [8, 9]. In tomato, resistance against *V. dahliae* race 1 strains is mediated by the Ve1 immune receptor, while race 2 strains escape recognition. To infer patterns of evolution and host adaptation, we compared the genome sequences of 10 recently sequenced tomato-pathogenic race 1 and 2 *V. dahliae* strains that were collected from various geographical regions and hosts [10], as well as the reference genome sequence of strain VdLs.17 [11] (Table S1).

Results and Discussion

Verticillium dahliae re-sequencing identifies core genomic regions with low genetic diversity

By assessing read mapping coverage of all sequenced *V. dahliae* strains over the VdLs.17 reference genome in 1 Kb windows, a core genome shared by all strains was identified that encompassed ~32 Mb of sequence containing 8,562 genes. Using the reference sequence for comparison, we subsequently assessed single nucleotide polymorphisms (SNPs) within the core genome, which ranged from 5,445 (for JR2, ~99.98% identity) to 163,602 (for St.100, ~99.5% identity) SNPs per strain (Table S2), collectively amounting to 236,785 non-redundant polymorphic sites. Of these SNPs, 78,342 (32.9%) occurred in protein-coding regions of which 55% were synonymous, not affecting the protein sequence, and 45% were non-synonymous (Table S2). To determine selection strength, the ratio of non-synonymous (dN) substitutions per non-synonymous

site (K_a) to the number of synonymous substitutions (dS) per synonymous site (K_s) was calculated within the coding regions for each of the 8,562 core genes [12,13] In total, only 28 genes are under positive selection (K_a/K_s >1; $P < 0.01$), of which 4 encode secreted proteins that are candidate effectors that modulate host physiology to enable host colonization [14] (Table S3).

Large-scale intra- and inter-chromosomal rearrangements within *V. dahliae*

To infer evolutionary relations within the population, all SNP positions were used to construct a phylogenetic tree (Figure 1). Despite the high degree of conservation between strains, JR2 formed a clearly separate cluster with the VdLs.17 reference strain and was analysed in further detail. To improve the *de novo* JR2 genome assembly, mate-pair sequencing (1 Gb) of a 5 Kb-insert library was performed, leading to a drastically improved contig N50 (minimum contig length such that the sum of contigs equal to this length or longer contains 50% of the entire assembly) from 59.4 Kb to 1.0 Mb, and decreased number of contigs from 4,753 to 267. We used optical mapping to place ~30% of the assembled contigs on 8 scaffolds, covering ~94% of the total assembly (34 Mb) (Table S4). Surprisingly, while the chromosome lengths of the reference strain vary between 3.1 and 6.0 Mb [11], the JR2 chromosome lengths were found to vary up to 8.9 Mb (Table S5). These remarkable differences reinforce the notion that *V. dahliae* strictly relies on asexual reproduction, as correct pairing of homologous chromosomes during the prophase of meiosis would be impossible [15]. To compare chromosomal content, pairwise alignments were used to identify collinear (synteny) blocks between the two strains. This analysis identified extensive regions of synteny which are repeatedly interrupted by intra- and inter-chromosomal rearrangements (Figure 2A). In total, 11 intra- and 17 inter-chromosomal rearrangements were identified between the two strains (Figure 2A). Some breakpoints could be pinpointed to the nucleotide and could be experimentally confirmed (Figure S1, Table S6; S7). However, most breakpoints were associated with assembly gaps that are likely caused by repeat-rich areas (Figure 2B, Table S7). Particularly, retrotransposons have been implicated in genome rearrangements through homologous recombination between repeated elements or by causing chromosomal breaks during excision or insertion [16, 17]. To investigate the role of repetitive elements in more detail, we identified all of them within the two genomes (Table S8) and observed significant correlation ($p < 0.01$) between synteny breakpoints and the presence of long terminal repeat (LTR) retrotransposons, but not other classes of repetitive elements.

Since JR2 is phylogenetically closest to the reference strain VdLs.17 (Figure 1), we expected that the extent of chromosomal rearrangements was greater in the other strains when compared to the reference. Therefore, we extended the analysis of chromosomal rearrangements to all sequenced strains by screening for breakpoints within the alignments to the reference strain. Although identification of such syntenic breakpoints in the small insert library genome assemblies was hampered due to the relatively short contig lengths, we could identify 3-8 synteny breakpoints for each of these strains. Furthermore, pulsed-field gel electrophoresis confirmed considerable chromosome length polymorphism between all strains (Figure 1; Figure S2). The extent of the karyotype variation is surprising since it does not concern absence-presence

polymorphisms of small dispensable chromosomes such as has previously been demonstrated for various other plant pathogens [18-21]. Although chromosome length polymorphism is common within fungal species, fungal karyotypes are typically mitotically stable [22]. The apparent variation observed among fungal karyotypes [22] may be the result of complex chromosomal rearrangements like those we identified in this study.

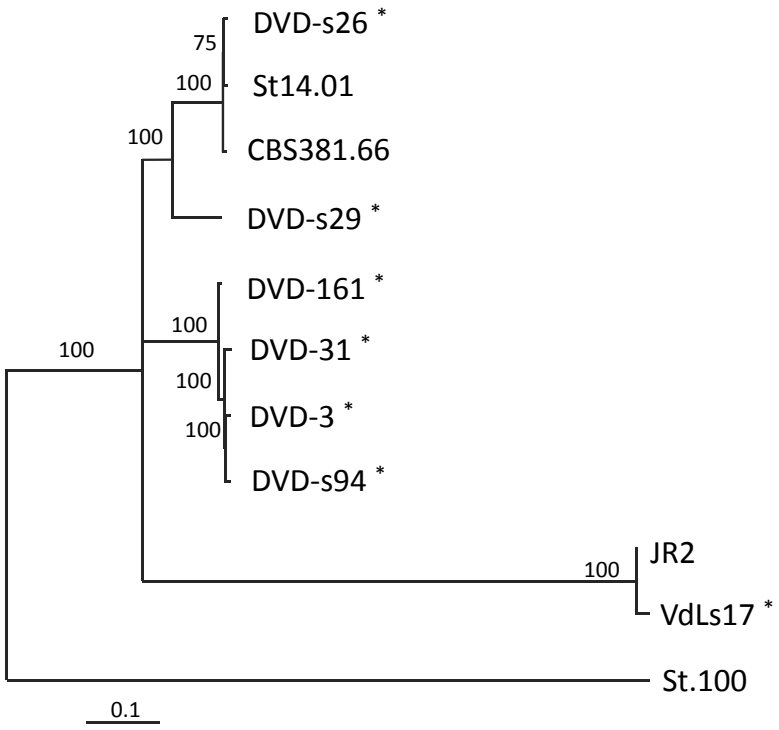


Figure 1. Population structure of sequenced *Verticillium dahliae* isolates. Unrooted maximum likelihood phylogenetic tree based on concatenation of 236,785 SNP sites relative to reference strain VdLs.17. Evolutionary distances based on the Jukes-Cantor method and bootstrap support (%)indicated at the nodes. Race 1 strains indicated by asterisks.

Chromosomal rearrangements contribute to genetic diversity and drive evolution

In addition to the core ~32 Mb genome, all strains carried up to ~1 Mb of genome sequence that was unique or shared by only few strains, composing a highly dynamic 'plastic' region of the genome encoding up to 1,500 genes. These plastic genomic regions are correlated ($p < 0.01$) with syntenic breakpoints in VdLs.17 and JR2 (Table S7). Recently, the race 1-specific effector that is recognized by the tomato Ve1 immune receptor was identified as Ave1; a secreted virulence factor that is required for full aggressiveness on plants lacking Ve1 [10]. Phylogenetic analysis supports a polyphyletic origin of race 1 and race 2 *V. dahliae* strains (Figure 1). Intriguingly, the *Ave1* gene is located within the ~1 Mb plastic region of race 1 strains, suggesting that these regions contribute to niche adaptation; pathogenicity on plant hosts [10]. To further establish the role of plastic regions in *V. dahliae* pathogenicity, we mined these regions, as well as the core genomic regions in VdLs.17 and JR2, for candidate effectors. No remarkable differences were observed in the amount of secretome components, Pfam protein domain composition, or gene density between core and plastic regions (Figure S3). However, 7 out of the 11 *V. dahliae* genes that were most highly induced during infection of *N. benthamiana* [11, 23] are located within the ~1 Mb plastic region of the JR2 genome and include

Ave1, while only 4 are present in the ~32 Mb core genome, indicating a significant over-representation of *in planta*-induced genes in the plastic regions (Figure 3). Among these 11 genes, 8 encode potential effectors of which 6 are located within the plastic regions (Figure 3). Similarly, evidence can be found for development of genes encoding novel virulence factors within the plastic regions of the VdLs.17 genome. Whereas all other *V. dahliae* strains that were sequenced contain 6 genes that encode LysM effectors [24,

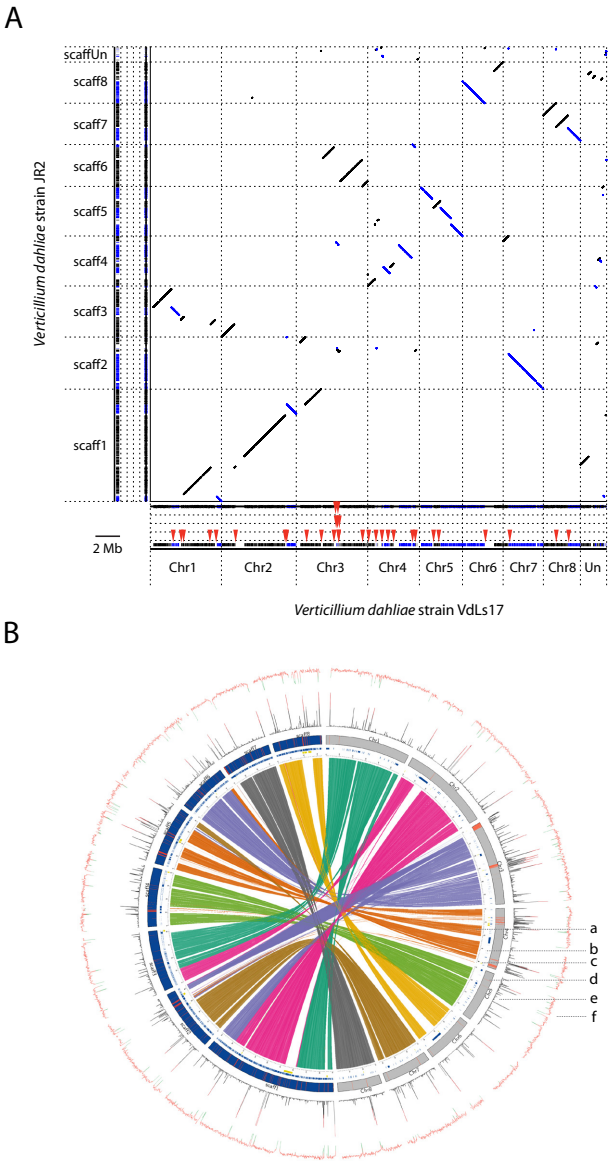


Figure 2. Whole-genome alignment of *Verticillium dahliae* strains VdLs.17 and JR2 reveals extensive chromosomal rearrangements. A) Whole-genome dot-plot comparison with forward-forward alignments (black) and inversions (blue). Red triangles mark syntenic breakpoints. Un: unplaced contigs during optical mapping. B) Circos diagram illustrating collinear blocks with alignments between VdLs.17 (grey) and JR2 (blue) chromosomes (a), sequence gaps (b), sequences aligning to unpositioned scaffolds (yellow) (c), unique sequences (red) (d), repeat density (% coverage of 10-Kb window) (e), and GC % (per 10-Kb window) (f).

25], the reference strain VdLs.17 contains an additional LysM effector gene within a VdLs.17-specific region. Remarkably, there was no significant expression *in planta* for any of the 6 conserved LysM effectors, and only expression of the VdLs.17-specific LysM effector gene *VDAG_05180* was found (Figure 4A). Targeted deletion confirmed that *VDAG_05180* is required for full virulence of VdLs.17 on tomato (Figure 4B, C). Notably, *VDAG_05180* is located within an extensively duplicated region, suggesting that duplication and subsequent diversification mediated its evolution (Figure S4).

Pathogen effector genes are frequently under selection pressure. Various mechanisms for natural variation have been described [26], including diversity in genomic locations enriched for transposons, mutation, and recombination in subtelomeric regions [27-29], co-regulated gene clusters [30, 31], mobile pathogenicity chromosomes [19], low gene density genomic regions [32], or AT-rich isochore-like regions [33, 34]. Often this involves fungi that can reproduce sexually, and of which the genomes were shaped by repeat-driven expansion [26]. In this study on the genome of the asexual haploid fungus *V. dahliae* that contains only a limited amount of repetitive DNA (4%, Table S8), we discovered a novel mechanism for evolution of pathogenicity. It has previously been suggested that karyotype variation, rather than being a mechanism of adaptation to generate novel virulence traits, occurs because non-deleterious genomic rearrangements are maintained due to the absence or rarity of a sexual cycle [22, 35]. However, our data challenge this hypothesis by showing that chromosomal plasticity, evidenced by extensive targeted chromosomal rearrangements and karyotype variability, is likely genetically non-neutral as it induces local

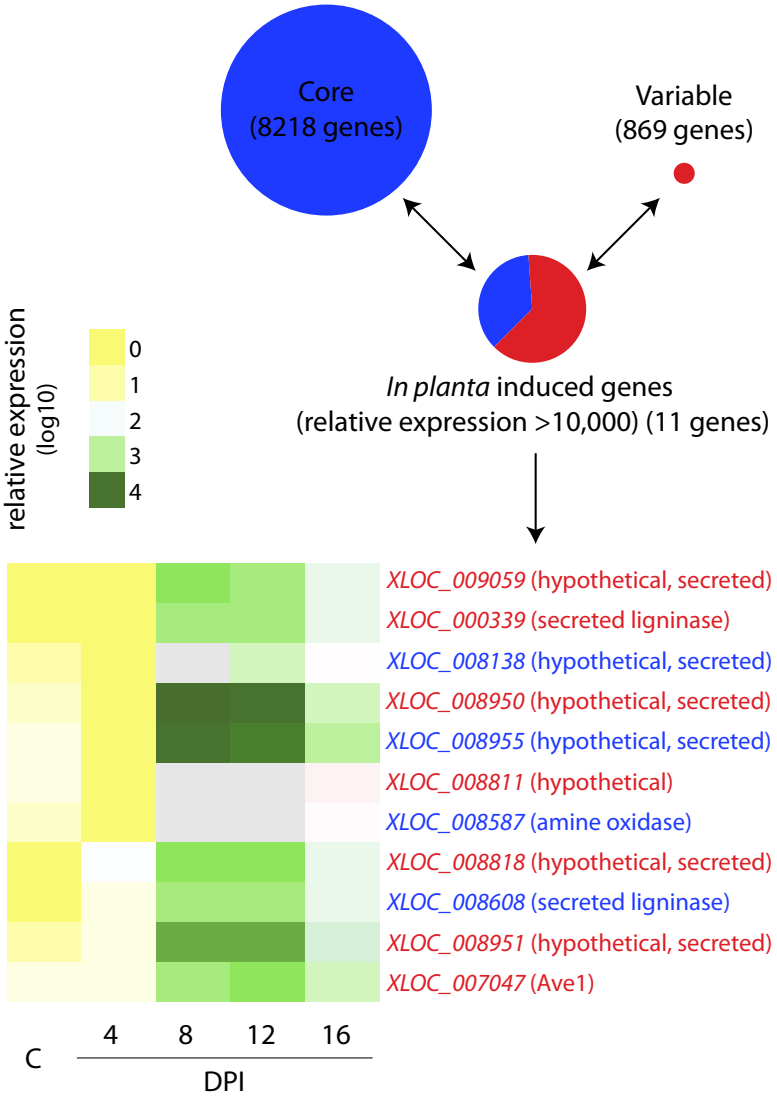


Figure 3. Variable genomic regions of *Verticillium dahliae* strain JR2 are enriched for *in planta*-expressed genes. Deep transcriptome sequencing of infected *Nicotiana benthamiana* plants harvested between 4 and 16 days post inoculation DPI and *in vitro* cultured (control, C) fungus. Reads were identified for 9,087 JR2 genes, of which 8,218 are on the core and 869 on the variable genome. Of the 11 most highly *in planta*-induced genes, 7 are within variable regions while 4 are in the core genome, demonstrating significant over-representation of *in planta*-expressed genes in variable regions.

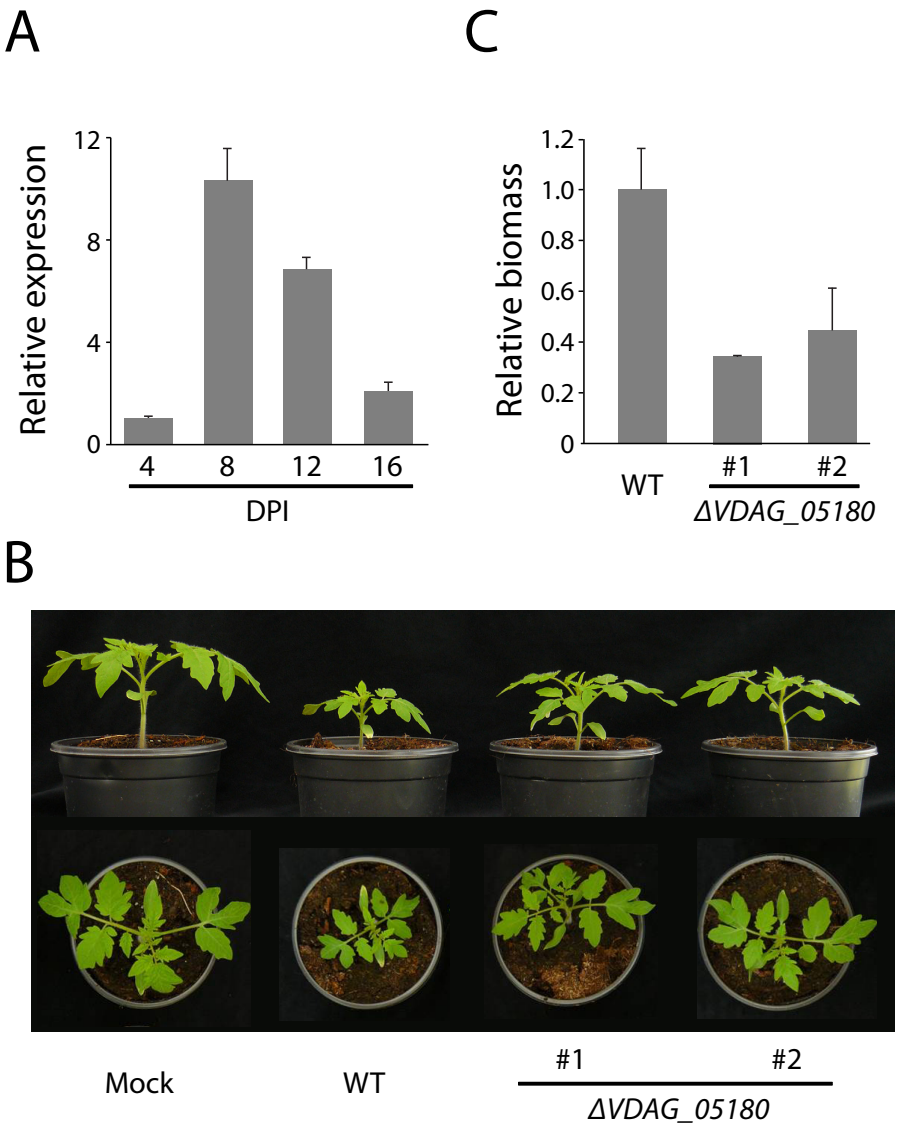


Figure 4. *Verticillium dahliae* strain VdLs.17-specific LysM effector VDAG_05180 is required for virulence on tomato. A) Expression of *VDAG_05180* during infection of *Nicotiana benthamiana* between 4 and 16 days post inoculation (DPI). B) Two independent *VDAG_05180* deletion strains (*ΔVDAG_05180*) show compromised virulence on tomato, evidenced by reduced stunting when compared with inoculation with wild-type *V. dahliae* (WT). Photographs are taken at 12 DPI. C). Reduced fungal biomass in plants inoculated with two independent *VDAG_05180* deletion strains when compared with wild-type *V. dahliae* (WT) at 8 DPI. Error bars represent standard error of three replicate experiments.

variation at syntenic breakpoints and increases adaptive capability. The highly dynamic plastic genomic regions are enriched for *in planta*-induced genes, including effector genes that contribute to virulence. Although it is generally believed that asexual reproduction limits genetic variation, and consequently limits adaptive capability, we here provide evidence for chromosomal plasticity as a mechanism that allows asexual haploid genomes to adapt to changing environments.

Acknowledgements

We thank X. Wang for technical assistance and T. Friesen, H. de Jong, F. Debets, P. de Wit, M. Joosten and J. van Kan for helpful discussions.

References

1. Goddard MR, Godfray HCJ, Burt A, **Sex increases the efficacy of natural selection in experimental yeast populations.** *Nature*, 2005, **434**:636-640.
2. Colegrave N, **Sex releases the speed limit on evolution.** *Nature*, 2002, **420**:664-666.
3. de Visser JAGM, Elena SF, **The evolution of sex: empirical insights into the roles of epistasis and drift.** *Nat Rev Genet*, 2007, **8**:139-149.
4. Agrawal AF, **Evolution of sex: why do organisms shuffle their genotypes?** *Curr Biol*, 2006, **16**:R696-704.
5. Heitman J, Kronstad JW, Taylor JW, L.A. Casselton, eds. ***Sex in Fungi: Molecular Determination and Evolutionary Implications***, 2007, Washington DC: ASM Press.
6. McDonald BA, Linde C, **Pathogen population genetics, evolutionary potential, and durable resistance.** *Annu Rev Phytopathol*, 2002, **40**:349-379.
7. Burt A, **Perspective: sex, recombination, and the efficacy of selection--was Weismann right?** *Evolution*, 2000, **54**:337-351.
8. Fradin EF, Thomma BPHJ, **Physiology and molecular aspects of *Verticillium* wilt diseases caused by *V. dahliae* and *V. albo-atrum*.** *Mol Plant Pathol*, 2006, **7**:71-86.
9. Klosterman SJ, Atallah ZK, Vallad GE, Subbaroa KV, **Diversity, pathogenicity, and management of *Verticillium* species.** *Annu Rev Phytopathol*, 2009, **47**:39-62.
10. de Jonge R, *et al.*, **Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing.** *Proc Natl Acad Sci USA*, 2012, **109**:5110-5115.
11. Klosterman SJ, *et al*., **Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens.** *PLoS Pathog*, 2011, **7**:e1002137.
12. Yang Z, Nielsen R, **Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models.** *Mol Biol Evol*, 2000, **17**:32-43.
13. Stukenbrock EH, *et al*., **The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species.** *Genome Res*, 2011, **21**:2157-2166.
14. de Jonge R, Bolton MD, Thomma BHPJ, **How filamentous pathogens co-opt plants: the ins and outs of fungal effectors.** *Curr Opin Plant Biol*, 2011, **14**:400-406.
15. Kistler HC, Miao VP, **New modes of genetic change in filamentous fungi.** *Annu Rev Phytopathol*, 1992, **30**:131-153.
16. Mieczkowski PA, Lemoine FJ, Petes TD, **Recombination between retrotransposons as a source of chromosome rearrangements in the yeast *Saccharomyces cerevisiae*.** *DNA Repair*, 2006, **5**:1010-1020.
17. Maxwell PH, Burhans WC, Curcio J, **Retrotransposition is associated with genome instability during chronological aging.** *Proc Natl Acad Sci USA*, 2011, **108**:20317-20324.
18. Coleman JJ, *et al.*, **The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion.** *PLoS Genet*, 2009, **5**:e1000618.
19. Ma L-J, *et al.,* **Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*.** *Nature*, 2010, **464**:367-373.
20. Stukenbrock EH, *et al.*, **Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen *Mycosphaerella graminicola*.** *PLoS Genet*, 2010, **6**:e1001189.
21. Goodwin SB, *et al.,* **Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis.** *PLoS Genet*, 2011, **7**:e1002070.
22. Zolan ME, **Chromosome-length polymorphism in fungi.** *Microbiol Rev*, 1995, **59**:686-698.
23. Faino L, de Jonge R, Thomma BPHJ, **The transcriptome of *Verticillium dahliae*-infected *Nicotiana benthamiana* determined by deep RNA sequencing.** *Plant Sig Behav*, 2012, **7**:9.

24. de Jonge R, Thomma BPHJ, **Fungal LysM effectors: extinguishers of host immunity?** *Trends Microbiol*, 2009, **17**, 151-157.

25. de Jonge R, *et al.*, **Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants.** *Science*, 2010, **329**:953-955.

26. Raffaele S, Kamoun S, **Genome evolution in filamentous plant pathogens: why bigger can be better.** *Nature Rev Microbiol*, 2012, **10**:417-430.

27. Fedorova ND, *et al*., **Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*.** *PLoS Genet*, 2008, **4**:e1000046.

28. McDonagh A, *et al*., **Sub-telomere directed gene expression during initiation of invasive Aspergillosis.** *PLoS Pathog*, 2008, **4**:e1000154.

29. Chuma I, *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.

30. Pallmer JM, Keller NP, **Secondary metabolism in fungi: does chromosomal location matter?** *Curr Opin Microbiol*, 2010, **13**:431-436.

31. Schirawski J, *et al.*, **Pathogenicity determinants in smut fungi revealed by genome comparison.** *Science*, 2010, **330**:1546-1548.

32. Raffaele S, *et al.,* **Genome evolution following host jumps in the Irish potato famine pathogen lineage.** *Science*, 2010, **330**:1540-1543.

33. van de Wouw AP, *et al.,* **Evolution of linked avirulence effectors in *Leptosphaeria maculans* is affected by genomic environment and exposure to resistance genes in host plants.** *PLoS Pathog*, 2010, **6**:e1001180.

34. Rouxel T, *et al.,* **Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by repeat-induced point mutations.** *Nature Commun*, 2011, **2**:202.

35. Talbot NJ, Salch YP, Ma M, Hamer JE, **Karyotypic variation within clonal lineages of the rice blast fungus, *Magnaporthe grisea*.** *Appl Environ Microbiol*, 1993, **59**:585-593.

Supplementary Materials

Materials and Methods; Supplementary Figures S1, S2, S3, Supplementary Tables S1, S2, S3, S4, S5, S6, S7, S8 and Supplementary References.

Materials and Methods

Mapping, SNP analysis and phylogeny.

Illumina reads (100 bp, paired-end with insert size 500 bp) were mapped onto the *Verticillium dahliae* VdLs.17 reference genome using GSNAP (version 2012-04-16; [1]) with default parameter settings. Mapping results were further processed by Picard (http://picard.sourceforge.net) to mark duplicates. SNPs were identified by samtools mpileup (-I, -E, -u) plus bcftools and filtered using vcfutils.pl varFilter (-Q20, -D100) (http://samtools.sourceforge.net). Finally we selected SNPs with a minimum allele frequency (AF) of 0.8. General SNP statistics were determined by VCFtools [2] and variant annotation and effect predictions were performed by snpEff (version 2.1b; http://snpeff.sourceforge.net).

Breath of coverage was calculated for each of the 10,535 genes [3]as the percentage of nucleotides with at least one read aligned using the BEDtools suite [4]. Genes were considered absent when breath was lower than 0.2. Similarly, breath of coverage was calculated for 1 Kb windows, and considered absent when breath was lower than 0.2.

Maximum likelihood phylogenetic analysis was performed in MEGA5 [5]. For each strain, as input we extracted for all non-redundant SNP position the respective base call, resulting in 11 sequences, each containing 236,785 base pairs.

Genome assembly, whole-genome sequence alignment and identification of re-arrangements

Draft assemblies generated previously [6] were used for all strains, except for JR2. Mate-pair library preparation and sequencing (50 bp, mate-pair with insert size ~5 Kb) of strain JR2 was performed by the Beijing Genome Institute (Hong Kong). For *de novo* assembly, Velvet [7] was used with the following settings: cov_cutoff=6, exp_cov=auto, k-mer=31 and shortMatePaired=yes, to generate a new assembly by including both the 500 bp paired-end library as well as the 5 Kb mate-pair library. Optical mapping, i.e. the construction of ordered genome-wide, high-resolution (>150X) restriction maps, are generated for single, stained DNA molecules, was performed by the Beijing Genome Institute (Hong Kong) using the Argus® System (OpGen, USA). The scaffolds generated by Velvet were subsequently placed on the optical map using MapSolver version 3.2 (OpGen, USA).

Whole-genome sequence alignments and dot plots were generated by MUMMER3 [8] using the Nucmer script with default settings (except for -l 15 and --maxmatch) and mummerplot, respectively. We used custom Perl scripts and the Dnadiff script, part of the MUMMER3 package to identify re-arrangements and the associated breakpoints. Core and variable genomic regions were determined from the whole-genome alignments, by assessing breadth of coverage of the alignments on 1 Kb non-overlapping windows [4]. In addition we determined the percentage of gaps for each region. Regions were considered absent when alignment breadth and the percentage of gaps was below 0.2 and the amount of overlapping repetitive sequence was not more than 80%. Core genomic regions were defined based on presence in all strains. Collinear (synteny) blocks, repeat content, percentage GC and the presence of variable genomic regions were plotted on the VdLs.17 and JR2 genomes using the Circos and GBrowse programs [9,10].

Multiple chromosomal re-arrangements were verified by polymerase chain reaction (PCR). To this end, primer pairs spanning predicted breakpoints were designed for both JR2 and VdLs.17 to selectively amplify breakpoint regions in either of the two genotypes, and then used in PCR reactions on genomic DNA of JR2 and VdLs.17 . Quality of input DNA and integrity of syntenic sequences flanking the breakpoints in JR2 and VdLs.17 were verified by control primer sets (Table 6). PCR conditions were as follows: an initial 95 C denaturation step for 5 min followed by denaturation for 15 s at 95 C and annealing for 30s at 58 C and extension at 72 C for 35 cycles.

Repeat identification

Repetitive elements were identified and classified by the RepeatModeler program (http://www.repeatmasker.org/RepeatModeler.html) which includes repeat identification using both RepeatScout [11] and Recon [12], and repeat classification using the RepBase library (version 16.12). For repeat identification we applied default RepeatModeler settings. Repetitive sequences were then used as an external library in the RepeatMasker (http://www.repeatmasker.

org) program to mask repetitive sequence applying the sensitive mode (-s). In addition, full-length long terminal repeat (LTR) retrotransposons were identified by the LTR_FINDER [13]. Statistical significant spatial correlation between repetitive sequences, breakpoints and variable genomic regions was assessed using the R package GenometriCorr (http://genometricorr.sourceforge.net).

Karyotyping

Mycelium of *V. dahliae* was prepared for protoplasting following the mycelium-based fungal biomass preparation method described by Mehrabi *et al*. [14]. Mycelium was digested in 1 M sorbitol containing 1% (w/v) glucanex (Sigma-Aldrich, USA) and 0.5% driselase (Sigma-Aldrich) at 32°C until protoplast concentration reached 108 ml-1 (~3 to 4 hours). Protoplast plugs were generated using the method of Mehrabi *et al*. [14] and stored at 4°C until use. Karyotyping was carried out using a CHEF Mapper XA pulsed field electrophoresis system (Bio-Rad, USA) outfitted with a cooling module (Bio-Rad) using the auto algorithm function with low molecular weight low set to 2 Mb and high molecular weight set to 6 Mb. All other parameters were default settings. Chromosomes were separated in 0.8% low EEO (EP) agarose (US Biological, USA) gels. Running buffer (1X TAE) was changed daily during each electrophoresis run. Chromosome size markers from *Hansenula wingei* and *Schizosaccharomyces pombe* (Bio-Rad) were included in each gel. After electrophoresis, gels were stained with ethidium bromide (1 µg ml-1) in water for 1 h, and then destained in water for 2 h.

Gene prediction, expression and annotation

Initially, *ab initio* gene predictions in *V. dahliae* strains were performed using the Augustus gene prediction software, applying the Fusarium training parameters [15].

For deep transcriptome sequencing and mapping, ~2 Gb of reads from *V. dahliae* strain JR2 infected *Nicotiana benthamiana* plants, and ~1 Gb of reads from *V. dahliae* JR2 cultured *in vitro* on Czapek Dox medium (control, C) were mapped on the *V. dahliae* JR2 genome using TopHat [6,16]. Cufflinks [17] was then used to assemble transcripts and isoforms from all of the mapped reads. Relative expression for each gene in each experiment was determined by Cufflinks and reported in the number of Fragments Per Kilobase of transcript per Million mapped reads (FPKM). For differential gene expression analysis we compared *in planta*-expression with *in vitro*-expression in *V. dahliae* using Excel and R, applying the CummeRbund package (http://compbio.mit.edu/cummeRbund/). Finally; aligned transcript evidence was used as hints for gene prediction by Augustus [18].

The predicted proteins were mined for candidate effectors [19] by functional annotation using BLASTp against the non-redundant database at NCBI (http://www.ncbi.nih.nlm.gov) and Pfam domain scanning [20]. Secreted proteins were predicted by a combination of SignalP4 [21] and WoLF PSORT [22].

Gene density was determined as previously described [23]. In short, 5'- and 3'-flanking intergenic regions were calculated, scored in two-dimensional bins and plotted. Filled contour plots were generated using R (http://www.r-project.org/).

***VDAG_05180* functional analysis**

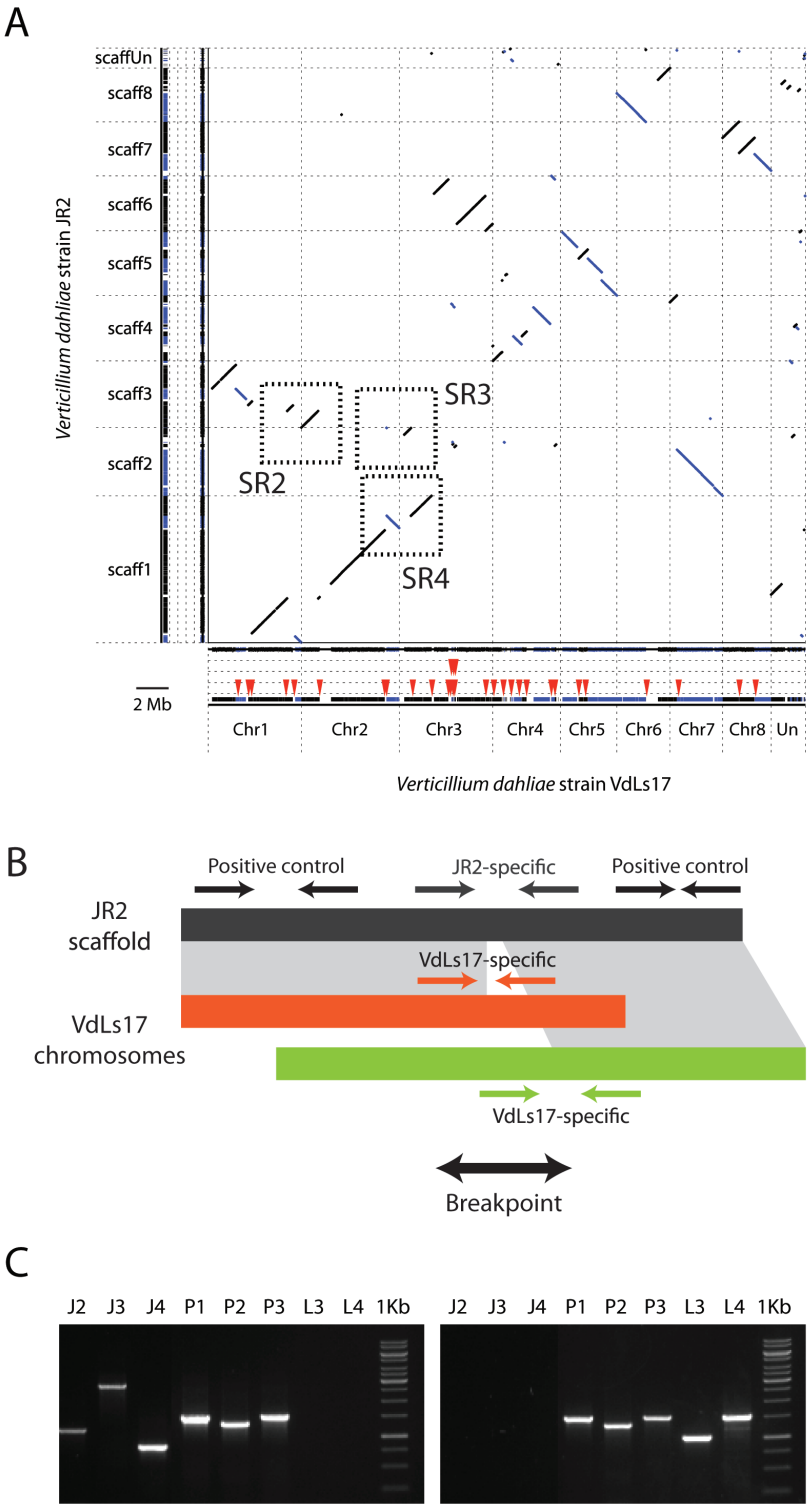
Targeted *VDAG_05180* knockouts were generated by amplifying the sequences flanking the *VDAG_05180* coding sequence using the primers KO-VDAG_05180-F1 and KO-VDAG_05180-R1, and the primers KO-VDAG_05180-F2 and KO-VDAG_05180-R2 (Table S5), and cloned into the vector pRF-HU2 as described [24]. *V. dahliae* transformation and subsequent inoculations on tomato cv. Motelle plants to assess the impact on virulence were performed as described previously [25]. Plants were regularly inspected during a two-week interval and 8 days and 12 days post inoculation photographs were taken.

For biomass quantification the roots and stem below cotyledons of three plants per *V. dahliae* genotype were flash-frozen in liquid nitrogen. The samples were ground to powder, of which an aliquot of approximately 100 mg was used for DNA isolation [26]. Real-time PCR was conducted with primers SlRub-F1 and SlRub-F2 for tomato *RuBisCo* and primers VdGAPH-F and VdGAPH-R for *V. dahliae GAPDH* (Table S5).

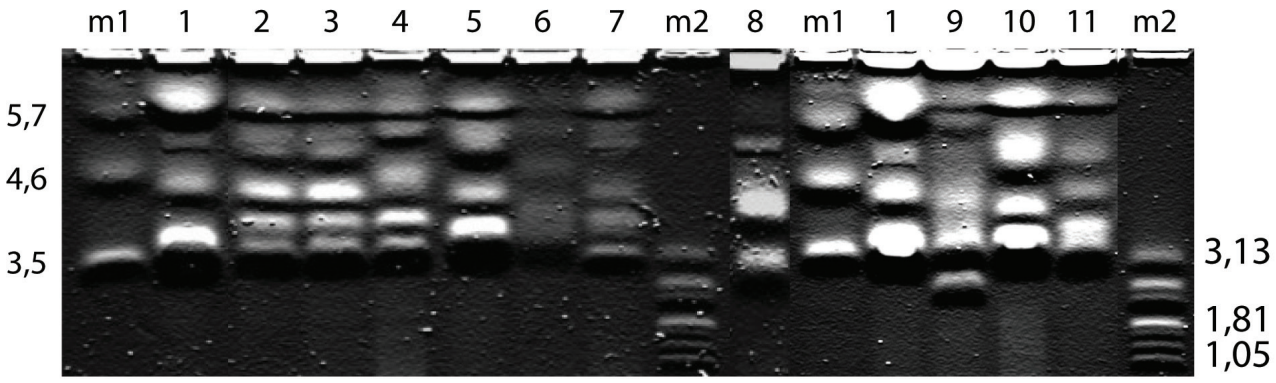
For expression analyses, three-week-old *Nicotiana benthamiana* plants were inoculated with strain VdLs.17 as previously described [25], harvested at 4, 8, 12 and 16 days post inoculation and flash frozen in liquid nitrogen. Total RNA was extracted using the RNeasy Mini Kit (Qiagen, USA), and cDNA was synthesized by SuperScript III (Invitrogen, USA).

). Real-time PCR was conducted with primers VdGAPH-F and VdGAPH-R for *V. dahliae GAPDH* and qVDAG_05180-F1 and qVDAG_05180-R1 for *V. dahliae VDAG_05180* (Table S5).

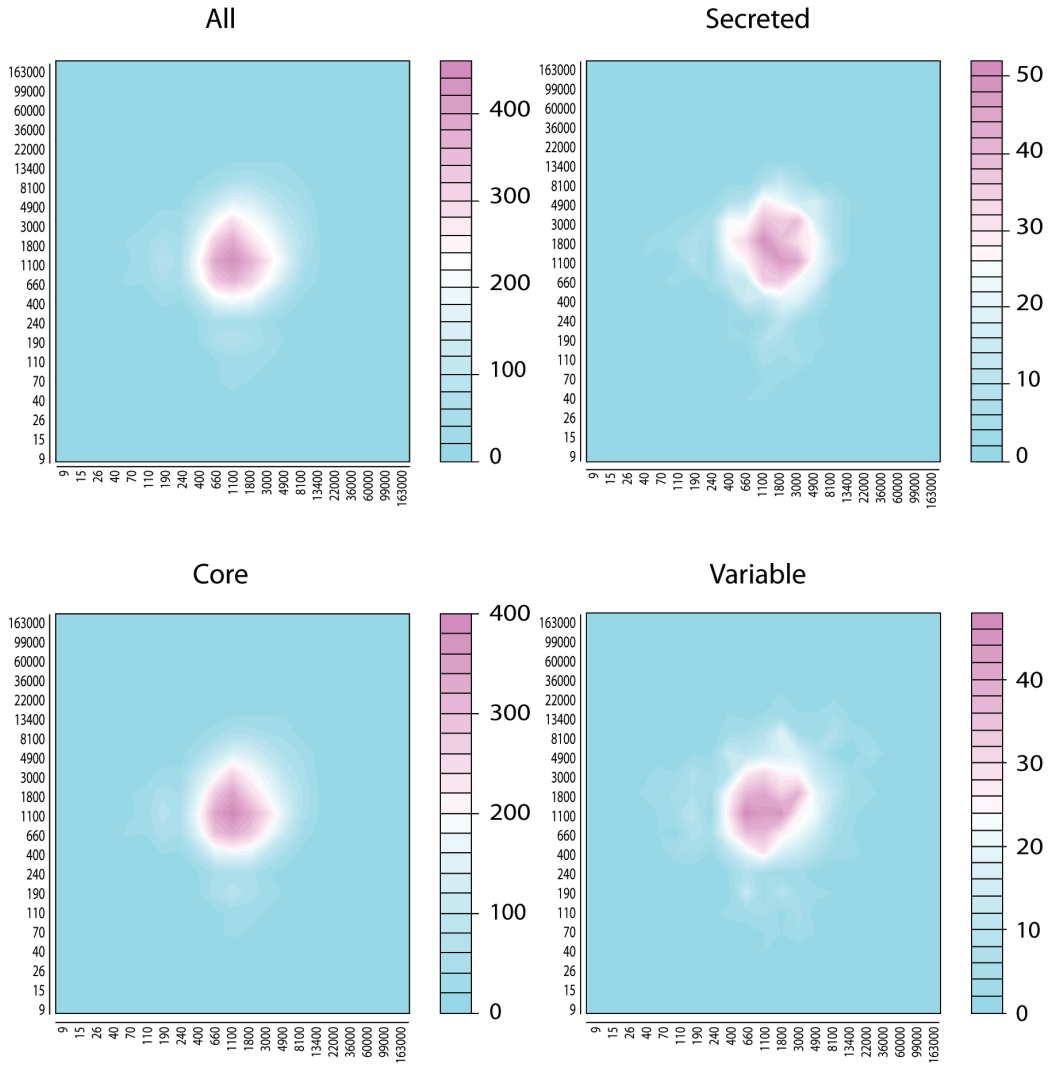
Real-time PCR was conducted using an ABI7300 PCR machine (Applied Biosystems, Foster City, USA) in combination with the qPCR SensiMix kit (BioLine, London, UK). Real-time PCR conditions were as follows: an initial 95°C denaturation step for 10 min followed by denaturation for 15 s at 95°C and annealing for 30s at 60°C and extension at 72°C for 40 cycles.



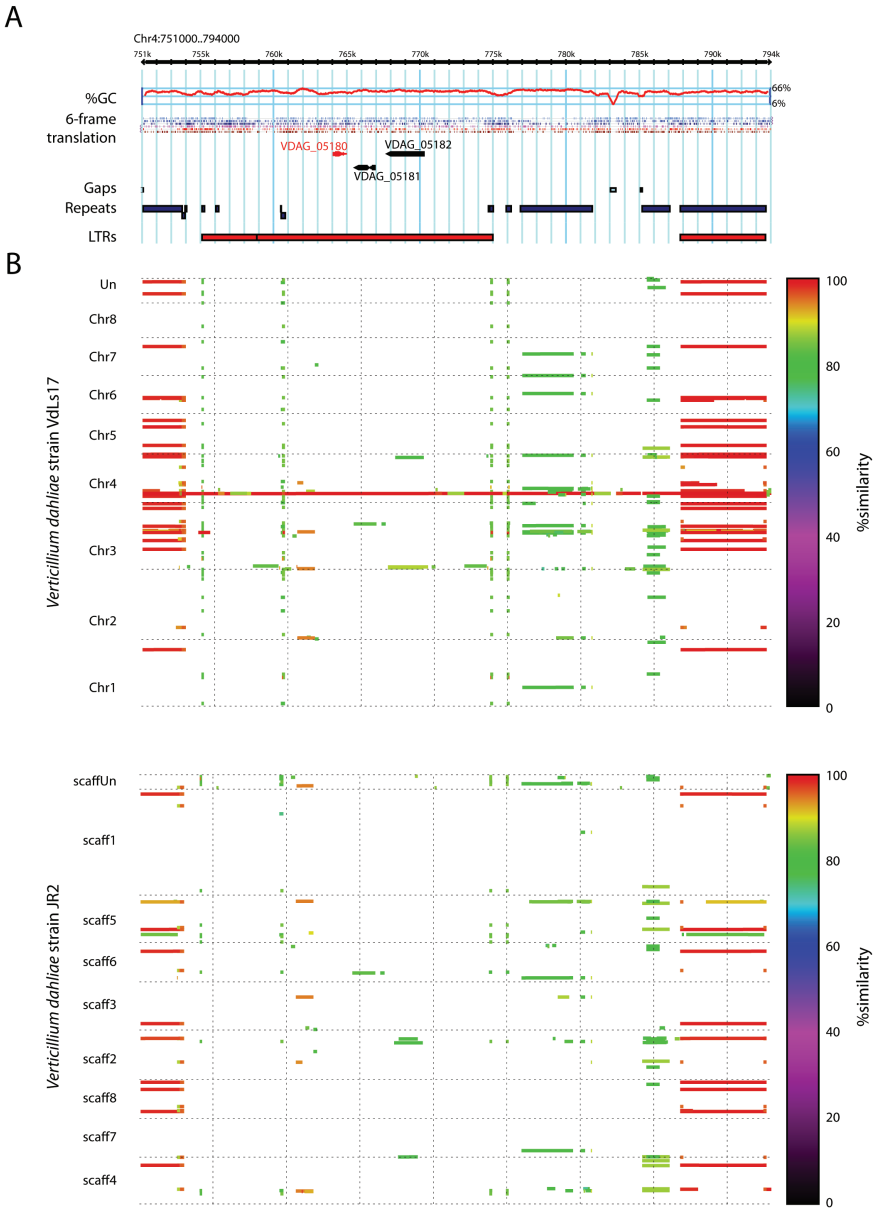
Supplemental Figure S1. Verification of chromosomal rearrangements by PCR. A) Whole-genome alignment of *Verticillium dahliae* strains VdLs.17 and JR2 highlighting three PCR-verified rearrangements. B) Schematic representation of the PCR set-up, illustrating the positions of both JR2 and VdLs.17-specific primer sets as well control primer sets. C) PCR amplification of structural rearrangements SR2, SR3 and SR4 in JR2 (left) and VdLs.17 (right) using JR2-specific primer sets J2, J3 and J4, respectively, and amplification of rearrangements SR3 and SR4 in JR2 and VdLs.17 using VdLs.17-specific primer sets L3 and L4, respectively. Control primer sets are shown by P1, P2 and P3.



Supplemental Figure S2. Pulsed-field gel electrophoresis of sequenced *V. dahliae* strains illustrating chromosome length polymorphisms. Loaded strains are: VdLs.17 (1), CBS381.66 (2), St14.01 (3), St100 (4), DVD-3 (5), DVD-31 (6), DVD-s26 (7), JR2 (8), DVD-s29 (9), DVD-s94 (10) and DVD-161 (11). Chromosomal DNA of *Schizosaccharomyces pombe* (m1) and *Hansenula wingei* (m2) were loaded as size makers.



Supplemental Figure S3. Gene density plots for *V. dahliae* strain JR2. 5'- and 3'-flanking intergenic regions were calculated and placed in two-dimensional bins. The figure was prepared using the *filled.contour* graph function in R.



Supplemental Figure S4. Genomic context of LysM effector gene *VDAG_05180*. A) Genomic location of *VDAG_05180* revealing the presence of flanking gaps, repeats and an overlapping predicted LTR retrotransposon. B) Repetitiveness of the genomic context of *VDAG_05180* throughout the genomes of VdLs.17 and JR2. The degree of conservation is indicated by a color scale, showing recent (red) and more ancient (orange to green) multiplications, especially of repeat elements. In addition, the figure shows presence of *VDAG_05181* (encoding a tetrahydroxynaphthalene reductase) in the JR2 genome, a duplication of this gene in the VdLs.17 genome, and presence of multiple *VDAG_05182* homologs (encoding a Kelch domain-containing protein) in both genomes. The LysM effector gene *VDAG_05180* is uniquely found in the VdLs.17 genome.

Supplemental Table S1. *Verticillium dahliae* strains used in this study.

Strain	Race	Originating host	Location	Year of collection	Tomato[#]	Tobacco*	Arabidopsis*
JR2	1	Tomato	ON, Canada	<1995	++	√	√
CBS381.66	1	Tomato	QC, Canada	1963	++	√	√
St14.01	1	Pistachio	CA, USA	?	++	√	√
St.100	n.a.	Soil	Belgium	?	-	√	√
DVD-3	2	Potato	Canada	1993	+	√	√
DVD-31	2	Tomato	Canada	1993	+	√	√
DVD-161	2	Potato	ON, Canada	1993	+/-	√	√
DVD-S26	2	Soil	Canada	1994	+	√	√
DVD-S29	2	Soil	Canada	1994	+/-	-	√
DVD-S94	2	Soil	Canada	1996	+/-	√	√
VdLs.17	2	Lettuce	USA, CA	1995 ~ 2001	+/-	√	√

[#]Aggressiveness on susceptible tomato ranging from highly aggressive (++; severe stunting, wilting and chlorosis/necrosis), to aggressive (+; stunting and intermediate wilting), mildly aggressive (+/-; intermediate stunting, mild wilting) and non-pathogenic (-; no symptoms visible).

*Ability to infect this host (√) or not (-)

Supplemental Table S2. Summary of SNPs when compared to *V. dahliae* reference strain VdLs.17.

Strain	#SNPs	#Unique SNPs	#SNPs in intergenic region	#SNPs in introns	#SNPs in exons	dN [(a)]	dS [(a)]
JR2	5,445	947	3,193	636	1,563	1,092	444
CBS381.66	117,364	166	67,640	8,775	39,062	16,702	21,974
St14.01	117,704	274	67,894	8,786	39,161	16,744	22,030
St.100	163,602	81061	94,516	12,108	54,219	23,350	30,361
DVD-3	118,528	211	68,773	8,716	39,135	16,662	22,098
DVD-31	119,025	316	68,973	8,786	39,381	16,827	22,170
DVD-161	117,277	178	67,831	8,656	38,776	16,526	21,879
DVD-S26	117,307	235	67,653	8,764	39,079	16,743	21,947
DVD-S29	122,057	5552	70,801	9,204	40,219	17,175	22,647
DVD-S94	119,060	205	68,993	8,782	39,303	16,763	22,161
All	*236,785*	*89,145*	*136,705*	*17,421*	*78,342*	*34,007*	*43,563*

a) Number of non-synonymous (dN) and synonymous substitutions (dS).

Supplemental Table S4. Summary of the optical mapping results.

Statistic	Result
# of contigs	273
# of significant contigs (>30 kb)	75
# of contigs placed	81
% of contigs placed	29.7%
Total size of placed contigs	32.0 Mb
Total size of unplaced contigs	1.2 Mb
% of genome covered	93.62
Number of gaps over 2 kb	51
Average gap size	50.4 kb
Total size of gaps	3.2 Mb

Supplemental Table S5. Chromosome length statistics.

	VdLs.17	JR2
Chr1 / scaff1	5,746,300	9,141,183
Chr2 / scaff2	6,048,892	4,240,912
Chr3 / scaff3	5,770,546	4,159,763
Chr4 / scaff4	4,180,501	4,033,922
Chr5 / scaff5	3,484,688	4,064,734
Chr6 / scaff6	3,28,934	3,411,043
Chr7 / scaff7	3,251,708	3,361,023
Chr8 / scaff8	2,995,396	3,353,035
Un / scaffUn	2,113,671	1,241,277

Supplemental Table S6. Primers used in this study.

Primer ID	Primer sequence	Target
J2-2-F	TCTCATGTCTGTCTTTACCGATT	JR2 re-arrangement SR2
J2-2-R	TATGAGGTTTAGGGTTACCGTTT	JR2 re-arrangement SR2
J3-F	TTTTGGTCGTGGTTGCAATA	JR2 re-arrangement SR3
J3-R	CGGCAAATCAGAAGAACCTC	JR2 re-arrangement SR3
J4-F	CCACACAAAGCATCACAACC	JR2 re-arrangement SR4
J4-R	CCACACATACAGACCGCATC	JR2 re-arrangement SR4
P1-F	GGCAGAAGTCAGTACGAGGA	Positive control JR2 and Ls17 (around SR3)
P1-R	TGACATCAATCTCCAAAGCCT	Positive control JR2 and Ls17 (around SR3)
P2-F	GTCACGCACTATACGGACCT	Positive control JR2 and Ls17 (around SR4)
P2-R	CCTACTTCAAGTTTATGCGCGTC	Positive control JR2 and Ls17 (around SR4)
L3-F	AATTCTGTACCTCTGTCCGT	VdLs.17 re-arrangement SR3
L3-R	CGCTACTTGATACTGTGAAAGG	VdLs.17 re-arrangement SR3
L4-F	GGAGAAAGATACGGAGAAATGG	VdLs.17 re-arrangement SR4
L4-R	GAGATTGAGATTGCGATGGGA	VdLs.17 re-arrangement SR4
KO-VDAG_05180-F1	GGACTTAAUAGTTTTGCCTGACAGTAGGT	Knock-out construct VDAG_05180
KO-VDAG_05180-R1	GGGTTTAAUAATTGATAGTGAACGGCTTC	Knock-out construct VDAG_05180
KO-VDAG_05180-F2	GGTCTTAAUGGTTTTCTTACGCCAGTATC	Knock-out construct VDAG_05180
KO-VDAG_05180-R2	GGCATTAAUTTGTCTGACATGTTTCTCGT	Knock-out construct VDAG_05180
SlRub-F1	GAACAGTTTCTCACTGTTGAC	Tomato *RuBisCo*
SlRub-R1	CGTGAGAACCATAAGTCACC	Tomato *RuBisCo*
VdGAPDH-F	CGAGTCCACTGGTGTCTTCA	*V. dahliae GAPDH*
VdGAPDH-R	CCCTCAACGATGGTGAACTT	*V. dahliae GAPDH*
qVDAG_05180-F1	CCGAAGGACATGCAGTCATACCGG	*V. dahliae VDAG_05180* (qRT-PCR)
qVDAG_05180-R1	TGCTGATATGGTTCCATTCCGTGAGG	*V. dahliae VDAG_05180* (qRT-PCR)

Supplemental Table S7. Location of breakpoints in VdLs.17 and JR2 compared to each other.

VdLs.17				JR2			
Chromosome	**Start**	**Stop**	**Note[a]**	**Chromosome**	**Start**	**Stop**	**Note[a]**
Chr1	1704928	1706838	IA	scaff1	464643	594004	IA
Chr1	2328810	2463411	IA	scaff1	2753856	2758331	IE
Chr1	2697567	2699850	IE	scaff1	6995538	7139184	IA
Chr1	4872136	4878089	IE	scaff1	7897589	7897588	IE
Chr1	5228212	5376603	IE	scaff2	2866260	3034775	IE
Chr2	1044841	1050896	IE[d]	scaff2	3115342	3116872	IE
Chr2	5151980	5262461	IE	scaff2	3192107	3197557	IE
Chr2	5280848	5280848	IE[e, f]	scaff2	3219697	3319456	IE
Chr3	723015	723015	IE[e, f]	scaff2	3663360	3663360	IE
Chr3	2001273	2150654	IE[c]	scaff2	4207425	4219600	IE
Chr3	3028284	3224254	IE	scaff3	1040257	1063663	IE
Chr3	3288396	3291734	IE	scaff3	1389038	1412796	IA
Chr3	3299414	3299973	IA	scaff3	1618226	1793692	IA
Chr3	3333364	3353845	IA	scaff3	2399925	2455297	IA
Chr3	3362443	3364549	IE	scaff4	931820	1275376	IE[b]
Chr3	3401115	3401126	IE	scaff4	1310827	1447892	IE[b]
Chr3	3500684	3566578	IE	scaff4	2288393	2307859	IA
Chr3	5314358	5337999	IA	scaff4	2853122	3027752	IA
Chr4	581875	582875	IE[b]	scaff5	527587	527588	IA[b]
Chr4	1149458	1159458	IE[b]	scaff5	935464	1028503	IA
Chr4	1797634	1806462	IA	scaff5	1485866	1533273	IA
Chr4	2088923	2098923	IA[b]	scaff5	3292265	3340655	IE
Chr4	2515463	2516463	IA[b]	scaff5	3527359	3647028	IE
Chr4	3611686	3624632	IE	scaff6	404908	434785	IA
Chr4	3841997	3844002	IE	scaff6	2153111	2308961	IA
Chr5	1030032	1138681	IA	scaff6	3184558	3191977	IE
Chr5	1685298	1706203	IA	scaff7	1345514	1430174	IA
Chr6	1789976	1803976	IA[c]	scaff7	2367775	2368961	IA
Chr7	423482	441582	IE	scaff8	1783883	1926229	IA
Chr8	1020271	1020372	IA[c]				
Chr8	1966542	1979754	IA				

a) Inter-chromosomal (IE) and intra-chromosomal (IA) re-arrangements

b) Variable region-associated breakpoint, large region delimited

c) Gap-associated translocation where breakpoint location is assessed and determined by manual inspection of the flanking sequence and the sequence in the unpositioned scaffolds.

d) Inter-chromosomal re-arrangement SR2 verified by PCR (Figure S1)

e) Inter-chromosomal re-arrangement SR3 verified by PCR (Figure S1)

f) Inter-chromosomal re-arrangement SR4 verified by PCR (Figure S1)

Supplemental Table S8. Classes of repetitive elements in *V. dahliae* genomes VdLs.17 and JR2.

	VdLs.17	JR2
Sequence summary:		
Number of chromosomes	9	9
total length	36,874,636	37,006,893
total length excluding Ns	32,903,115	33,040,648
GC level (%)	55.85	55.52
bases masked	1,331,148 (4.05%)	1,429,941 (4.33%)
Repetitive Elements (number of elements, length occupied, percentage of total genome sequence):		
SINEs	0; 0 bp; 0%	0; 0 bp; 0%
LINEs	71; 116,971 bp; 0.36%	63; 61,416 bp; 0.19%
LTR elements	395; 586,279 bp; 1.78%	600; 671,062 bp; 2.03 %
DNA elements	93; 88,379 bp; 0.27%	70; 61,486 bp; 0.19%
Unclassified	1,030; 311,043 bp; 0.95%	1,311; 408,003 bp; 1.23%
Total interspersed repeats	1,589; 1,102,672 bp; 3.35%	2,044; 1,201,967 bp; 3.64%
Small RNA	0; 0 bp; 0%	0; 0 bp; 0%
Satellites	0; 0 bp; 0%	0; 0 bp; 0%
Simple repeats	3,240; 154,570 bp; 0.47%	3,206; 151,249 bp; 0.46%
Low complexity	1,483; 76,153 bp; 0.23%	1,598; 79,866 bp; 0.24%

Supplementary References

1. Wu TD, Nacu S, **Fast and SNP-tolerant detection of complex variants and splicing in short reads.** *Bioinformatics*, 2010, **26**:873-881.
2. Danecek P, *et al*., **The variant call format and VCFtools.** *Bioinformatics*, 2011, **27**:2156-2158.
3. Klosterman SJ, *et al*., **Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens.** *PLoS Pathog*, 2011, **7**:e1002137.
4. Quinlan AR, Hall IM, **BEDTools: a flexible suite of utilities for comparing genomics features.** *Bioinformatics*, 2010, **26**:841-842.
5. Tamura K, *et al*., **MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Mol Biol Evol*, 2011, **28**:2731-2739.
6. de Jonge R, *et al.*, **Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing.** *Proc Natl Acad Sci USA*, 2012, **109**:5110-5115.
7. Zerbino DR, Birney E, **Velvet: algorithms for de novo short read assembly using de Bruijn graphs.** *Genome Res*, 2008, **18**:821-829.
8. Kurtz S, *et al*., **Versatile and open software for comparing large genomes.** *Genome Biol*, 2004, **5**:R12.
9. Krzywinski M. *et al*., **Circos: an information aesthetic for comparative genomics.** *Genome Res*, 2009, **19**:1639-1645.
10. Stein SD, *et al.,* **The generic genome browser: a building block for a model organism system database.** *Genome Res*, 2002, **12**:1599-1610.
11. Price AL, Jones NC, Pevzmer PA, ***De novo*** **identification of repeat families in large genomes.** *Bioinformatics*, 2005, **21**:i351-358.
12. Bao Z, Eddy SR, **Automated de novo identification of repeat sequence families in sequenced genomes.** *Genome Res*, 2002, **12**:1269-1276.
13. Xu Z, Wang H, **LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons.** *Nucl Acid Res*, 2007, **35**:W265-268.
14. Mehrabi R, *et al*., **Karyotyping methods for fungi.** *Methods Mol Biol*, 2012, **835**:591-602.
15. Stanke M, Diekhans M, Baertsch R, Haussler D, **Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding.** *Bioinformatics*, 2008, **24**:637-644.
16. Trapnell C, Pachter L, Salzberg SL, **TopHat: discovering splice junctions with RNA-Seq.** *Bioinformatics*, 2009, **25**:1105-1111.
17. Trapnell C, *et al*., **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.** *Nat Biotech*, 2010, **28**:511-515.
18. Stanke M, Tzvetkova A, Morgenstern B, **AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome.** *BMC Genome Biol*, 2006, **7**:S11.
19. de Jonge R, ***In silico*** **identification and characterization of effector catalogs.** *Methods Mol Biol*, 2012, **835**:415-425.
20. Finn RD, *et al*., **The Pfam protein families database.** *Nucl Acid Res*, 2008, **36**:D281-288.
21. Petersen TN, Brunak S, von Heijne G, Nielsen H, **SignalP 4.0: discriminating signal peptides from transmembrane regions.** *Nat Meth*, 2011, **8**:785-786.
22. Horton P, *et al*., **WoLF PSORT: protein localization predictor.** *Nucl Acid Res*, 2007, **35**:W585-587.
23. Raffaele S, *et al.,* **Genome evolution following host jumps in the Irish potato famine pathogen lineage.** *Science*, 2010, **330**:1540-1543.
24. Frandsen RJ, Andersson JA, Kristensen MB, Giese H, **Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi.** *BMC Mol Biol*, 2008, **9**:70.

25. Fradin EF, *et al.*, **Genetic dissection of *Verticillium* wilt resistance mediated by tomato Ve1.** *Plant Physiol*, 2009, **150**:320-332.

26. Fulton TM, Chunwongse J, Tanksley SD, **Microprep protocol for extraction of DNA from tomato and other herbaceous plants.** *Plant Mol Biol Rep*, 1995, **13**:207-209.

Chapter 6

General discussion: Plant pathogen effectors revealed by next-generation genomics

de Jonge R

Abstract

Filamentous eukaryotic pathogens are causal agents of disease in animals and plants. Research on plant-pathogen interactions mostly focuses on identification and characterization of novel effectors; secreted molecules that modulate host physiology to enable host colonization. Discovery of new effectors has been greatly boosted by the *in silico* identification from genome sequences. Recent advancements in sequencing technologies facilitate genome sequencing of many isolates of a single species, facilitating evolutionary analyses of pathogen populations. Such analyses can quickly provide insight in the diversity and evolution of candidate effector genes and allows for the identification of effectors that are likely to be most critical for host colonization.

Introduction

The effectors of filamentous pathogens

Eukaryotic plant pathogens secrete an arsenal of effector proteins that modulate host physiology and enable successful host colonization. Accordingly, research on plant-pathogen interactions has mainly focused on discovery and functional analysis of effectors. Effector molecules can be divided into two classes; those that remain extracellular and those that are translocated into host cells [1]. Extracellular effectors are common to most pathogens, whereas host-translocated or cytoplasmic effectors have been identified only for a subset of plant pathogens. Especially pathogens that establish an intimate interaction with the host through the formation of specialized infection structures, such as the haustorium of oomycetes [2], rusts and powdery mildews [3,4], are known to produce cytoplasmic effectors. Typically, effector molecules belonging to the class of small secreted proteins are lineage-specific, although some have homologs in other pathogens. Little is known about the function of these effector molecules, as only for a few their role in virulence has been elucidated. These include a number of extracellular effectors from *Cladosporium fulvum* and *Phytophthora infestans* that inhibit host proteases [5,6], effectors from *C. fulvum*, *Zymoseptoria tritici* (synonym of *Mycosphaerella graminicola*), *M. fijiensis* and *Magnaporthe oryzae* that protect fungi against chitin-triggered host defense responses [7-12], and the *Ustilago maydis* effector Pep1 that inhibits a host peroxidase to block the oxidative burst [13,14]. Cytoplasmic effectors are thought to play a crucial role in suppression of host defense responses. Such a role has been found for *P. infestans* Avr3a that stabilizes a host ubiquitin ligase to suppress host cell death [15,16], and *Fusarium oxysporum* Avr1 that suppresses host immunity triggered upon recognition of Avr2 and Avr3 by the immune receptors I2 and I3, respectively [17].

Effector identification

Many effector proteins have been cloned based on their recognition by plant immune receptors, and consequently are also known as avirulence factors. Although demonstrated only in particular cases, avirulence factors are generally assumed to contribute to virulence as proposed in the guard model [17]. Effector proteins were also identified by random insertional mutagenesis [19-23], while more recently,

transcriptomics [24-26] and proteomics [27-30] approaches have been used to identify candidate effectors, often in combination with bioinformatics-based prediction of small secreted proteins [24,31-33]. However, subsequent functional analysis to confirm a role in pathogen virulence for such effectors is often hampered by the vast number of candidates that is obtained, and thus further selection to identify the most relevant effectors is necessary. Also, the inability or inefficiency of genetic manipulation in particular fungi as well as functional redundancy can complicate the validation of candidate effectors. However, alternative methods that can be used to demonstrate the contribution of a particular effector to virulence have been reported, including heterologous expression of candidate effectors in plants [5,8] and bacteria [34,35].

Genomic location of effectors

Frequently, effector genes are found in repeat-rich genomic regions that evolve more rapidly than other regions of the genome [36], such as the AT-rich isochore-like regions in the blackleg fungus *Leptosphaeria maculans* that are mainly composed of transposons and effector genes [37,38]. Similarly, effector genes in *M. oryzae* are found in transposon-rich, sub-telomeric regions [39].

Identification of novel effector candidates by comparative genomics

In comparative genomics, genome sequences from different individuals are compared through alignment in order to study the relationship between sequence and function. Over 150 whole genome sequences of eukaryotic plant pathogens have been determined over the last few years, and many comparative genomics studies have been performed to reveal the similarities and differences between selected genomes and subsequently relate these to phenotypic differences [36]. In the following section, studies that employed comparative genomics to identify candidate effectors in various plant pathogens are described, including vascular wilt fungi, smut fungi and oomycetes of the *Phytophthora* genus.

Lineage-specific regions of vascular fungi revealed by comparative genomics

F. oxysporum is a species complex, composed of pathogenic lineages, that cause disease in only a narrow range of plant species, and non-pathogenic lineages. Comparative analysis of the genome sequence of a *F. oxysporum* strain causing tomato vascular wilt with those of phylogenetically related *Fusarium* pathogens on cereals (*F. graminearum* and *F. verticillioides*) and pea (*F. solani*), resulted in the identification of conserved, core genomic regions that are shared between all species, and a number of non-conserved, lineage-specific (LS) genomic regions [40]. These LS regions include four complete chromosomes that account for ~40% of the *F. oxysporum* genome and that are enriched for transposable elements and (putative) effector genes. The effector genes *Six1* and *Six3* [41,42], and various other candidate effector genes [28] that are all conserved in *F. oxysporum* strains causing tomato wilt [43], are all located on LS chromosome 14. Consequently, it was proposed that LS chromosome 14 carries the main determinants for pathogenicity of *F. oxysporum* on tomato [40]. This hypothesis was corroborated by comparative analysis with the genome sequence of a second *F. oxysporum* strain, pathogenic on Arabidopsis, and with EST sequences of a cotton infecting

strain, as for both only homology to the core genome was observed [40]. Based on phylogenetic analysis of LS-encoded proteins it was further proposed that the LS regions were acquired by horizontal gene transfer. Experimentally, horizontal transfer of complete LS chromosomes from a tomato pathogenic isolate to a non-pathogenic isolate was shown by co-cultivation of these two isolates. Furthermore, these experiments showed that pathogenicity of *F. oxysporum* towards tomato can be specifically attributed to the acquisition of LS chromosome 14 [40].

Comparative genome analysis of the vascular wilt fungi *Verticillium dahliae* and *V. albo-atrum* [44] revealed four LS regions in *V. dahliae* that were missing in *V. albo-atrum*. Unlike in *F. oxysporum*, LS regions did not cover complete chromosomes; they rather represented islands in largely syntenic chromosomes. LS regions in *V. dahliae* were also enriched for transposable elements, and flexibility of these regions was reflected by extensive gene duplications. Nucleic acid hybridizations using probes from four different genes (one from each of the four LS regions) revealed substantial genetic variation among tested *V. dahliae* strains [44]. Unlike *F. oxysporum* LS-regions, no enrichment for candidate effector genes was observed for *V. dahliae* LS regions, although they were devoid of housekeeping genes and enriched for genes related to iron/lipid metabolism that are known to play a role in host-pathogen interactions [45].

Islands of low similarity encode effectors required for virulence of smut fungi

Smut fungi are biotrophic pathogens that primarily infect grasses (*Graminaceae*), including economically important cereal crops such as maize, corn, barley and wheat. Despite overall high sequence similarity, gene-by-gene comparisons of the genome sequences of the closely related maize smuts *U. maydis* and *Sporisorium reilianum* revealed 43 distinct genomic islands containing genes with low sequence similarity [46]. Most of the genes (71%) in these divergent islands were found in both species, whereas 10% were *U. maydis*-specific and 19% were *S. reilianum*-specific. Moreover, islands were significantly enriched for secreted proteins, suggesting involvement in pathogenicity. Functional analysis confirmed a role in pathogenicity for three of these regions, as loss of these regions resulted in reduced virulence on maize seedlings [46]. Notably, increased divergence of smut effectors as compared to the rest of the proteome was also found by comparisons between *U. maydis* and *U. hordei* proteins, a smut of barley [47].

Gene-sparse regions of phytophthora infestans are enriched for candidate effectors

Re-sequencing of *P. infestans*, causal agent of late blight of potato and tomato, and comparison to the genomes of the three sister species *P. ipomoeae*, *P. mirabilis* and *P. phaseoli* that are pathogenic on morning glory, four-o'clock and lima bean, respectively, revealed considerable variation in evolutionary rates across its genome [48]. Whereas genes in most genomic regions were highly conserved, genes in repeat-rich, gene-sparse regions were highly divergent as evidenced by high levels of copy number variation, presence/absence polymorphisms and non-synonymous substitutions. These gene-sparse regions were highly enriched for genes induced *in planta,* particularly effector genes.

Identification of effectors by population genomics

Population genetics of plant pathogens focuses on the analysis of allele frequency and distribution of effector genes in a population. These analyses can provide insight in the selective constrains that act on effector genes in diverse ecological settings, such as in agriculture by the deployment of resistance genes. Many population genetic studies have found that effector genes are under positive selection, due to their role in the plant-pathogen interactions [49]. Population genomics uses the power of next-generation sequencing to determine the genome sequences of multiple isolates from a single pathogenic species and simultaneously assesses the evolution of many genes, including effector genes and marks a next step in comparative genomics to identify plant pathogen effectors. In the following section, recent studies are described that applied population genomics to identify effector genes by assessing natural variation within a pathogen population.

Identification of effectors by genome-wide association analyses

Genome-wide association studies typically try to link single nucleotide polymorphisms (SNPs) and copy number variation with phenotypic traits. Few studies on microbial pathogens have applied association genetics to identify novel effectors. In one such study, SNPs in candidate *M. oryzae* effectors were examined that could be associated with avirulence activities on a panel of rice cultivars harboring different resistance genes, resulting in the identification of three novel avirulence genes [50]. Although their contribution to pathogen virulence has not yet been established, they likely encode effectors. In a recent study, we applied genome-wide association analyses to identify the *V. dahliae* race-specific effector that activates Ve1-mediated resistance in tomato [Chapter 4]. To this end, we sequenced race 1 and 2 isolates using next-generation sequencing and compared their genome sequences, resulting in the identification of one region that was found exclusively in race 1 isolates. Deep transcriptome sequencing of race 1 *V. dahliae*-infected plants subsequently resulted in the identification of the highly expressed *Ave1* gene within this region. Functional analysis confirmed that Ave1 activates Ve1-mediated resistance and demonstrated that Ave1 is required for full virulence on susceptible tomato plants [Chapter 4].

Population genome sequencing reveals effectors under adaptive evolution

Population genome sequencing of *Z. tritici*, causal agent of Septoria tritici blotch of wheat, and comparative analysis to closely related sister species that are pathogenic on wild grasses was used to study host adaptation and speciation [51-53]. Through comparative analyses, the evolutionary rates across the *Z. tritici* genome were assessed by determination of the number of nucleotide substitutions [52,53]. By assessing the ratio of non-synonymous and synonymous substitutions across protein-coding genes, various effector genes under positive selection were identified [52,53].

Recently we applied a similar approach to identify candidate effectors in the genome of *V. dahliae* [Chapter 5]. Comparative analyses of 10 *V. dahliae* isolates with the reference genome VdLs.17 [44] revealed high sequence similarity across the complete genome, including protein-coding regions, and led to the

identification of four candidate effectors with elevated levels of non-synonymous mutations. Preliminary analysis suggests that at least one of them contributes to virulence of *V. dahliae* on tomato (unpublished data).

Frequent gain and loss of effector genes of plant pathogens revealed by population genomics

Comparative analyses of 10 *V. dahliae* isolates revealed numerous transposon-rich LS regions in each isolate, but analysis of the composition of these LS regions did not reveal enrichment of candidate effectors [Chapter 5]. However, by analyses of *in planta* gene expression we found a clear overrepresentation of highly induced genes within the LS regions, including the previously identified race 1 effector *Ave1* and additional putative effectors. Accordingly, these proteins are unique or shared by only a subset of strains (Figure 1), and we speculate that frequent presence/absence polymorphisms of effector genes in LS regions is mediated by the flexibility and instability of these regions. Preliminary analysis suggests that also at least one of these putative effectors contributes to virulence of *V. dahliae* on tomato (unpublished data). Flexibility of LS regions is reflected by enrichment in transposable elements, extensive gene duplications and association with chromosomal breakpoints [Chapter 5, 44]. Presence/absence polymorphisms and translocation of effector genes that are associated with unstable genomic regions have also been observed in asexual lineages of *M. oryzae* [50,54], and it was hypothesized that parasexual recombination facilitates the exchange of effector genes between these lineages [54,55]. Parasexual recombination involves fusion of haploid fungal hyphae followed by karyogamy to form a diploid nucleus. By mitotic recombination and haploidization these nuclei

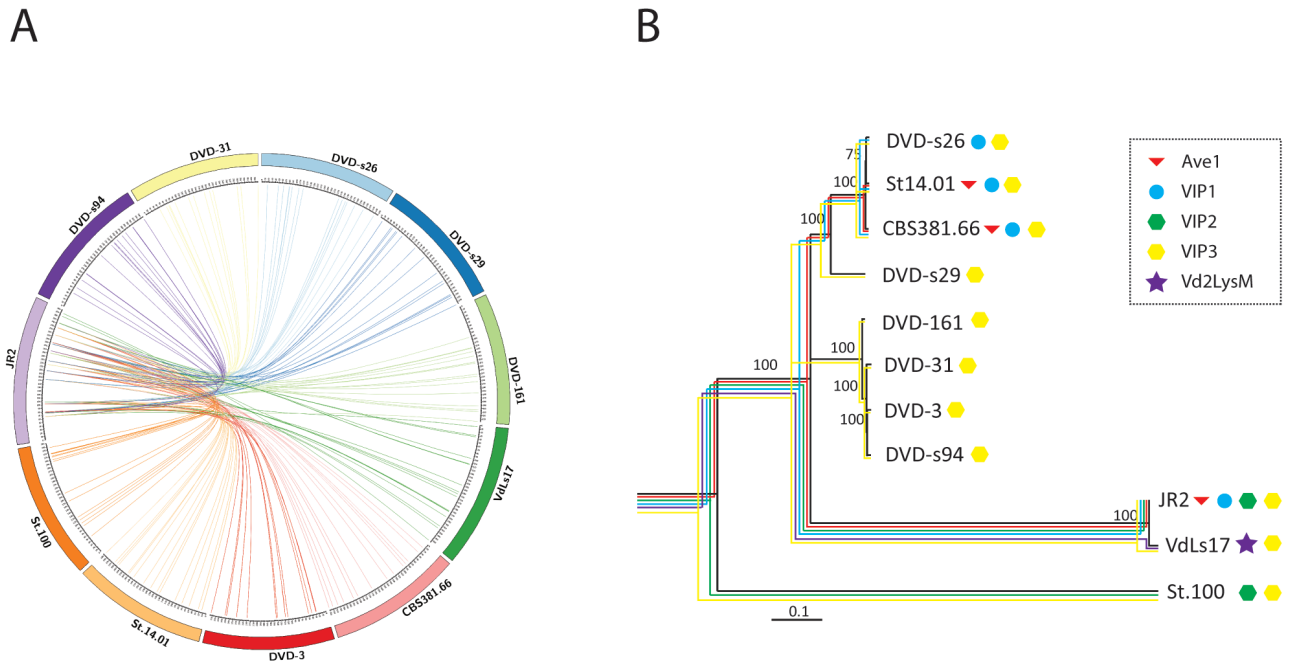


Figure 1. *Verticillium dahliae* effector dynamics. A) Circular representation of the genome sequences of a population of 10 *V. dahliae* isolates. Each isolate is drawn around the outer edge of the circle. Lines connect *V. dahliae* JR2 effector genes with homologues in other isolates. B) Strain-specific presence of effector genes in the phylogenetic tree of the *V. dahliae* population.

return to their normal haploid state, resulting in non-sexual exchange of genetic material in the absence of meiosis. Parasexuality has been identified in many, mostly asexually propagating fungi during *in vitro* culturing, and it was shown to significantly increase *in vitro* to adaptive capabilities in the model fungus *Aspergillus nidulans* [56]. Parasexual recombination could also be involved in the exchange of effector genes between asexual lineages of *V. dahliae*. Laboratory studies have shown that transgenic *V. dahliae* strains that carry auxotrophic markers are capable of hyphal fusion and subsequent exchange of genetic material, depending on their vegetative compatibility [44]. Likewise, transfer of chromosomes between asexual lineages of *F. oxysporum* [40] could be mediated by a parasexual cycle [57]. However, the significance of parasexuality in nature and the genetic factors that control parasexual compatibility between lineages are poorly understood [58]. Nevertheless, evidence for parasexual recombination in field populations of *M. oryzae* has been reported [59].

Frequent gain and loss of effector genes is probably driven by local selective constrains such as the deployment of resistance genes. Our phylogenetic analysis based on whole-genome sequence data, has demonstrated that race 1 and race 2 isolates of *V. dahliae* form separate clades, and the fact that we did not observe any sequence variation for *Ave1* within and between *Verticillium* species strongly suggests that *Ave1* has been lost several times [Chapter 4, 5]. Deletion of effector genes is common in plant pathogens [36,49,60], which demonstrates the importance of population genomics to identify the complete effector repertoire of a species.

Perspective of next-generation genomics for detection of effector variation and identification of novel resistance traits

Identification of conserved effector targets for development of durable resistance by population genomics

In agriculture, resistance genes are commonly deployed in crop plants to restrict disease development. However, resistance is often broken because pathogen populations evolve to overcome resistance. Insight in the frequency and distribution of effectors can be useful to predict targets for durable resistance if one assumes that conserved effectors are likely important for pathogen fitness and thus less likely to be lost. Such a strategy was recently used to identify effectors of the bacterial pathogen *Xanthomonas axonopodis* pv. *manihotis,* causal agent of cassava bacterial blight, that are conserved [61]. Whole genome sequencing of 65 isolates and comprehensive assessment of known effector genes resulted in the identification of nine conserved effectors that are found in all isolates and that can be used to screen for potential durable resistance traits directed against *Xanthomonas* [61]. The declining cost of sequencing makes population genomics an attractive tool to identify targets for durable resistance in plants.

Next-generation genomics reveals the origin of emerging microbial diseases and can be used to find novel resistance traits

In 1999 the highly aggressive race Ug99 of *Puccinia graminis*, the causal agent of stem rust disease on wheat, emerged in Uganda. Since then, Ug99 has spread throughout eastern Africa and recently reached the Middle-East. Over 90% of the wheat cultivars currently grown worldwide are susceptible to this race, and Ug99 was recognized as a major treat to wheat production and food security [62,63]. Next-generation genomics can quickly provide insight into the evolution of such an emerging pathogen. Genome sequencing of various Ug99 strains has been performed and comparative analyses with the reference strain [64] revealed that Ug99 and the reference strain share only 70-80% of their genomes [65]. Instead of looking for conserved, core effectors that could be possible targets for durable resistance, in this case research focuses on the identification of polymorphic regions which can be exploited for diagnostics as well as for the identification of novel effector genes that contribute to the high aggressiveness of the Ug99 race. Candidate effectors can subsequently be used to screen for recognition in wheat germplasm, accelerating the process of finding novel resistance traits.

Concluding remarks

Comparative genomics has revealed considerable variation in the genome sequences of plant pathogens. Intriguingly, effector genes are often located in variable regions, allowing for their rapid evolution. Applying comparative genomics on pathogen populations further corroborates these observations, and can result in the identification of effectors that are under accelerated evolution, illustrated by copy number variation and positive selection. Population genomic sequencing can furthermore be applied to detect polymorphisms in effector genes across the population in order to select conserved effectors that could be potential targets for development of durable resistance. The high speed and low costs of these analyses make comparative population genomics a valuable tool to study emerging diseases, and can be used to get a better understanding of the evolution of plant pathogens in natural and agricultural settings.

References

1. de Jonge R, Bolton MD, Thomma BHPJ, **How filamentous pathogens co-opt plants: the ins and outs of fungal effectors.** *Curr Opin Plant Biol*, 2011, **14**:400-406.
2. Kamoun S, **A catalogue of the effector secretome of plant pathogenic oomycetes.** *Annu Rev Phytopathol*, 2006, **44**:41-60.
3. Voegele RT, Mendgen K, **Rust haustorial: nutrient uptake and beyond.** *New Phytol*, 2003, **159**:93-100.
4. Panstruga R, Dodds PN, **Terrific protein traffic: the mystery of effector protein delivery by filamentous plant pathogens.** *Science*, 2009, **324**:748-750.
5. van Esse HP, *et al*., **The *Cladosporium fulvum* virulence protein Avr2 inhibits host proteases required for basal defense.** *Plant Cell*, 2008, **20**:1948-1963.
6. Song J, *et al*., **Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3.** *Proc Natl Acad Sci USA*, 2009, **106**:1654-1659.
7. van den Burg HA, *et al*., ***Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection.** *Mol Plant Microbe Interact*, 2006, **19**:1420-1430.

8. van Esse HP, *et al*., **The chitin-binding *Cladosporium fulvum* effector protein Avr4 is a virulence factor.** *Mol Plant Microbe Interact*, 2007, **20**:1092-1101.
9. de Jonge R, *et al*., **Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants.** *Science*, 2010, **329**:953-955.
10. Stergiopoulos I, *et al*., **Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots.** *Proc Natl Acad Sci USA*, 2010, **107**:7610-7615.
11. Marshall R, *et al.*, **Analysis of two in planta expressed LysM effector homologs from the fungus *Mycosphaerella graminicola* reveals novel functional properties and varying contributions to virulence on wheat.** *Plant Physiol*, 2011, **156**:756-769.
12. Mentlak TA, *et al*., **Effector-mediated suppression of chitin-triggered immunity by *Magnaporthe oryzae* is necessary for rice blast disease.** *Plant Cell*, 2012, **24**:322-335.
13. Doehlemann G, *et al*., **Pep1, a secreted effector protein *Ustilago maydis*, is required for successful invasion of plant cells.** *PLoS Pathog*, 2009, **5**:e1000290.
14. Hemetsberger C, *et al*., **The *Ustilago maydis* effector Pep1 suppresses plant immunity by inhibition of host peroxidase activity.** *PLoS Pathog*, 2012, **8**:e1002684.
15. Bos JL, *et al*., **The C-terminal half of *Phytophthora infestans* RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in *Nicotiana benthamiana*.** *Plant J*, 2006, **48**:165-176.
16. Bos JL, *et al*., ***Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1.** *Proc Natl Acad Sci USA*, 2010, **107**:9909-9914.
17. Houterman PM, Cornelissen BJC, Rep M, **Suppression of plant resistance gene-based immunity by a fungal effector.** *PLoS Pathog*, 2008, **4**:e1000061.
18. van der Biezen EA, Jones JDG, **Plant disease-resistance proteins and the gene-for-gene concept.** *Trends Biochem Sci*, 1998, **23**:454-
19. Betts MF, *et al*., **Development of a high throughput transformation system for insertional mutagenesis in *Magnaporthe oryzae*.** *Fungal Genet Biol*, 2007, **44**:1035-1049.
20. Blaise F, *et al*., **A critical assessment of *Agrobacterium tumefaciens*-mediated transformation as a tool for pathogenicity gene discovery in the phytopathogenic fungus *Leptosphaeria maculans*.** *Fungal Genet Biol*, 2007, **44**:123-138.
21. Jeon J, *et al*., **Genome-wide functional analysis of pathogenicity genes in the rice blast fungus.** *Nat Genet*, 2007, **39**:561-565.
22. Seong K, *et al*., **Random insertional mutagenesis identifies genes associated with virulence in the wheat scab fungus *Fusarium graminearum*.** *Phytopathol*, 2005, **95**:744-750.
23. Michielse CB, *et al*., **Insight into the molecular requirements for pathogenicity of *Fusarium oxysporum* f. sp. *lycopersici* through large-scale insertional mutagenesis.** *Genome Biol*, 2009, **10**:R4.
24. Saitoh H, *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.
25. Stassen JH, *et al*., **Effector identification in the lettuce downy mildew *Bremia lactucae* by massively parallel transcriptome sequencing.** *Mol Plant Pathol*, 2012, doi: 10.1111/j.1364-3703.2011.00780.x.
26. Kleemann J, *et al*., **Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen *Colletotrichum higginsianum*.** *PLoS Pathog*, 2012, **8**:e1002643.
27. Kim ST, *et al*., **Proteomic analysis of pathogen-responsive proteins from rice leaves induced by rice blast fungus, *Magnaporthe grisea*.** *Proteomics*, 2004, **4**:3569-3578.
28. Houterman PM, *et al*., **The mixed proteome of *Fusarium oxysporum*-infected tomato xylem vessels.** *Mol Plant Pathol*, 2007, **8**:215-221.

29. Bindschedler LV, *et al*., **In planta proteomics and proteogenomics of the biotrophic barley fungal pathogen *Blumeria graminis* f. sp. *hordei*.** *Mol Cell Prot*, 2009, **8**:2368-2381.
30. Vincent D, *et al*., **Proteomic techniques for plant-fungal interactions.** *Methods Mol Biol*, 2012, **835**:75-96.
31. Mueller O, *et al*., **The secretome of the maize pathogen *Ustilago maydis*.** *Fungal Genet Biol*, 2008, **45**:S63-70.
32. de Jonge R, **In silico identification and characterization of effector catalogs.** *Methods Mol Biol*, 2012, **835**:415-425.
33. Hacquard S, *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**:279-293.
34. Sohn KH, Lei R, Nemri A, Jones JD, **The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in *Arabidopsis thaliana*.** *Plant Cell*, 2007, **19**:4077-4090.
35. Rentel MC, Leonelli L, Dahlbeck D, Zhao B, Staskawicz BJ, **Recognition of the *Hyaloperonospora parasitica* effector ATR13 triggers resistance against oomycete, bacterial, and viral pathogens.** *Proc Natl Acad Sci USA*, 2008, **105**:1091-1096.
36. Raffaele S, Kamoun S, **Genome evolution in filamentous plant pathogens: why bigger can be better.** *Nat Rev Microbiol*, 2012, **10**:417-430.
37. van de Wouw AP, *et al*., **Evolution of linked avirulence effectors in *Leptosphaeria maculans* is affected by genomic environment and exposure to resistance genes in host plants.** *PLoS Pathog*, 2010, **6**:e1001180.
38. Rouxel T, *et al*., **Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations.** *Nat Commun*, 2011, **2**:202.
39. Orbach MJ, *et al*., **A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*.** *Plant Cell*, 2000, **12**:2019-2032.
40. Ma L-J, *et al.,* **Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*.** *Nature*, 2010, **464**:367-373.
41. Rep M, *et al.*, ***Fusarium oxysporum* evades I-3-mediated resistance without altering the matching avirulence gene.** *Mol Plant Microbe Interact*, 2005, **18**:15-23.
42. Houterman PM, *et al*., **The small effector protein Avr2 secreted in xylem by a vascular wilt fungus interacts with its cognate resistance protein inside plant cells.** *Plant J*, 2009, **58**:970-978.
43. van der Does HC, *et al*., **Expression of effector gene *SIX1* of *Fusarium oxysporum* requires living plant cells.** *Fungal Genet Biol*, 2008, **45**:1257-1264.
44. Klosterman SJ, *et al*., **Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens.** *PLoS Pathog*, 2011, **7**:e1002137.
45. Expert D, Enard C, Masclaux C, **The role of iron in plant host-pathogen interactions.** *Trends Microbiol*, 1996, **4**:232-237.
46. Schirawski J, *et al.*, **Pathogenicity determinants in smut fungi revealed by genome comparison.** *Science*, 2010, **330**:1546-1548.
47. Laurie JD, *et al*., **Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements.** *Plant Cell*, 2012, **24**:1733-1745.
48. Raffaele S, *et al.,* **Genome evolution following host jumps in the Irish potato famine pathogen lineage.** *Science*, 2010, **330**:1540-1543.
49. Stukenbrock EH, McDonald BA, **Population genetics of fungal and oomycete effectors involved in gene-for-gene interactions.** *Mol Plant Microbe Interact*, 2009, **22**:371-380.
50. Yoshida K, *et al*., **Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen *Magnaporthe oryzae*.** *Plant Cell*, 2009, **21**:1573-1591.
51. Goodwin SB, *et al*., **Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity and stealth pathogenesis.** *PLoS Genet*, 2011, **7**:e10020170.
52. Stukenbrock EH, *et al*., **Whole-genome and chromosome evolution associated with host adaptation and speciation of**

the wheat pathogen *Mycosphaerella graminicola*. *PLoS Genet*, 2010, **6**:e1001189.

53. Stukenbrock EH, *et al*., **The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species.** *Genome Res*, 2011, **21**:2157-2166.
54. Chuma I, *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.
55. Noguchi MT, Yasuda N, Fujita Y, **Evidence of genetic exchange by parasexual recombination and genetic analysis of pathogenicity and mating type of parasexual recombinants in rice blast fungus, Magnaporthe oryzae.** *Phytopathol*, 2006, **96**:746-750.
56. Schoustra SE, Debets AJ, Slakhorst M, Hoekstra RF, **Mitotic recombination accelerates adaptation in the fungus *Aspergillus nidulans*.** *PLoS Genet*, **3**:e68.
57. Correll JC, **The relationship between formae speciales, races, and vegetative compatibility groups in *Fusarium oxysporum*.** *Phytopathology*, 1991, **81**:1061-1064.
58. Glass NL, Jacobson DJ, Shiu PK **The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycete fungi.** *Annu Rev Genet*, 2000, **34**:165-186.
59. Zeigler RS, *et al*., **Evidence of parasexual exchange of DNA in the rice blast fungus challenges its exclusive clonality.** *Phytopathol*, 1997, **87**:284-294.
60. van den Ackerveken GF, van Kan JAL, de Wit PJGM, **Molecular analysis of the avirulence gene *avr9* of the fungal tomato pathogen *Cladosporium fulvum* fully supports the gene-for-gene hypothesis.** *Plant J*, 1992, **2**:359-366.
61. Bart R, *et al*., **High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance.** *Proc Natl Acad Sci USA*, 2012, **109**:E1972-1979.
62. Singh RP, *et al*., **The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production.** *Annu Rev Phytopathol*, 2011, **49**:465-481.
63. Fisher MC, *et al*., **Emerging fungal threats to animal, plant and ecosystem health.** *Nature*, 2012, **484**:186-194.
64. Duplessis S, *et al*., **Obligate biotrophy features unraveled by the genomic analysis of rust fungi.** *Proc Natl Acad Sci USA*, 2011, **108**:9166-9171.
65. Szabo LJ, Cuomo C, **Sequencing Ug99 and other stem rust races: progress and results.** *Proceedings, oral papers and posters,* 2009, *Technical Workshop, Borlaug Global Rust Initiative, March, 2009 [ed. by McIntosh, R.]. Ithaca, USA: Borlaug Global Rust Initiative,* 47-48, http://www.globalrust.org/db/attachments/resources/737/10/BGRI_2009_proceedings_cimmyt_isbn.pdf.

Summary

Research on effectors secreted by pathogens during host attack has dominated the field of molecular plant–microbe interactions in the last decade. Effectors are defined as molecules secreted by plant pathogens to modulate host physiology to enable host colonization. In contrast to most bacterial effectors that are delivered by the type III machinery inside the host cytoplasm, fungal and oomycete effectors are delivered extracellularly, and we are gradually learning more about their functions. While some function outside the host cell to compromise defence, others exploit host cellular uptake mechanisms to suppress defence or stimulate the release of nutrients. In **Chapter 1** we describe the function and evolution of effectors from filamentous plant pathogens, guided by the consecutive stages occurring during disease establishment.

In **Chapter 2**, the occurrence and characteristics of a family of effectors that we named LysM effectors are described, and we show that this family is conserved throughout the fungal kingdom. LysM effectors are secreted proteins that contain no other recognizable protein domains than Lysin motifs (LysMs) that have been recognized as carbohydrate-binding protein domains. We propose that LysM effectors have a role in sequestration of chitin oligosaccharides, breakdown products of fungal cell walls that are released during invasion and act as triggers of host immunity, to dampen host defence.

In **Chapter 3** we investigated the function of the LysM effector Ecp6 from the plant pathogenic fungus *Cladosporium fulvum*. We show that Ecp6 binds to chitin and prevents the induction of chitin-triggered host defence responses, such as the alkalinisation of tomato and tobacco cell suspensions and the production of reactive oxygen species in tomato and tobacco leaf disks upon chitin treatment. Consistent with a role as suppressor of chitin-triggered immunity, Ecp6 was found to successfully compete with the rice receptor for binding of chitin oligosaccharides. In conclusion, we show that Ecp6 mediates virulence through scavenging of chitin oligosaccharides. As LysM effectors are widely conserved in the fungal kingdom, this may represent a common strategy of host immune evasion by fungal pathogens.

In **Chapter 4** we describe the identification of the race 1 elicitor activating Ve1-mediated resistance in tomato. By high-throughput population genome sequencing of both race 1 and race 2 isolates, a single 50 Kb sequence stretch was identified that only occurs in race 1 strains. Subsequent transcriptome sequencing of *Verticillium*-infected *Nicotiana benthamiana* plants revealed only a single highly expressed ORF in this region, designated *Ave1* (for *Avirulence on Ve1* tomato). Functional analyses confirmed that Ave1 activates Ve1-mediated resistance and demonstrated that Ave1 markedly contributes to fungal virulence, not only on tomato but also on Arabidopsis. Interestingly, we found that Ave1 is homologous to a widespread family of plant natriuretic peptides that, beside plants, are also found in the plant pathogenic fungi *Colletotrichum higginsianum*, *Cercospora beticola* and *Fusarium oxysporum* f. sp. *lycopersici*, as well as in the bacterial plant pathogen *Xanthomonas axonopodis*. The distribution of Ave1 homologs, which coincides with the presence of *Ave1* within a flexible genomic region, strongly suggests that *Verticillium* acquired *Ave1* from plants through horizontal gene transfer. Remarkably, by transient expression we show that also the Ave1 homologs from *F. oxysporum* and *C. beticola* can activate *Ve1*-mediated resistance. In line with this observation, *Ve1*

was found to mediate resistance toward *F. oxysporum* in tomato, showing that this immune receptor is involved in resistance against multiple fungal pathogens.

In **Chapter 5**, a comparative genomics approach was used to study sequence diversity within a population of *V. dahliae* isolates that is known to reproduce asexually. We found that sequence diversity is generally low among *V. dahliae* isolates. However, comparative analyses by pairwise alignment between the two highly similar isolates VdLs.17 and JR2 (>99.9% identity) revealed regions of extensive synteny that are repeatedly interrupted by intra- and inter-chromosomal rearrangements. Syntenic breakpoints were associated with the presence of retrotransposons and frequently flanked by lineage-specific sequences. Syntenic breakpoints and lineage-specific sequences were found in all isolates, and pulsed-field gel electrophoresis further confirmed considerable chromosome length polymorphism among all sequenced isolates. Apparently, chromosomal rearrangement establishes highly dynamic 'plastic' regions that lead to variation. Interestingly, the highly dynamic plastic genomic regions are enriched for *in planta*-induced genes, including effector genes that contribute to virulence such as the *Ave1* effector in strain JR2 and a LysM effector in strain VdLs.17. Although it is generally assumed that asexual reproduction limits genetic variation, and consequently also adaptive potential, we propose that chromosomal plasticity is a mechanism that allows asexual haploid genomes to adapt to changing environments.

A perspective on next-generation genomics in relation to plant pathogen research and identification of effector genes is provided in **Chapter 6**. We illustrate the power of comparative genomics, and discuss recent studies describing comparative population genomics to identify effector genes. Studying natural variation by next-generation sequencing can be useful to analyse the evolutionary potential of a pathogen population and can be exploited to develop durable resistance in crop species.

Samenvatting

Onderzoek naar effectoren die worden gesecreteerd door ziekteverwekkers tijdens infectie van hun gastheer domineert het onderzoeksveld van moleculaire plant-microbe interacties. Effectoren zijn moleculen die door pathogenen uitgescheiden worden om de gastheerfysiologie te beïnvloeden en zodoende infectie mogelijk te maken. In tegenstelling tot bacteriële type III effectoren die direct in het cytoplasma van de gastheer worden geïnjecteerd, worden effectoren van schimmels en oömyceten gesecreteerd. Echter, voor de meeste effectoren is niet bekend op welke manier ze bijdragen aan virulentie. Terwijl sommige effectoren hun functie buiten de gastheercel uitoefenen, door bepaalde aspecten van de verdediging van de gastheer uit te schakelen, gebruiken andere effectoren een natuurlijk opnamesysteem van gastheercellen om de verdediging intracellulair te onderdrukken of voedingsstoffen vrij te maken. In **hoofdstuk 1** beschrijven we de functie en de evolutie van effectoren van filamenteuze plantpathogenen, aan de hand van de opeenvolgende fasen die tijdens een infectie doorlopen worden.

In **hoofdstuk 2** worden de verspreiding en de eigenschappen van een familie van effectoren, die we LysM effectoren hebben genoemd, beschreven, en laten we zien dat deze familie in het hele schimmelrijk voorkomt. LysM effectoren zijn gesecreteerde eiwitten die enkel lysine motieven (LysMs) bevatten. LysMs staan bekend als koolhydraat-bindende domeinen en komen voor in eiwtten van bacteriën, planten en schimmels. We stellen dat LysM effectoren van schimmels een belangrijke rol spelen bij het wegvangen van chitine oligomeren, afbraakproducten van de schimmelcelwand die vrijkomen tijdens infectie en herkend worden door het immuunsysteem van de gastheer, om op die manier activatie van dit systeem te voorkomen.

In **hoofdstuk 3** onderzochten we de functie van de LysM effector Ecp6 van de plantpathogene schimmel *Cladosporium fulvum*. We laten zien dat Ecp6 chitine bindt en daarmee voorkomt dat chitine fragmenten het immuunsysteem van de plant activeren. In overeenstemming met deze functie kon Ecp6 succesvol concurreren met een plantreceptor voor chitine. Dus, Ecp6 draagt bij aan virulentie van de schimmel door het wegvangen van chitine fragmenten. Aangezien LysM effectoren voorkomen in het hele schimmelrijk stellen we dat het wegvangen van celwandcomponenten een algemene strategie van schimmels kan zijn om de afweer van de gastheer te onderdrukken.

In **hoofdstuk 4** beschrijven we de identificatie van de fysio 1 elicitor die Ve1-gemedieerde weerstand in tomaat tegen *Verticillium* activeert. Door de genoomsequenties van een populatie van fysio 1 en fysio 2 isolaten te bepalen met behulp van zogenaamde "tweede generatie" sequencing technieken is een 50 Kb DNA sequentie gevonden die alleen voorkomt in fysio 1 isolaten. Vervolgens hebben we het transcriptoom van *Verticillium*-geïnfecteerde *Nicotiana benthamiana* planten bepaald, wat resulteerde in de identificatie van één enkel open leesraam dat hoog tot expressie komt in deze regio, die we *Ave1* (voor *Avirulentie op Ve1* tomaat) hebben genoemd. Functionele analyse bevestigde dat Ave1 Ve1-gemedieerde weerstand activeert, en toonde aan dat Ave1 belanrijk is voor virulentie van *Verticillium*, niet alleen op tomaat, maar ook op *Arabidopsis thaliana*. Ave1 homologen zijn wijd verspreid in planten, maar komen ook voor in de

plantpathogene schimmels *Colletotrichum higginsianum*, *Cercospora beticola* en *Fusarium oxysporum* f. sp. *lycopersici*, en in het bacteriële plantpathogeen *Xanthomonas axonopodis*. De verspreiding van Ave1 homologen, en de aanwezigheid van Ave1 in een flexibel genomisch gebied, suggereert dat *Verticillium Ave1* verkregen heeft uit planten door middel van horizontale genoverdracht. Transiënte expressie van de *Ave1* homologen van *F. oxysporum* en *C. beticola* kan Ve1-gemedieerde weerstand activeren. Ve1 geeft dan ook resistentie tegen *F. oxysporum* in tomaat.

In **hoofdstuk 5** hebben we een "vergelijkende genomica"-benadering beschreven om genoomdiversiteit te bestuderen in een populatie van *V. dahliae* isolaten. De sequentiediversiteit binnen de *V. dahliae* isolaten, was over het algemeen laag . Echter, paarsgewijze vergelijking tussen de twee isolaten VdLs.17 en JR2 die sterk op elkaar lijken(> 99,9% identiteit) liet zien dat grote syntenische regio's herhaaldelijk worden afgewisseld met intra- en inter-chromosomale herschikkingen. Breekpunten in syntenie werden geassocieerd met de aanwezigheid van retrotransposons, en zijn vaak geflankeerd door isolaat-specifieke sequenties. Syntenie breekpunten en isolaat-specifieke sequenties kwamen voor in alle isolaten en karyotypering bevestigde aanzienlijke variatie in chromosoomlengtes tussen alle isolaten. We stellen dat chromosomale herschikking zorgt voor dynamische, 'plastische' regio's in het genoom die leiden tot variatie. De plastische regio's zijn verrijkt met *in planta*-geïnduceerde genen, inclusief effectorgenen die bijdragen aan virulentie zoals de *Ave1* effector in JR2 en een LysM effector in VdLs.17. Hoewel algemeen wordt aangenomen dat ongeslachtelijke voortplanting genetische variatie beperkt, in vergelijking met geslachtelijke voortplanting, en daarmee adaptief vermogen, stellen wij dat chromosomale plasticiteit een mechanisme is waarmee asexuele haploïde genomen zich kunnen aanpassen aan veranderende omgevingen.

Een perspectief op "volgende generatie genomica" in relatie tot onderzoek aan plantpathogenen en selectie van effectorgenen wordt besproken in **hoofdstuk 6**. We illustreren de kracht van de vergelijkende genomica, en bespreken recente studies die vergelijkende populatiegenomica gebruiken om effectorgenen te identificeren. Het bestuderen van natuurlijke variatie met behulp van volgende-generatie sequentie technieken kan bruikbaar zijn om het evolutionaire potentieel van een pathogeenpopulatie te analyseren en kan een wetenschappelijke basis geven bij de ontwikkeling van duurzame resistentie in gewassen.

Dankwoord

Om te beginnen zou ik iedereen die dit leest willen bedanken voor zijn/haar hulp tijdens de ruim 4 ½ jaar dat ik bezig ben geweest met het uitvoeren van mijn promotieonderzoek en het schrijven van dit proefschrift! In het bijzonder wil ik nog een aantal mensen specifiek bedanken.

Allereerst wil ik mijn dagelijks begeleider en co-promoter, Bart, ten zeerste bedanken voor de enorme inspiratie en het volle vertrouwen dat je mij hebt gegeven tijdens het uitvoeren en opschrijven van mijn promotieonderzoek. Altijd had je tijd voor mijn vragen, ongeacht welk onderwerp ter discussie stond of de tijd waarop de vraag werd gesteld. Nogmaals enorm bedankt hiervoor!! ;-) Dit proefschrift zou zeker niet tot stand zijn gekomen dankzij mijn promoter, Pierre, en ik wil je dan ook bedanken voor alle goede adviezen, zowel voor als tijdens het schrijven van mijn proefschrift. Verder wil ik je ook bedanken voor het snelle en zorgvuldige nakijken van verschillende manuscripten.

Beste Grardy, Ursula, Emilie, Sajid, Peter, Anja, Andrea, Partha, Mireille, Zhao, Mojtaba, Ermis, Dirk-Jan, Luigi, Jordi & Koste, oftewel, beste (ex)-Verticillium group-members, ik wil jullie allemaal super bedanken voor een top tijd! *Koste, many, many thanks for the great times we shared during the time we worked together. Galatoomi!!* Of Dank U! *I greatly enjoyed our trip to Ethiopia and visits to your home town, family and friends. I will certainly remember our trip to the Bale Mountains, visit to Tulu Dimtu, sightings of the Ethiopian Wolf and the Mountain Nyala and the numerous birds that we spotted (I know you enjoyed those as well ;-)). Cya soon mate!!* Peter, bedankt voor de goede en vruchtbare samenwerking van de afgelopen jaren, welke resulteerde in verscheidene 'top'-verhalen en hopelijk kunnen we die lijn nog even doortrekken. Alvast succes gewenst met de volgende stap in je carriere; ben benieuwd waar die stap zal zijn. *Partha; man I enjoyed working with you! It's been an honor. You're a great friend and I wish you a whole lot of success and happiness back in India and I hope sometime soon we can visit you and make a grand tour through your country. Needless to say; hope to see you soon on your defense! Same to you, Andrea, many thanks for the nice conversations, especially during our 'skilifts' trips ;) and Emilie, many thanks for supervising my first steps @ phyto and the pleasant and plentiful chats we had since then.*

En natuurlijk mijn paranimfen, Anja en Grardy. Anja, reuze bedankt voor de super gezellige tijd op fyto, de toffe gesprekken en natuurlijk de goede samenwerking! Alvast veel succes gewenst met het schrijven van jouw proefschrift! Grardy, zonder jouw hulp waren talloze projecten nog nergens en bleven theoriën gewoon theoriën; hiervoor mijn dank. Maar nog veel belangrijker, ik heb zeer genoten van de tijd dat we samenwerkten. De gesprekken over van-alles-en-nog-wat onder de koffie 's morgens waren zeer fijn en een welkom begin van de dag, ik zal ze gaan missen (en dat is al begonnen).

Verder wil ik iedereen op fyto bedanken voor de topsfeer, prettige samenwerkingen en goede discussies. In het bijzonder mijn dank aan Jan, Ate, Matthieu en Harrold voor de nuttige inhoudelijke discussies maar ook voor de sportieve ondernemingen.

Over sportieve ondernemingen gesproken, graag onderbrak ik het werk om een weekend of week op pad te gaan om straffe klimmen af te wisselen met technische single-tracks en zinderende afdalingen.

Hiervoor mijn dank aan Robert, Lucas, Wouter D., Wouter S., Karel, Raoul, Laurent, Jan, Euan, Johan en natuurlijk alle anderen die er met enige regelmaat bij waren!

Hierbij wil ik ook mijn familie bedanken, en in het bijzonder mijn paps en mams, die mij altijd de vrijheid hebben gegeven om mijn dromen te verwezelijken. Al vroeg verliet ik het nest om te gaan studeren en nu pas, na 10 jaar, ben ik weer even in de buurt komen wonen. Ontzettend bedankt voor jullie steun en toeverlaat wanneer dat nodig was. Dennis en Annemarie, al vroeg zijn wij alledrie een andere richting ingeslagen en misschien genoot ik daarom nog wel des te meer van de gezellige gesprekken wanneer ik weer eens in het Zeeuwse was. Bedankt daarvoor!

Tot slot, de belangrijkste persoon in mijn leven. Lieve Manon, ontzettend bedankt voor je schitterende glimlach, liefde en nimmer aflatende steun. Bedankt voor je geduld wanneer ik weer eens tot laat bleef werken, thuis of in't lab, of in het weekend nog even door moest om een manuscript of mijn proefschrift af te ronden. Nu storten we ons vol in het Belgische avontuur en ik hoop dat we hier samen veel plezier gaan beleven.

Curriculum Vitae

Ronnie de Jonge is geboren op 1 september 1984 te Goes (NL). Na het behalen van zijn VWO diploma in 2002 begon hij aan de BSc opleiding Biologie met de specialisatie "Celbiologie", gevolgd door de MSc opleiding Plant Biotechnology, met de specialisatie "Plant Pathology and Pest Control", beide aan de Wageningen Universiteit. Tijdens zijn eerste afstudeeronderzoek onderzocht hij het gebruik van RNA-gemedieerde silencing in de plant pathogene schimmels *Cladosporium fulvum* en *Verticillium dahliae* om potentiële effectoren te bestuderen onder leiding van Ir. Emilie Fradin en Dr. Ir. Bart Thomma bij de leerstoelgroep Fytopathologie aan de Wageningen Universiteit. Tijdens zijn afstudeerstage onderzocht hij de expressie van hoog licht induceerbare genen in cyanobacteriën en diatomeën onder leiding van Dr. Ir. Jacco Kromkamp aan het Nederlands Instituut voor Ecologie te Yerseke (NL). In zijn 2e afstudeervak heeft hij de expressie van type-III effectoren in de plant pathogene bacterie *Pectobacterium carotovorum* tijdens verschillende infectiestadia onderzocht onder leiding van Dr. Ir. Eleanor Gilroy en Prof. Dr. Ir. Paul Birch aan het Scottish Crop Research Institute te Dundee (GB). Na succesvolle afronding van zijn MSc opleiding in november 2007 heeft hij als assistent in opleiding onderzoek gedaan aan effectoren van de plant pathogene schimmels *C. fulvum* en *V. dahliae* bij de leerstoelgroep Fytopathologie aan de Wageningen Universiteit, onder begeleiding van Dr. Ir. Bart Thomma. De resultaten van dit onderzoek zijn beschreven in dit proefschrift. In september 2012 is hij begonnen als postdoctoraal onderzoeker aan het Vlaams Instituut voor Biotechnologie te Gent (BE) onder begeleiding van Prof. Dr. Ir. Yves van de Peer met als doel de evolutie van effectoren in schimmelpopulaties te bestuderen.

List of publications

Zhang Z, Fradin EF, de Jonge R, van Esse HP, Smit P, Liu C-M, Thomma BPHJ, **Optimized agroinfiltration and virus-induced gene silencing to study Ve1-mediated *Verticillium* resistance in tobacco.** *Mol Plant Microbe Interact* (Posted online on 19 Sep 2012; doi: http://dx.doi.org/10.1094/MPMI-05-12-0138-R)

Liebrand TWH, Smit P, Abd-El-Haliem A, de Jonge R, Cordewener JHG, America AHP, Sklenar J, Jones AME, Robatzek S, Thomma BPHJ, Tameling WIL, Joosten MHAJ, **ER-quality control chaperones facilitate the biogenesis of Cf receptor-like proteins involved in pathogen resistance of tomato.** *Plant Physiol*, 2012, **159**:1819-1833

Faino L, de Jonge R and Thomma BPHJ, **The transcriptome of *Verticillium dahliae*-infected *Nicotiana benthamiana* determined by deep RNA sequencing.** *Plant Signal Behav*, 2012, **7**:1-5

de Jonge R*, van Esse HP*, Maruthachalam K, Santhanam P, Saber MK, Zhang Z, Usami T, Lievens B, Subbarao KV and Thomma BPHJ, **Tomato Immune Receptor Ve1 Recognizes Effector of Multiple Tomato Fungal Wilt Pathogens Uncovered by Genome and RNA Sequencing.** *Proc Natl Acad Sci USA*, 2012, **109**:5110-5115

de Jonge R, ***In silico* identification and characterization of effector catalogs.** *Methods Mol Biol*, 2012, **835**:415-425

Klosterman SJ, Subbarao KV, Kang S, Veronese P, Gold SE, Thomma BPHJ, Chen Z, Henrissat B, Lee Y-H, Park J, Garcia-Pedrajas MD, Barbara DJ, Anchieta A, de Jonge R, Santhanam P, Maruthachalam K, Atallah Z, Amyotte SG, Paz Z, Inderbitzin P, Hayes RJ, Heiman DI, Young S, Zeng Q, Engels R, Galagan J, Cuomo C, Dobinson KF and Ma LJ, **Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens.** *PLoS Pathog*, 2011, **7**:e1002137

de Jonge R, Bolton MB and Thomma BPHJ, **How filamentous pathogens co-opt plants; the ins and outs of fungal effectors**. *Curr Opin Plant Biol*, 2011, **14**:400-406

de Jonge R, van Esse HP, Kombrink A, Shinya Y, Desaki Y, Bours R, van der Krol S, Shibuya N, Joosten MHAJ and Thomma BPHJ, **Conserved fungal LysM Effector Ecp6 inhibits chitin-triggered immunity in plants.** *Science*, 2010, **329**:953-955

de Jonge R and Thomma BPHJ, **Fungal LysM effectors: extinguishers of host immunity?** *Trends Microbiol*, 2009, **17**:151-157

Ellendorff U*, Fradin EF*, de Jonge R* and Thomma BPHJ, **RNA silencing is required for Arabidopsis defence against Verticillium wilt disease.** *J Exp Bot*, 2009, **60**:591-602

Bolton MD, van Esse HP, Vossen JH, de Jonge R, Stergiopoulos I, Stulemeijer IJE, van den Berg GCM, Borrás-Hidalgo O, Dekker HL, de Koster CG, de Wit PJGM, Joosten MHAJ and Thomma BPHJ, **The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species.** *Mol Microbiol*, 2008, **69**:119-136

*shared first authorship

Education Statement of the Graduate School

Experimental Plant Sciences

The Graduate School EXPERIMENTAL PLANT SCIENCES

Issued to: **Ronnie de Jonge**
Date: **9 November 2012**
Group: **Laboratory of Phytopathology, Wageningen University & Research Centre**

1) Start-up phase	*date*
► **First presentation of your project**	
Analysis of the Verticillium - Arabidopsis interaction making use of whole-genome tools	Feb 05, 2008
► **Writing or rewriting a project proposal**	
Analysis of the Verticillium - Arabidopsis interaction making use of whole-genome tools	Jan 10, 2008
► **Writing a review or book chapter**	
In silico identification/characterization of effector catalogs (MiMB; 2012)	Autumn 2010
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*9,5 credits**

2) Scientific Exposure	*date*
► **EPS PhD student days**	
EPS PhD student day, Naturalis Museum, Leiden	Feb 26, 2009
EPS PhD student day, Wageningen University	May 20, 2011
► **EPS theme symposia**	
EPS Theme 2: Interactions between Plants and Biotic Agents & Willie Commelin scholten day, Utrecht University	Jan 22, 2009
EPS Theme 2: Interactions between Plants and Biotic Agents & Willie Commelin scholten day, Amsterdam	Feb 03, 2011
► **NWO Lunteren days and other National Platforms**	
EPS meeting 'Experimental Plant Sciences', Lunteren	Apr 07-08, 2008
ALW meeting 'Molecular Biology and Genetics', Lunteren	Sep 18-19, 2008
NBIC meeting, Lunteren	Mar 17-18, 2009
EPS meeting 'Experimental Plant Sciences', Lunteren	Apr 06-07, 2009
ALW meeting 'Molecular Biology and Genetics', Lunteren	Oct 15-16, 2009
EPS meeting 'Experimental Plant Sciences', Lunteren	Apr 19-20, 2010
► **Seminars (series), workshops and symposia**	
Genomics of *Stagonospora nodorum;* Genes, Genomes and Growers	Oct 16, 2008
Seminar of Prof.dr. T. Nürnberger, University of Tübingen, ZMBP - Plant Biochemistry, Germany: Patterns and receptors in plant immunity	Dec 17, 2008
Seminar of Dr. Pieter van West, University of Aberdeen: Saprolegnia parasitica an oomycete with a fishy appetite, new challenges for an old problem	Feb 03, 2009
Seminar of Dr. Rays H.Y. Jiang, Broad Institute: Host-pathogen interaction drives genome plasticity in animal and plant pathogens	Jun 10, 2009
Plant Sciences Seminar by Prof.dr.ir. Pierre de Wit and Prof.dr. Fred van Eeuwijk, Wageningen University	Nov 10, 2009
Seminar of Dr. Brigitte Mauch-Mani: Grapevine and downy mildew - Wine is not the only difference between grapevine and Arabidopsis, Université de Neuchâtel, CH-2009 Neuchâtel, Switzerland	May 31, 2010
Seminar of Prof. Felix Mauch: Old fashioned secondary metabolites save Arabidopsis from Phytophthora brassicae, University of Fribourg, CH-1700 Fribourg, Switzerland	May 31, 2010
Seminar of Prof. Naoto Shibuya : PAMP receptor (CBEIC and CERK1)	Sep 09, 2010
Joint meeting CBS-Phytopathology on Bioinformatics and Medical Mycology	Nov 12, 2010
New Frontiers in Pattern Recognition Receptors symposium	Nov 05-06, 2009
► **International symposia and congresses**	
Verticillium genome Sequencing Meeting (APS)	Jul 21-26, 2008
Benelux Bioinformatics Conference 2008	Dec 14-16, 2008
IM9 Congres on The Biology of Fungi	Aug 01-06, 2010
Fungal Genetics in Asilomar (USA, California)	Mar 01-06, 2011
2nd Joined Program Phytopathology, Wageningen UR - MPI Marburg, Marburg	Jan 30-31, 2012
► **Presentations**	
Poster presentation APS	Jul 21-26, 2008
Oral presentation APS	Jul 21-26, 2008
Oral presentation EPW	Apr 06-07, 2009
Oral presentation EPW	Apr 19-20, 2010
Poster presentation IMC9	Aug 01-06, 2010
Poster presentation FungalGenetics, Asilomar	Mar 01-06, 2011
Oral presentation EPS Student Day, Wageningen University	May 20, 2011
Oral presentation 2nd Joined Program Phytopathology - MPI Marburg, Marburg	Jan 30-31, 2012
► **IAB interview**	Feb 18, 2011
► **Excursions**	
Subtotal Scientific Exposure	*21,8 credits**

3) In-Depth Studies	*date*
► **EPS courses or other PhD courses**	
Summer School 'On the Evolution of Plant Pathogen Interactions: from Principles to Practice'	Jun 18-20, 2008
Course "Next Generation Sequencing (NGS) data analysis" (Leiden, MGC/CMSB/NBIC)	Sep 29-Oct 01, 2009
Spring School 'RNAi & the World of Small RNA Molecules' (Wageningen, EPS)	Apr 14-16, 2010
► **Journal club**	
Participate in literature discussion group (Phytopathology) (Weekly)	2007-2011
► **Individual research training**	
Two weeks visit to Toulouse, Group of Jerome Gouzy. Genomics of Verticillium and other fungi	2009
Two weeks visit to Toulouse, Group of Yves Marco. uArray of Arabidopsis	2009

►	1 1/2 weeks visit to Marburg, Max Planck Institute, Group of Eva Stukenbrock	2011
	Subtotal In-Depth Studies	*8,4 credits**

4) Personal development		*date*
►	**Skill training courses**	
	Minisymposium: How to write a world class paper.	Apr 19, 2011
	EPS Expectations Career day	Nov 19, 2010
	Adobe Indesign (Hans de Jong and Boudewijn van Veen)	Nov 15, 2010
	Advanced Course Guide to Scientific Artwork (Hans de Jong and Boudewijn van Veen)	May 10-11 2010
	Career Assessment (Meijer en Meijaard)	Jun 2012
►	**Organisation of PhD students day, course or conference**	
►	**Membership of Board, Committee or PhD council**	
	PhD representative Phytopathology (4 annual meetings, 4 years)	Jan 2008 - Jan 2012
	Subtotal Personal Development	*3,9 credits**

TOTAL NUMBER OF CREDIT POINTS*	**43,6**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits

** A credit represents a normative study load of 28 hours of study.*

This research was conducted in the Laboratory of Phytopathology of Wageningen University and was financially supported by the Netherlands Organization for Scientific Research (NWO). The printing of this thesis was financially supported by the J.E. Jurriaanse Stichting.

Front cover: Tomato plants infected with the vascular wilt fungus *Verticillium dahliae*. The left plant is healthy because it carries the Ve1 immune receptor, while the right plant that lacks this receptor is stunted. Plants are represented in DNA sequence to illustrate identification of the *V. dahliae Ave1* gene that encodes the ligand of the Ve1 immune receptor using next-generation sequencing. Intriguingly, *V. dahliae* acquired *Ave1* from plants through horizontal gene transfer.

Back cover: Whole-genome dot-plot comparison between *Verticillium dahliae* strains VdLs.17 and JR2 with forward-forward alignments in white and inversions in green reveals extensive chromosomal rearrangements.

Illustrations by Ronnie de Jonge

Printed at Wöhrmann Print Service, Zutphen, The Netherlands.