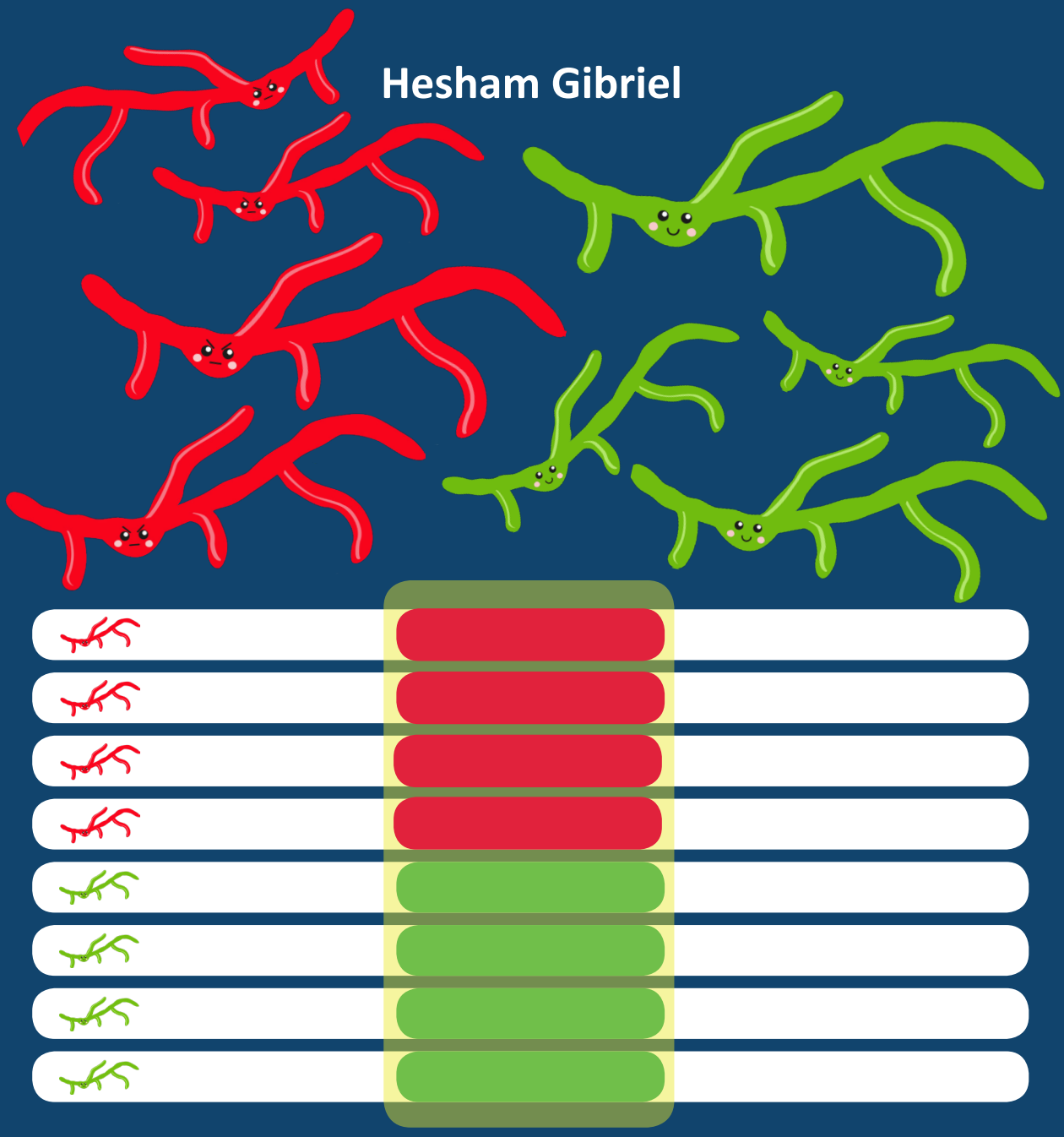
Comparative pathogenomics in the ascomycete plant pathogens *Zymoseptoria tritici* and *Verticillium dahliae*

Hesham Gibriel



Propositions

1. Improper gene annotation of *Zymoseptoria tritici* hampered the discovery of AvrStb6 for years.
   (this thesis)

2. *Verticillium dahliae* strains infecting the same host plant utilize divergent effector repertoires.
   (this thesis)

3. Artificial insulin-releasing cells may be the key therapy for diabetes in the future.

4. Establishments of networks between weather forecasters in the developed world and governmental organizations in developing regions can diminish the impact of future weather disasters.

5. Scientists should engage in reporting their research results publicly on social media to raise awareness of the value of science.

6. Obesity should not be blamed on genes, but on nutritional habits.

7. Considering the availability of vast amounts of solar energy in the Middle East, solar energy should become the main source of power for the region.

8. Bodybuilding athletes should consider healthy diets as replacement for protein supplements.

Propositions belonging to the thesis, entitled

'Comparative pathogenomics in the ascomycete plant pathogens *Zymoseptoria tritici* and *Verticillium dahliae'.*

Hesham Gibriel

Wageningen, 20 February 2019

Comparative pathogenomics in the ascomycete plant pathogens *Zymoseptoria tritici* and *Verticillium dahliae*

Hesham Gibriel

Thesis committee

Promotor
Prof. Dr B.P.H.J. Thomma
Professor of Phytopathology
Wageningen University & Research

Co-promotors
Dr M.F. Seidl
Assistant Professor, Laboratory of Phytopathology
Wageningen University & Research

Prof. Dr G.H.J. Kema
Special Professor Tropical Phytopathology
Wageningen University & Research

Other members
Prof. Dr J.E. Kammenga, Wageningen University & Research
Dr A. Sanchez-Vallet, ETH Zurich, Switzerland
Dr M.H. Medema, Wageningen University & Research
Dr Ronnie de Jonge, Utrecht University

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

Comparative pathogenomics in the ascomycete plant pathogens *Zymoseptoria tritici* and *Verticillium dahliae*

Hesham Gibriel

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 20 February 2019
at 4 p.m. in the Aula.

Hesham Gibriel
Comparative pathogenomics in the ascomycete plant pathogens
Zymoseptoria tritici and *Verticillium dahliae*,
238 pages.

PhD thesis, Wageningen University, Wageningen, The Netherlands (2019)
With references, with summary in English

DOI: 10.18174/467538
ISBN: 978-94-6343-572-7

Table of contents

Chapter 1	General introduction	7
Chapter 2	The age of effectors: genome-based discovery and applications	21
Chapter 3	*Verticillium dahliae* strains that infect the same host plant display highly divergent effector catalogs	43
Chapter 4	The *Verticillium dahliae* effector Sun1 quantitatively contributes to virulence on sunflower	77
Chapter 5	Signatures of adaptation to bread or durum wheat in a global collection of *Zymoseptoria tritici* isolates	107
Chapter 6	Stress and sexual reproduction affect the dynamics of the wheat pathogen effector AvrStb6 and strobilurin resistance	153
Chapter 7	General discussion: Effector repertoires of plant pathogens revealed by comparative genomics	209
Summary		225
Acknowledgements		227
About the author		229
List of publications		231
Education statement		233

Chapter 1

General introduction

HOST-MICROBE INTERACTIONS

Genomes are highly variable and differ significantly between, and even within, species. This genetic variability is established by a variety of mechanisms ranging from single nucleotide polymorphisms (SNPs) to large-scale structural variations, and provides the basis for species to continuously evolve under changing environmental conditions (Seidl and Thomma 2014). This is particularly relevant for microbial pathogens as they are engaged in co-evolutionary arms races with their hosts, where the pathogen tries to establish the parasitic relationship while the host tries to ward off the pathogen (Dong et al. 2015; Raffaele and Kamoun 2012; Seidl and Thomma 2014).

Microbial plant pathogens cause devastating diseases on important crops, leading to billions of dollars of losses annually and threatening global food security (Fisher et al. 2012). To defend themselves against invaders, plants have evolved extracellular and intracellular receptors that can detect invading microbes through recognition of so-called invasion patterns that betray the presence or activity of an invader (Cook et al. 2015). To establish themselves, invaders secrete so-called effectors that suppress host immune responses and facilitate the symbiosis (Cook et al. 2015; Rovenich et al. 2014). However, also effectors can act as invasion patterns when they are detected by plant immune receptors (Cook et al. 2015). Therefore, pathogens need to continuously diversify their effector repertoires to avoid recognition by the plant immune system by evolving novel effectors, or losing or modifying existing ones, while the plant simultaneously evolves novel receptors to intercept pathogens (Jones and Dangl 2006). Thus, plants and pathogens are engaged in co-evolutionary arms races that shape their evolution. It can be anticipated that this co-evolution ultimately leads to highly specialized pathogens that can only infect few or even a single host. However, some pathogens are successful pathogens of a broad range of hosts. Such broad host range pathogens, arguably, have to engage in arms races with multiple hosts. This leads to the question how broad host range pathogens are able to adapt to be competitive in the arms race with multiple hosts.

In my research, I have used two ascomycete plant pathogens in order to study their adaptation on a narrow versus a broad host range. More specifically, I focussed on *Zymoseptoria tritici* that has a narrow host range with wheat as its major host although it occasionally infects other grasses as well. Additionally, I studied *Verticillium dahliae* that is able to colonize a broad range of plant hosts, encompassing hundreds of species (Figure 1).

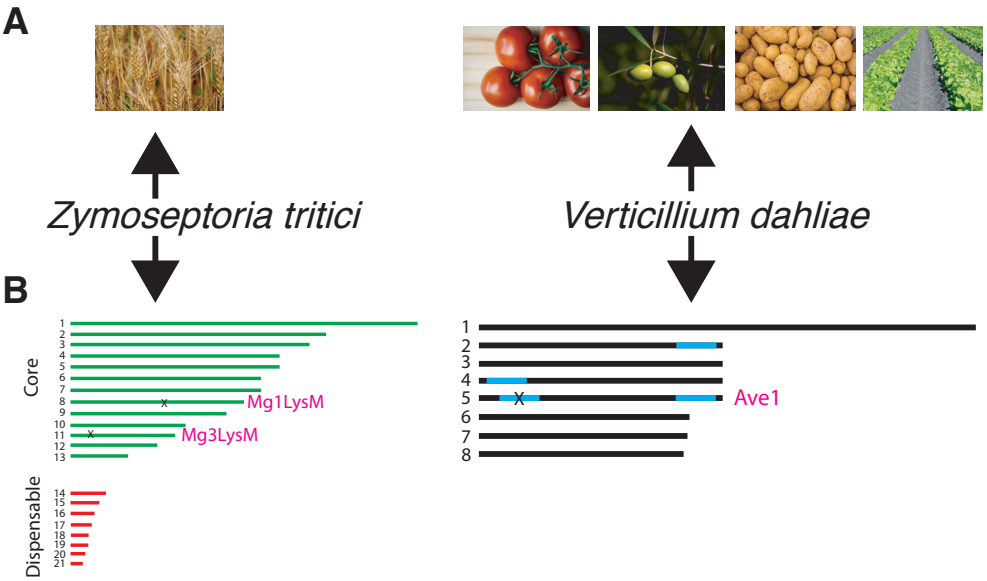


Figure 1. Two selected ascomycete plant pathogens used in this thesis, which have divergent host ranges and encode effector genes located on different genomic regions. (A) On the left side, the fungal wheat pathogen *Z. tritici* infects only few host plants, while on the right side *V. dahliae* colonizes a diverse host species including tomato, olive, potato, and lettuce. **(B)** The genomes and the location of a selection of previously characterized effector genes of these two pathogens are shown. On the left side, the genome of *Z. tritici* is shown that contains 13 core chromosomes (green) and eight dispensable ones (red). Effector genes in *Z. tritici* are located on the core chromosomes. For example, the effector genes *Mg3LysM* and *Mg1LysM* are located on chromosome 11 and chromosome 8, respectively. No effector genes are known to reside on the dispensable ones. On the right side, the genome of *V. dahliae* is shown, which contains 8 chromosomes. This pathogen contains lineage-specific regions (blue), which encode effector genes. For example, the effector gene *Ave1* is localized within such an LS region.

THE FUNGAL WHEAT PATHOGEN *ZYMOSEPTORIA TRITICI*

Wheat is one of the world's major crops that provides about 20% of all calories consumed by people worldwide (Shiferaw et al. 2013). However, wheat is susceptible to a variety of pathogens including the notorious fungal pathogen *Z. tritici* that causes Septoria tritici leaf blotch (STB). This disease is considered the economically most damaging disease of wheat worldwide (Fones and Gurr 2015). In Europe, the disease leads to 1 billion euro losses per year (Kettles and Kanyuka 2016). STB disease management strongly relies on fungicide usage (Torriani et al. 2015). However, the extensive use of fungicides has resulted in the emergence of fungicide resistant *Z. tritici* strains over the last decades (Cools and Fraaije 2008). Therefore, screening for STB resistance in wheat germplasm

is an important cornerstone for disease management. To date, 21 *Stb* resistance genes have been genetically mapped (Brown et al. 2015). *Stb6*, which is a common resistance gene that is exploited in European wheat cultivars (Chartrain et al. 2005), is the first resistance gene that was recently cloned (Saintenac et al. 2018).

The life cycle of *Z. tritici* includes both sexual and asexual reproduction. Sexual reproduction involves the production of air-borne sexual ascospores year-round with peaks at the onset of the growing season, thus providing genetically diverse *Z. tritici* populations (Hunter et al. 1999). Subsequently, asexual conidia are produced and disseminated over short distances (Wittenberg et al. 2009). The infection process of *Z. tritici* on wheat plants is initiated by sexual ascospores or asexual pycnidiospores that land onto the leaf surface by rain splash. Subsequently, spores germinate and the fungus colonizes the apoplastic space between mesophyll cells without apparent damage to host cells (Ponomarenko et al. 2011). This symptomless phase is referred to as the biotrophic phase. After 7-10 days, the fungus switches to a symptomatic phase, referred to as necrotrophic, where lesions appear in which pycnidia develop (Ponomarenko et al. 2011).

Z. tritici is pathogenic on both bread wheat (*Triticum aestivum*) and durum wheat (*T. turgidum*) cultivars (Annone et al. 1996). However, individual *Z. tritici* isolates exhibit strong host species specificity (Kema and van Silfhout 1997). This means that the majority of *Z. tritici* isolates originating from durum wheat are virulent on the majority of durum wheat cultivars, but avirulent on bread wheat cultivars, and *vice versa*. Moreover, individual *Z. tritici* isolates exhibit cultivar specificity, meaning that *Z. tritici* isolates are virulent or avirulent on a particular cultivar within a wheat species. In this manner, *Z. tritici* isolate IPO94269 is virulent on bread wheat cultivar Shafir that carries *Stb6*, while *Z. tritici* isolate IPO323 is avirulent on this cultivar. However, so far, *Z. tritici* gene(s) associated with host and cultivar species specificity have not been identified.

Evolution of virulence in Z. tritici

Advances in high-throughput sequencing technologies have facilitated rapid sequencing of entire genomes of fungal plant pathogens (Metzker 2010). This has allowed phytopathologists to gain insight into the evolution of virulence by studying genomic plasticity and determining effector catalogs. The genome of isolate IPO323 was the first *Z. tritici* isolate to be sequenced. It has a size of 39.69 Mb, divided over 21 chromosomes (Goodwin et al. 2011). Notably, *Z. tritici* has 13 core chromosomes (CCs, chromosome 1-13) that are present in all isolates, while eight chromosomes are considered to be conditionally dispensable chromosomes (CDCs, chromosome 14-21). These latter can be absent in an isolate without an obvious effect on fungal fitness (Goodwin et al. 2011). Over the past few years, there has been a growing interest to investigate the origin and role of these CDCs in the pathogenicity of *Z. tritici*. A comparative genomics study of species closely related to *Z. tritici* has suggested that at least some CDCs originated prior to the emergence of *Z. tritici* as a specialized wheat pathogen (Stukenbrock et al. 2010).

Moreover, it has been shown that CDCs can undergo chromosome fusions, resulting in newly formed chromosomes, and thus can contribute to the rapid adaptive evolution (Croll et al. 2013). The genome of *Z. tritici* is considered to be complete, i.e. without any gaps, and includes telomeres and centromeres (Goodwin et al. 2011), which should facilitate effector gene discovery (Thomma et al. 2016). However, thus far there are no known effector genes that are located on *Z. tritici* CDCs, and thus the role of these CDCs in pathogenicity remains poorly understood.

Comparative genomics of *Z. tritici* identified genomic compartments that harbour candidate effector genes. Comparative genomics between *Z. tritici* reference isolate IPO323 and the highly virulent isolate 3D7 identified orphan genomic regions that are unique to one of the two genomes (Plissonneau et al. 2016). The orphan regions in *Z. tritici* isolate 3D7 contain genes encoding secreted proteins (Plissonneau et al. 2016). Additionally, a recombination map was generated for *Z. tritici* to identify recombination hotspots, i.e. regions that display a high recombination frequency. Notably, these recombination hotspots are located in proximity to telomeres, and are enriched for genes encoding secreted proteins (Croll et al. 2015).

Genomes can be mined for homologs of effectors with known functions, as it can be anticipated that these homologs play similar roles in other fungal plant pathogens. Analysis of the *Z. tritici* genome indicated the presence of three effector genes encoding LysM proteins (Mg3LysM, Mg1LysM, and MgxLysM) (Marshall et al. 2011), which are homologs of *Ecp6* (for extracellular protein 6), a LysM effector gene of the fungal tomato pathogen *Cladosporium fulvum* (de Jonge et al. 2010). Two of these effectors (Mg1LysM and Mg3LysM) have been functionally analysed, and both effectors are able to bind chitin and protect fungal hyphae against plant hydrolytic enzymes (Marshall et al. 2011). In addition, Mg3LysM is able to block the activation of chitin-induced immunity in host plants (Marshall et al. 2011).

Z. tritici genome and transcriptome analysis revealed many additional effector candidates that may be involved in virulence (Goodwin et al. 2011; Grandaubert et al. 2015; Rudd et al. 2015). For example, a genome-wide mining for genes encoding small, secreted, cysteine-rich proteins that are induced *in planta*, commonly used criteria to define effectors in fungal pathogens (Gibriel et al. 2016), revealed 78 effector candidates (Mirzadi Gohari et al. 2015). However, none of them have been functionally characterized as effector yet, by showing either a role in immune suppression or a contribution to virulence (Mirzadi Gohari et al. 2015). Additionally, three *Z. tritici* genes were found to be subject to positive selection and to play a role in virulence (Poppe et al. 2015). However, only one of them encodes a signal peptide and thus qualifies as a *bona fide* effector (Poppe et al. 2015).

In order to identify full effector repertoires of plant pathogens, it is required to have properly annotated gene models (Gibriel et al. 2016). However, for *Z. tritici*, many of

the previously annotated gene models have incorrect exon/intron boundaries and lack either a start or a stop codon (Grandaubert et al. 2015). Arguably, such incorrect gene annotation compromises discovery of the complete effector catalog in this fungal pathogen.

VERTICILLIUM WILT DISEASES

It is unknown how broad host range pathogens engage in arms races with multiple hosts (Dong et al. 2015). Here, I focus on *V. dahliae*, a broad host range pathogen, to understand processes of host adaptation. This pathogen belongs to the genus *Verticillium* that comprises ten species of soil-borne ascomycete fungi (Inderbitzin et al. 2011). Few of these species collectively cause Verticillium wilt diseases on hundreds of plant hosts across the world (Inderbitzin et al. 2011; Klimes et al. 2015). For example, *V. dahliae*, which is the most notorious plant pathogenic species within this genus, is able to infect more than 200 plant hosts including economically important crops such as tomato, potato, strawberry, lettuce and cotton (Fradin and Thomma 2006; Inderbitzin et al. 2011). This pathogen infects the plant by penetrating the root, after which it enters the xylem and starts to produce conidia that are carried upward with the water flow to distal plant parts (Fradin and Thomma 2006). In contrast to *V. dahliae*, the other pathogenic *Verticillium* species (*V. albo-atrum*, *V. alfalfae*, *V. non-alfalfae*, and *V. longisporum*) have considerably narrower host ranges.

Control of Verticillium wilt diseases is difficult due to the long viability of its resting structures and the inability of fungicides to affect the pathogens once they enter the xylem (Fradin and Thomma 2006). Therefore, the use of resistant cultivars is the preferred strategy for disease management. Thus far, tomato Ve1 is the only immune receptor that provides resistance against particular *V. dahliae* and *V. alfalfae* strains (Fradin et al. 2009). Recently, functional homologs of tomato *Ve1* have also been identified in tobacco, potato, wild eggplant and hop (Song et al. 2016).

Evolution of virulence in V. dahliae

Sexual reproduction is considered an important driver for adaptive genome evolution. However, sexual reproduction has never been observed in *V. dahlia*e although it is a successful pathogen that evolved the capacity to infect hundreds of hosts (Fradin and Thomma 2006; Inderbitzin and Subbarao 2014). Previously, genomic studies have provided insights into the evolution of virulence of *V. dahliae*. Genome comparisons of multiple *V. dahliae* strains revealed chromosomal length polymorphisms between closely related *V. dahliae* strains (de Jonge et al. 2013). These length polymorphisms are caused by extensive genome rearrangements that are mediated by erroneous double-strand break repair pathways, often utilizing abundant transposable elements as a substrate for repair (Faino et al. 2016). Moreover, these chromosomal rearrangements facilitate the formation of highly dynamic, repeat-rich, lineage-specific (LS) regions that

are only present in a subset of *V. dahliae* strains (de Jonge et al. 2013; Faino et al. 2016). Notably, these LS regions can amount to up to ~4 Mb of the total genome size and contain up to 1,000 genes (de Jonge et al. 2013). Interestingly, many *in planta*-induced effector genes are localized within these highly dynamic LS regions (de Jonge et al. 2013). For example, *Ave1*, which is an important effector gene that is involved in virulence on many host plants, was found to be localized within an LS region (de Jonge et al. 2012). Similarly, the gene encoding the LysM domain-containing Vd2LysM effector that is involved in virulence on tomato is located in an LS region (Kombrink et al. 2017). By definition, LS regions differ significantly between *V. dahliae* strains, and thus different *V. dahliae* strains commonly carry distinct sets of effector genes in their LS regions (de Jonge et al. 2013). However, *V. dahliae* strains with differences in their effector gene repertoires have the ability to infect the same host plant. For example, *V. dahliae* strain JR2 carries the *Ave1* effector gene, which is involved in virulence in tomato plants lacking *Ve1*, while *V. dahliae* strain VdLs17 lacks this gene. Loss of this gene in VdLs17 reduces its virulence on tomato, although VdLs17 remains the capacity to infect this host plant (de Jonge et al. 2012). This *V. dahliae* strain uses other effectors than Ave1, such as Vd2LysM, to infect this host (Kombrink et al. 2017). Therefore, different LS effectors enable *V. dahliae* strains to infect the same host.

MAIN RESEARCH QUESTION

In this thesis research, I have used comparative population genomics to analyse the ascomycete plant pathogens *Z. tritici* and *V. dahliae*. My aim was to mine for genomic characteristics of these two pathogens that can explain their narrow or broad host range, respectively.

THESIS OUTLINE

In **Chapter 2**, I review recent advances in genome sequencing technologies, genome assembly, gene annotation, as well as effector identification methods to reveal complete effector repertoires. Finally, I explain my view of how the knowledge of effector diversity can be exploited to develop sustainable resistance breeding strategies.

In **Chapter 3**, I show that *V. dahliae* strains that are able to infect the same host plant harbour highly divergent LS effector repertoires. First, I determined effector genes localized within the core and lineage-specific (LS) genomic regions of a collection of *V. dahliae* strains. Subsequently, I determined the ability of these strains to infect a collection of plant species that belong to various plant families.

In **Chapter 4**, I show that *V. dahliae* strain 85S that infect sunflower plant harbour a host-specific effector that contribute to disease establishment. First, I conducted comparative

genomics of a sunflower pathogenic strain with a collection of non-pathogenic strains and identified an effector gene that exclusively occur in the sunflower pathogenic strain and that is highly induced during host colonization. Functional analysis revealed that this effector gene quantitatively contributes to *V. dahliae* virulence on sunflower, but not on other host plants.

In **Chapter 5**, I show extensive genetic diversity amongst *Z. tritici* isolates that are adapted to infect either bread wheat (BW) or durum wheat (DW). First, I conducted whole-genome sequencing of a worldwide collection of 136 *Z. tritici* isolates and phenotyping assays on a set of BW and DW cultivars. Subsequently, I assessed genome-wide differences between BW and DW isolates and identified four effector genes carrying non-synonymous single nucleotide polymorphisms in BW isolates.

In **Chapter 6**, I show the cloning of the first avirulence effector gene of *Z. tritici*; *AvrStb6*, and show that this effector gene is maintained in *Z. tritici* populations. I first conducted comparative population genomics combined with genetic mapping and identified a polymorphic genomic region that harbours *AvrStb6*. The paradigm states that deployment of Stb6 leads to removal of this avirulence effector gene from pathogen populations. However, when we crossed *Z. tritici* on wheat, we found that sex occurs even with an avirulent parent, thus leading to the maintenance of *AvrStb6* in subsequent populations.

In **Chapter 7**, the major findings of this thesis are discussed in a broader context and directions for future research are suggested.

REFERENCES

Brown, J. K. M., Chartrain, L., Lasserre-Zuber, P., and Saintenac, C. 2015. Genetics of resistance to *Zymoseptoria tritici* and applications to wheat breeding. Fungal Genetics and Biology 79:33-41.

Chartrain, L., Brading, P. A., and Brown, J. K. M. 2005. Presence of the *Stb6* gene for resistance to septoria tritici blotch (*Mycosphaerella graminicola*) in cultivars used in wheat-breeding programmes worldwide. Plant Pathology 54:134-143.

Cook, D. E., Mesarich, C. H., and Thomma, B. P. H. J. 2015. Understanding plant immunity as a surveillance system to detect invasion. Annual Review of Phytopathology 53:541-563.

Cools, H. J., and Fraaije, B. A. 2008. Are azole fungicides losing ground against Septoria wheat disease? Resistance mechanisms in *Mycosphaerella graminicola*. Pest Management Science 64:681-684.

Croll, D., Zala, M., and McDonald, B. A. 2013. Breakage-fusion-bridge cycles and large insertions contribute to the rapid evolution of accessory chromosomes in a fungal pathogen. PLoS Genetics 9:e1003567.

Croll, D., Lendenmann, M. H., Stewart, E., and McDonald, B. A. 2015. The impact of recombination hotspots on genome evolution of a fungal plant pathogen. Genetics 201:1213-1228.

de Jonge, R., Bolton, M. D., Kombrink, A., van den Berg, G. C., Yadeta, K. A., and Thomma, B. P. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Research 23:1271-1282.

de Jonge, R., Peter van Esse, H., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., van der Krol, S., Shibuya, N., Joosten, M. H. A. J., and Thomma, B. P. H. J. 2010. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. Science 329:953-955.

de Jonge, R., Peter van Esse, H., Maruthachalam, K., Bolton, M. D., Santhanam, P., Saber, M. K., Zhang, Z., Usami, T., Lievens, B., and Subbarao, K. V. 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. Proceedings of the National Academy of Sciences of the United States of America 109:964-975.

Dong, S., Raffaele, S., and Kamoun, S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Current Opinion in Genetics & Development 35:57-65.

Faino, L., Seidl, M. F., Shi-Kunne, X., Pauper, M., van den Berg, G. C. M., Wittenberg, A. H. J., and Thomma, B. P. H. J. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome Research 26:1091-1100.

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484:186-194.

Fones, H., and Gurr, S. 2015. The impact of Septoria tritici Blotch disease on wheat: An EU perspective. Fungal Genetics and Biology 79:3-7.

Fradin, E. F., and Thomma, B. P. H. J. 2006. Physiology and molecular aspects of Verticillium wilt diseases caused by *V. dahliae* and *V. albo-atrum*. Molecular Plant Pathology 7:71-86.

Fradin, E. F., Zhang, Z., Juarez Ayala, J. C., Castroverde, C. D., Nazar, R. N., Robb, J., Liu, C. M., and Thomma, B. P. 2009. Genetic dissection of Verticillium wilt resistance mediated by tomato Ve1. Plant Physiology 150:320-332.

Gibriel, H. A. Y., Thomma, B. P. H. J., and Seidl, M. F. 2016. The age of effectors: genome-based discovery and applications. Phytopathology 106:1206-1212.

Goodwin, S. B., Ben M'Barek, S., Dhillon, B., Wittenberg, A. H. J., Crane, C. F., Hane, J. K., Foster, A. J., Van der Lee, T. A. J., Grimwood, J., Aerts, A., Antoniw, J., Bailey, A., Bluhm, B., Bowler, J., Bristow, J., van der Burgt, A., Canto-CanchÈ, B., Churchill, A. C. L., Conde-Ferr‡ez, L., Cools, H. J., Coutinho, P. M., Csukai, M., Dehal, P., De Wit, P., Donzelli, B., van de Geest, H. C., van Ham, R. C. H. J., Hammond-Kosack, K. E., Henrissat, B., Kilian, A., Kobayashi, A. K., Koopmann, E., Kourmpetis, Y., Kuzniar, A., Lindquist, E., Lombard, V., Maliepaard, C., Martins, N., Mehrabi, R., Nap, J. P. H., Ponomarenko, A., Rudd, J. J., Salamov, A., Schmutz, J., Schouten, H. J., Shapiro, H., Stergiopoulos, I., Torriani, S. F. F., Tu, H., de Vries, R. P., Waalwijk, C., Ware, S. B., Wiebenga, A., Zwiers, L.-H., Oliver, R. P., Grigoriev, I. V., and Kema, G. H. J. 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLoS Genetics 7:e1002070.

Grandaubert, J., Bhattacharyya, A., and Stukenbrock, E. H. 2015. RNA-seq-Based gene annotation and comparative genomics of four fungal grass pathogens in the genus *Zymoseptoria* identify novel orphan genes and species-specific invasions of transposable elements. G3 5:1323-1333.

Hunter, T., Coker, RR., and Royle, DJ. 1999. The teleomorph stage, *Mycosphaerella graminicola*, in epidemics of septoria tritici blotch on winter wheat in the UK. Plant Pathology 48:51-57.

Inderbitzin, P., and Subbarao, K. V. 2014. *Verticillium* systematics and evolution: how confusion impedes Verticillium wilt management and how to resolve it. Phytopathology 104:564-574.

Inderbitzin, P., Bostock, R. M., Davis, R. M., Usami, T., Platt, H. W., and Subbarao, K. V. 2011. Phylogenetics and taxonomy of the fungal vascular wilt pathogen *Verticillium*, with the descriptions of five New species. PLoS One 6:e28341.

Jones, J. D. G., and Dangl, J. L. 2006. The plant immune system. Nature 444:323-329.

Kema, G. H. J., Annone, J. G., Sayoud, R., Van Silfhout, C. H., Van Ginkel, M., and De Bree, J. 1996. Genetic variation for virulence and resistance in the wheat-*Mycosphaerella graminicola* pathosystem. I. Interactions between pathogen isolates and host cultivars. Phytopathology 86, 200– 212.

Kema, G. H. J., and van Silfhout, C. H. 1997. Genetic variation for virulence and resistance in the wheat-*Mycosphaerella graminicola* pathosystem III. Comparative seedling and adult plant experiments. Phytopathology 87:266-272.

Kettles, G. J., and Kanyuka, K. 2016. Dissecting the molecular interactions between wheat and the fungal pathogen *Zymoseptoria tritici*. Frontiers in Plant Science 7: 508.

Klimes, A., Dobinson, K. F., Thomma, B. P., and Klosterman, S. J. 2015. Genomics spurs rapid advances in our understanding of the biology of vascular wilt pathogens in the genus *Verticillium*. Annual Reviews of Phytopathology 53:181-198.

Kombrink, A., Rovenich, H., Shi-Kunne, X., Rojas-Padilla, E., van den Berg, G. C. M., Domazakis, E., de Jonge, R., Valkenburg, D.-J., Sánchez-Vallet, A., Seidl, M. F., and Thomma, B. P. H. J. 2017. *Verticillium dahliae* LysM effectors differentially contribute to virulence on plant hosts. Molecular Plant Pathology. 18:596-608.

Marshall, R., Kombrink, A., Motteram, J., Loza-Reyes, E., Lucas, J., Hammond-Kosack, K. E., Thomma, B. P. H. J., and Rudd, J. J. 2011. 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 Physiology 156:756-769.

Metzker, M. L. 2010. Sequencing technologies - the next generation. Nature Reviews Genetics 11:31-46.

Mirzadi Gohari, A., Ware, S. B., Wittenberg, A. H. J., Mehrabi, R., Ben M'Barek, S., Verstappen, E. C. P., van der Lee, T. A. J., Robert, O., Schouten, H. J., de Wit, P. P. J. G. M., and Kema, G. H. J. 2015. Effector discovery in the fungal wheat pathogen *Zymoseptoria tritici*. Molecular Plant Pathology 16:931-945.

Plissonneau, C., Stürchler, A., and Croll, D. 2016. The evolution of orphan regions in genomes of a fungal pathogen of wheat. mBio 7: e01231-16.

Ponomarenko, A., Goodwin, S. B., and Kema, G. H. 2011. Septoria tritici blotch (STB) of wheat. Plant Health Instructor: 1-7.

Poppe, S., Dorsheimer, L., Happel, P., and Stukenbrock, E. H. 2015. Rapidly evolving genes are key players in host specialization and virulence of the fungal wheat pathogen *Zymoseptoria tritici* (*Mycosphaerella graminicola*). PLoS Pathogens 11:e1005055.

Raffaele, S., and Kamoun, S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nature Reviews Microbiology 10:417-430.

Rovenich, H., Boshoven, J. C., and Thomma, B. P. H. J. 2014. Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. Current Opinion in Plant Biology 20:96-103.

Rudd, J. J., Kanyuka, K., Hassani-Pak, K., Derbyshire, M., Andongabo, A., Devonshire, J., Lysenko, A., Saqi, M., Desai, N. M., Powers, S. J., Hooper, J., Ambroso, L., Bharti, A., Farmer, A., Hammond-Kosack, K. E., Dietrich, R. A., and Courbot, M. 2015. Transcriptome and metabolite profiling of the infection cycle of *Zymoseptoria tritici* on wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic

lifestyle definition. Plant Physiology 167:1158-1185.
Saintenac, C., Lee, W.-S., Cambon, F., Rudd, J. J., King, R. C., Marande, W., Powers, S. J., Bergès, H., Phillips, A. L., Uauy, C., Hammond-Kosack, K. E., Langin, T., and Kanyuka, K. 2018. Wheat receptor-kinase-like protein Stb6 controls gene-for-gene resistance to fungal pathogen *Zymoseptoria tritici*. Nature Genetics. 51:291.
Seidl, M. F., and Thomma, B. P. H. J. 2014. Sex or no sex: evolutionary adaptation occurs regardless. BioEssays 36:335-345.
Shiferaw, B., Smale, M., Braun, H.-J., Duveiller, E., Reynolds, M., and Muricho, G. 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Security 5:291-317.
Song, Y., Zhang, Z., Seidl, M. F., Majer, A., Jakse, J., Javornik, B., and Thomma, B. P. H. J. 2016. Broad taxonomic characterization of Verticillium wilt resistance genes reveals ancient origin of the tomato Ve1 immune receptor. Molecular Plant Pathology 19:77-89.
Stukenbrock, E. H., Jørgensen, F. G., Zala, M., Hansen, T. T., McDonald, B. A., and Schierup, M. H. 2010. Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen *Mycosphaerella graminicola*. PLoS Genetics 6:e1001189.
Thomma, B. P. H. J., Seidl, M. F., Shi-Kunne, X., Cook, D. E., Bolton, M. D., van Kan, J. A. L., and Faino, L. 2016. Mind the gap; seven reasons to close fragmented genome assemblies. Fungal Genetics and Biology 90:24-30.
Torriani, S. F., Melichar, J. P., Mills, C., Pain, N., Sierotzki, H., and Courbot, M. 2015. *Zymoseptoria tritici*: A major threat to wheat production, integrated approaches to control. Fungal Genetics Biology 79:8-12.
Wittenberg, A. H. J., van der Lee, T. A. J., Ben M'Barek, S., Ware, S. B., Goodwin, S. B., Kilian, A., Visser, R. G. F., Kema, G. H. J., and Schouten, H. J. 2009. Meiosis drives extraordinary genome plasticity in the haploid fungal plant pathogen *Mycosphaerella graminicola*. PLoS One 4:e5863.

Chapter 2

The age of effectors: genome-based discovery and applications

Hesham A.Y. Gibriel, Bart P.H.J. Thomma, Michael F. Seidl

Laboratory of Phytopathology, Wageningen University,
Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.

Published in Phytopathology 106:1206-1212 (2016)

ABSTRACT

Microbial pathogens cause devastating diseases on economically and ecologically important plant species, threatening global food security, and causing billions of dollars of losses annually. During the infection process, pathogens secrete so-called effectors that support host colonization, often by deregulating host immune responses. Over the last decades, much of the research on molecular plant-microbe interactions has focused on the identification and functional characterization of such effectors. The increasing availability of sequenced plant pathogen genomes has enabled genomics-based discovery of effector candidates. Nevertheless, identification of full plant pathogen effector repertoires is often hampered by erroneous gene annotation and the localization of effector genes in genomic regions that are notoriously difficult to assemble. Here, we argue that recent advances in genome sequencing technologies, genome assembly, gene annotation, as well as effector identification methods hold promise to disclose complete and correct effector repertoires. This allows to exploit complete effector repertoires, and knowledge of their diversity within pathogen populations, to develop durable and sustainable resistance breeding strategies, disease control and management of plant pathogens.

INTRODUCTION

Plant-associated microbes, which include bacteria, fungi, oomycetes and nematodes, play important roles in agriculture and in natural ecosystems. Some of these microbes cause devastating diseases that lead to >40% yield losses on major crops (Fisher et al. 2012; Pennisi 2010), and therefore threaten global food security. To contain diseases caused by plant pathogens, farmers rely on multiple strategies including adaptation of cultural practices such as crop rotation, but also deployment of resistant crop varieties and the use of chemical control agents such as fungicides. Nevertheless, pathogens may evolve to rapidly overcome such control measures (McDonald and Linde 2002; Stukenbrock and McDonald 2008).

To establish a successful infection, plant pathogens need to overcome the plant immune system. Plant cells deploy extracellular and intracellular receptors to detect invasion patterns that indicate the presence or activity of pathogens to subsequently mount immune responses (Cook et al. 2015). In turn, pathogens secrete so-called effector proteins to deregulate these host responses (Cook et al. 2015; Rovenich et al. 2014). However, effectors themselves can become invasion patterns when they become recognized by plant immune receptors that trigger an immune response, leading to an incompatible interaction between the pathogen and the host (Cook et al. 2015). To avoid recognition by the plant immune system, pathogens in turn evolve novel effectors or modify or lose existing ones.

Plant pathogen genomes generally encode a plethora of effectors. Effector diversification is mediated by genomic mechanisms that range from single nucleotide polymorphisms to structural variations that can affect chromosomal shape and gene content (Dong et al. 2015; Raffaele and Kamoun 2012; Seidl and Thomma 2014). In the past, various methods have been used to identify effector repertoires, and to determine their diversity. Here, we discuss the recent progress in genomic-based methods for effector discovery, while focusing on effectors from plant pathogens with a biotrophic or hemi-biotrophic lifestyle. Furthermore, we highlight how effector discovery can contribute to the development of resistance breeding strategies and disease management.

GENOMICS-BASED EFFECTOR DISCOVERY

More than a decade ago, the genome of the first plant associated microbe, the plant pathogenic bacterium *Xylella fastidiosa*, was sequenced (Simpson et al. 2000). Five years later, the rice blast fungus *Magnaporthe oryzae* was the first eukaryotic plant pathogen for which a genome sequence became available (Dean et al. 2005), soon followed by the genome sequences of two oomycete plant pathogens; *Phytophthora sojae* and *P. ramorum* (Tyler et al. 2006). All these genomes were sequenced using Sanger sequencing, which is laborious, costly, and low in throughput (Kircher and Kelso

2010). Subsequent advances in high throughput sequencing technologies, termed next-generation sequencing (NGS), have allowed rapid sequencing of entire genomes at significantly reduced costs (Metzker 2010). This development also brought genomic sciences within reach of research laboratories that study non-model organisms. Consequently, the number of available genome sequences has increased exponentially. Whereas researchers were fortunate to have access to a genome sequence of their organism of interest a decade ago, by now the availability of genome sequences for a multitude of genotypes of the same species is becoming the standard. Phytopathologists are especially interested in population genomic data which allows to determine core and lineage-specific effector repertoires of their pathogen of interest, allowing to investigate effector dynamics and evolution. However, reliable assessment of effector repertoires remains a considerable challenge.

Pathogens secrete effectors that exert their activity either in the host's extracellular environment (apoplastic effectors) or inside host cells (cytoplasmic effectors) (Giraldo and Valent 2013). Effectors are generally secreted via the endoplasmic reticulum ER/Golgi route, which requires a hydrophobic N-terminal signal peptide (von Heijne 1990). To date, most experimentally verified fungal and oomycete effectors contain such an N-terminal secretion signal (Lo Presti et al. 2015; Stergiopoulos and de Wit 2009). Therefore, genome-wide studies aiming to identify effector candidates commonly first predict the secretome, i.e. the repertoire of all secreted proteins, by querying for the presence of N-terminal signal peptides (Lo Presti et al. 2015).

Upon secretion into the extracellular space, effectors may subsequently be translocated into the host cytoplasm. In oomycetes, so far two distinct classes of cytoplasmic effectors have been described, the RXLRs and the Crinklers (CRNs). Both contain conserved amino acid motifs in the N-terminal region downstream of the signal peptide, which have been proposed to facilitate host-cell translocation (Schornack et al. 2010; Whisson et al. 2007). RXLR effectors contain an RXLR motif that is often followed by a (D)EER motif, while CRNs contain an LXLFLAK motif (Haas et al. 2009; Torto et al. 2003; Whisson et al. 2007). The presence of these motifs can be identified using computational methods (Haas et al. 2009; Jiang et al. 2008).

Most fungal and oomycete effectors that have been functionally characterized to date are smaller than 300 amino acids with four or more cysteine residues (Stergiopoulos et al. 2013; Stergiopoulos and de Wit 2009). Importantly, however, these characteristics cannot be used as criteria to identify effectors, as also well-characterized effectors often lack these properties (Lo Presti et al. 2015; Sperschneider et al. 2015). Therefore, to exhaustingly describe effector repertoires of plant pathogens, universal features of effector candidates, which are their secretion and differential expression *in planta*, should be considered (Sperschneider et al. 2015). Automatic classification methods that do not solely rely on strict rules can rank proteins according to their likelihood of being effectors by integrating multiple effector properties, e.g. protein size and cysteine

content, and thus can further aid effector identification (Saunders et al. 2012). Recently, 'EffectorP' has been introduced which is the first program that utilizes machine learning to improve the prediction of fungal effector proteins from secretomes (Sperschneider et al. 2016). 'EffectorP' uses sequence-derived features to predict fungal effectors from secretomes (Sperschneider et al. 2016), thereby limiting the direct dependency on the above mentioned rules of effector selection. However, fungal effectors are a very heterogeneous group of proteins and, even though novel approaches such as 'EffectorP' hold promises in improving effector discovery, effector proteins remain notoriously difficult to predict.

Comparative genomics-based effector discovery

The availability of genome sequences of closely related pathogens, or of different strains of the same pathogenic species, permit comparative studies to identify effectors and to determine core and lineage-specific effector repertoires of the pathogen of interest, allowing to investigate effector dynamics and evolution. Genome comparisons of the closely related fungal maize pathogens *Ustilago maydis* and *Sporisorium reilianum* revealed 43 genomic regions with decreased sequence conservation, some of which were completely absent in one of the species (Schirawski et al. 2010). Notably, these divergent regions harbour *in planta* induced genes encoding secreted proteins, likely representing effectors (Kamper et al. 2006; Schirawski et al. 2010). Deletion analyses in these effector-rich regions confirmed their contribution to pathogen virulence (Kamper et al. 2006; Schirawski et al. 2010), thereby highlighting the merit of comparative genomics to flag genes as candidate effectors.

Comparisons between related pathogens can also provide insights into core and lineage-specific effector repertoires. For example, genome comparison of multiple strains of the vascular wilt pathogen *Verticillium dahliae* identified large lineage-specific genomic regions, which are regions that are only present in a subset of the *V. dahliae* population (de Jonge et al. 2013). These regions are enriched for secreted, *in planta* induced effector candidates that contribute to plant colonization (de Jonge et al. 2013). Genome comparisons of 65 geographically diverse strains of the cassava bacterial blight pathogen *Xanthomonas axonopodis* pv. *manihotis* revealed multiple lineage-specific effector candidates (Bart et al. 2012). Moreover, multiple effectors were found to be highly conserved and have been maintained throughout the world in all field samples (Bart et al. 2012).

Effector genes in plant pathogens such as *Leptosphaeria maculans* and *Fusarium oxysporum* can also be located on distinct chromosomes that are under specific conditions not required for growth, and therefore have been referred to as conditionally dispensable chromosomes (CDCs) (Balesdent et al. 2013; Ma et al. 2010). Comparisons between CDCs of related pathogens have been shown to facilitate effector discovery (Williams et al. 2016). For example, comparisons between CDCs from legume–infecting *Fusarium oxysporum* f. sp. *medicaginis* with CDCs *of F. oxysporum* f. sp. *pisi* and *F.*

oxysporum f. sp. *ciceris* revealed small conserved genomic regions that contain *in planta* expressed genes encoding secreted proteins, which likely represent conserved effectors with roles in pathogenicity on legume hosts (Williams et al. 2016).

COMPARATIVE POPULATION GENOMICS- AND TRANSCRIPTOMICS-BASED DISCOVERY OF AVIRULENCE EFFECTOR GENES

Genome comparisons can also be utilized to identify effectors that are recognized by the plant immune system, so called avirulence effectors. Here, comparisons of avirulent and virulent strains can lead to the identification of single nucleotide polymorphisms (SNPs) in the protein-coding regions of effectors or reveal presence/absence polymorphisms, i.e. genomic regions that are exclusive to avirulent or virulent strains. For example, comparative genomics of multiple virulent and avirulent *V. dahliae* strains uncovered a 50-kb region encoding a single highly expressed effector gene that is present in avirulent strains but absent in virulent strains (de Jonge et al. 2012). This effector was subsequently experimentally shown to be the virulence factor Ave1 that is recognized by *Ve1*-carrying tomato plants (de Jonge et al. 2012). In the melon wilt fungus *F. oxysporum* f. sp. *melonis*, the effector AVRFOM2, which is recognized by melon Fom-2, was recently identified by comparative genomics of virulent and avirulent *F. oxysporum* f. sp. *melonis* strains (Schmidt et al. 2016). Similar to comparative analyses that exploit genome sequences, the sequencing and subsequent analyses of transcriptomes can lead to the identification of effectors. For example, combining *in silico* predictions with transcriptome comparisons of virulent and avirulent *Cladosporium fulvum* strains led to the identification of two polymorphic effector candidates, one of which was later confirmed to encode Avr5, which is recognized by the tomato immune receptor Cf5 (Mesarich et al. 2014).

ERRONEOUS GENOME ASSEMBLY AND ANNOTATION HAMPER EFFECTOR DISCOVERY

Nowadays, genome sequencing projects that utilize second-generation sequencing technologies generate millions of short (<500 bp) sequence fragments, so-called reads, of relatively high quality (1 sequence error per 1 kb) (Metzker 2010). However, the assembly of these reads into contiguous sequences, ideally chromosomes, is significantly hampered by highly similar repetitive sequences such as transposable elements (TEs). Sequencing reads that are shorter than repetitive sequences lead to their collapse, or to fragmented genome assemblies. Moreover, repetitive sequences in ascomycetes are often targeted by a genome-defense mechanism that induces specific C to T mutations, called repeat induced point mutations (RIP) (Selker 2002). RIP thereby locally elevates the AT content, which is a known bias for many second-generation technologies leading to low sequencing yields in these regions which further impact

genome assemblies (Dohm et al. 2008). Notably, in many plant pathogens, effector genes reside in repeat-rich genomic sub-compartments where repeats facilitate rapid effector diversification by promoting structural variations (Dong et al. 2015; Raffaele and Kamoun 2012; Seidl and Thomma 2014). Moreover, effectors are often flanked by repetitive elements that can be targeted by RIP, which can lead to rapid effector gene diversification as for example demonstrated in *L. maculans* (Gout et al. 2007). Based on the limitations inherent to second-generation sequencing technologies, effector-rich regions are often not completely assembled, thereby hampering the complete and accurate identification of effector repertoires and their diversity (Figure 1A) (Thomma et al. 2016).

Advances in sequencing technologies have opened possibilities to overcome the limitations of short-read sequencing leading to significant improvement in genome assemblies. Third-generation sequencing technologies such as Single-Molecule Real-Time (SMRT) and Nanopore sequencing generate long sequencing reads (>1 kb) that can span entire repetitive sequences and therefore enable more contiguous genome assemblies (Ashton et al. 2015; Kim et al. 2014; Powers et al. 2013; Shin et al. 2013). However, reads from third generation sequencing platforms display a relatively high error rate (~15%). To correct these sequencing errors, two alternative approaches are commonly applied. Hybrid error corrections such as implemented in Nanocorr (Goodwin et al. 2015) or PacBio Corrected Reads (PBcR) use high-quality short reads aligned to third-generation long reads to correct for sequencing errors (Koren et al. 2012). Alternatively, errors in long reads can be faithfully corrected by consensus-based approaches such as provided by HGAP (Chin et al. 2013) or Nanocorrect (Loman et al. 2015), if sequencing is performed at considerable depth (≥50X genomic coverage) (Faino et al. 2015). The costs involved in third-generation sequencing, in particular for large genomes, make this technology not yet widely adopted. Therefore, with the expected advances in sequencing technologies, i.e. longer reads with lower error rate and reduction in costs, sequencing and assembly of larger and complex (repeat-rich) genomes will become feasible. Sequence contiguity can be further increased by the application of optical mapping, which is a technique that constructs high-resolution restriction maps from a single DNA molecule (Levy-Sakin and Ebenstein 2013). Subsequently, the optical map is aligned to an *in silico* restriction map of the genome assembly, thereby facilitating the placement of assembled sequences into contiguous regions. This technology has been successfully applied for the genome assemblies of multiple plant pathogens such as *Botrytis cinerea*, *F. oxysporum*, *P. infestans* and *V. dahliae* (Amselem et al. 2011; Haas et al. 2009; Klosterman et al. 2011; Ma et al. 2010). Recently, the combination of long-read sequencing using SMRT sequencing technology and optical mapping yielded the complete and gapless genome assemblies of two *V. dahliae* strains (Faino et al. 2015). Significantly, comparison to previous genome assemblies revealed a 3-fold increase in repetitive elements (~12%) compared to previous estimate of ~4% (de Jonge et al. 2013; Faino et al. 2015; Klosterman et al. 2011; Seidl et al. 2015), indicating that repeats in the previous genome assemblies have been under-represented. Therefore, complete

genome assemblies mark a crucial step towards the optimal prediction of full effector repertoires in repeat-rich regions.

Apart from fragmented genome assemblies, incorrect gene annotation can hamper effector discovery (Figure 1B). Many of the previously annotated gene models in the fungal wheat pathogen *Z. tritici* have incorrect exon/intron boundaries and about 735 lack either start or stop codons (Gibriel et al. unpublished), which likely compromised the study of effectors candidates that had been selected based on these erroneous gene models (Mirzadi Gohari et al. 2015). Read data obtained from NGS-based RNA sequencing (RNA-seq) can be mapped to the reference genome, which provides a valuable guide for gene annotation as it reveals the exon-intron boundaries, and correct transcription start/stop sites (Figure 1B) (Yandell and Ence 2012). To achieve a high-quality gene prediction, *ab initio* or alignment based gene predictions can be supplemented with RNA-seq data (Faino and Thomma 2014). There has been a steady progress in the suite of gene-prediction tools that combine RNA-seq data with other lines of evidence including Maker2 (Holt and Yandell 2011), EVidenceModeler (EVM) (Haas et al. 2008), and BRAKER1 (Hoff et al. 2015). By combining *ab initio* gene predictions with RNA-seq supported data, the previous gene annotation of *Z. tritici* was recently re-assessed, leading to the correction of 1,555 gene models (Grandaubert et al. 2015). Next to transcriptomics, also proteomics can be used to validate and correct gene annotation, in particular of those genes encoding secreted proteins. For example, a large-scale proteomic study aiming to profile the extracellular proteome of *P. infestans* validated the extracellular nature of 254 proteins previously predicted to be secreted (Meijer et al. 2014). Moreover, this analyses also led to the correction of about 150 open reading frames in the *P. infestans* genome which were previously incorrectly annotated (Meijer et al. 2014).

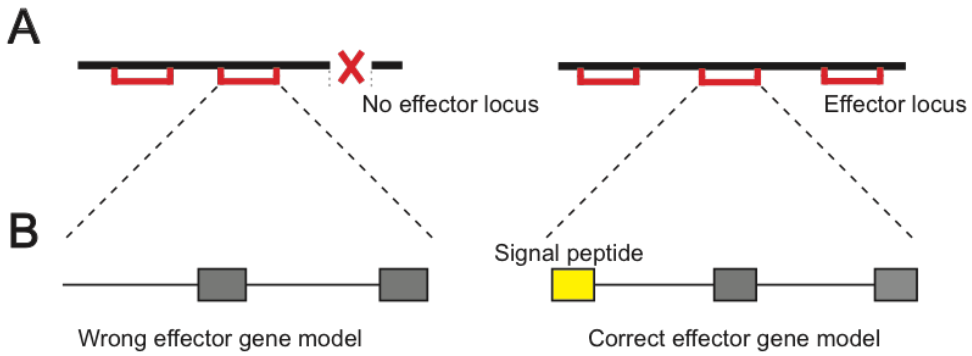


Figure 1. Erroneous genome assembly and annotation hamper effector discovery. Correct effector repertoires are obtained with proper assembly and annotation. **(A)** A stretch of DNA (black bar) that contains multiple effector genes (red blocks) is shown. In a fragmented genome assembly, an effector gene might be missing, as effectors are often located in

repeat-rich regions. In gapless genome assemblies, all effector genes are present. **(B)** Identification of effector genes in the predicted gene set is influenced by the quality of gene annotation. Applying effector selection criteria will miss or wrongly classify effector candidate in erroneous gene models that e.g. miss the correct start/stop codon. Mining of correctly annotated gene models with correct start/stop codon and signal peptides will yield full effector repertoires.

USING EFFECTORS AS A TOOL TO SCREEN FOR DISEASE RESISTANCE

Effector repertoires determined by genomics-based discoveries can be exploited to aid breeding for disease resistance. 'Effectoromics' is a large-scale screening approach that uses effector candidates to identify host resistance (*R*) genes (Vleeshouwers et al. 2008). This approach is based on transient expression of candidate effectors in plant leaves e.g. by agroinfiltration (*Agrobacterium tumefaciens* transient transformation assay) and/or a virus vector such as Potato virus X (PVX) (Vleeshouwers et al. 2008). Subsequently, plants are screened for the occurrence of hypersensitive cell death responses (HR), indicating the recognition of the effector by a matching plant immune receptor. 'Effectoromics' can contribute to resistance breeding by accelerating the identification of immune receptors when the matching effectors are available, as it replaces the slow process of generating stable transformants (Du and Vleeshouwers 2014; Vleeshouwers et al. 2008).

Effectors can target genes that critically facilitate compatible interactions between pathogens and their hosts, and thus these targets have often been referred to as susceptibility genes (van Schie and Takken 2014). Previously, susceptibility genes were identified by screening plant germplasm for recessive resistance (Bai et al. 2005), or by induced mutations that confer loss-of-susceptibility towards pathogens, which led to the identification of about 30 susceptibility genes (Gawehns et al. 2013). Alternatively, knowledge on the effector repertoire enables the usage of effectors as guides to identify plant susceptibility genes, for example by using protein-protein interaction assays such as yeast-two hybrid screening (Bos et al. 2010). Upon identification of a susceptibility gene, inactivation by subsequent mutations may confer disease resistance (Lewis et al. 2012), as this might interfere with the capability of the effectors to interact with their targets, and thereby limit the pathogen's ability to survive (Gawehns et al. 2013; van Schie and Takken 2014). Durability of resistance is a priority in plant breeding. One way to achieve durable resistance against plant pathogens is by the incorporation of multiple immune receptor genes into a single cultivar (gene pyramiding) as these can detect multiple effectors of the same pathogen (Michelmore et al. 2013). However, breeding multiple immune receptor genes into a cultivar is a time-consuming process and requires a sufficient number of previously identified receptor genes (Michelmore et al. 2013). Alternatively, susceptibility genes can provide more durable resistance than deployment of immune receptors, as they are essential for pathogen survival (van Schie

and Takken 2014). However, mutations in susceptibility genes often have pleiotropic effects such as dwarfism or sensitivity to other stresses, thereby limiting their utilization in agriculture (Gawehns et al. 2013).

Knowledge on effectors can also be used to develop protection strategies against pathogens, as silencing these effectors might lead to reduced pathogen development. Host-induced gene silencing (HIGS) is based on exploiting an RNA silencing mechanism to target a selected pathogen gene, for example an effector gene, via the host plant (Nowara et al. 2010). Plants can be transformed with hairpin RNA constructs that target selected pathogen effectors to suppress their expression. For example, HIGS in barley expressing dsRNA targeting the effector gene *Avra10* from the fungus *Blumeria graminis* led to reduced disease incidence (Nowara et al. 2010). To avoid pathogen evasion by loss of the targeted effector, conserved genes that are unlikely to be lost by the pathogen can be targeted. For example, HIGS in barley targeting the cytochrome P450 gene *CYP51* of the wheat pathogen *Fusarium graminearum* leads to complete plant immunity (Koch et al. 2013). Even though HIGS has the potential to become a powerful tool for disease control, its application requires the usage of genetically modified plants, which is restricted under current legislative and public consent in many countries worldwide.

MONITORING PATHOGEN FIELD POPULATIONS TO ASSESS EFFECTOR DIVERSITY

DNA and RNA sequencing technologies allow to monitor pathogen populations in fields, and thereby enable pathogen surveillance. For example, field transcriptome sequencing of *Puccinia striiformis* f. sp. *tritici* (PST) infected wheat revealed a dramatic shift in the pathogen population structure in the UK (Hubbard et al. 2015). Similarly, genome and transcriptome analysis of different strains of *P. infestans* provided insights into the emergence of the aggressive lineage 13_A2 in northwest Europe (Cooke et al. 2012). Genome sequencing enables researchers to rapidly assess effector diversity, which can indicate the emergence of new pathogens that have the potential to overcome host resistance in the field (Figure 2). Therefore, monitoring of effector diversity, and subsequent deployment of plants carrying the matching resistance gene, would greatly contribute to disease management. The application of sequencing technologies to monitor pathogen migration and diversity will likely further accelerate in the future. Rapid, nearly real-time genome sequencing of pathogen populations can be applied on field isolates using the MiniION system, which is a portable sequencer that sequences individual DNA molecules using Nanopore technology. MiniION can presently sequence bacterial and viral genomes in as little as 48 hours (Check Hayden 2015; Quick et al. 2015; Quick et al. 2016). For example, MinION was used for real-time detection of *Salmonella enterica* from clinical isolates during a hospital outbreak, and yielded reliable clinical information in less than half a day (Quick et al. 2015). Similarly, during the 2015 Ebola

virus outbreak in West Africa, real-time genomic sequencing revealed insights into virus diversity, and provided genomic information for epidemiological investigations to monitor disease dissemination (Quick et al. 2016). In the future, plant pathogen genomes can be rapidly sequenced with this device, disclosing information on their effector repertoires. Further fueled by technological advances of the MinION platform, we anticipate that real-time disease monitoring by genome sequencing in agricultural settings will provide unprecedented opportunities to monitor the spread of plant pathogens, and will be a prerequisite to install timely and suitable countermeasures.

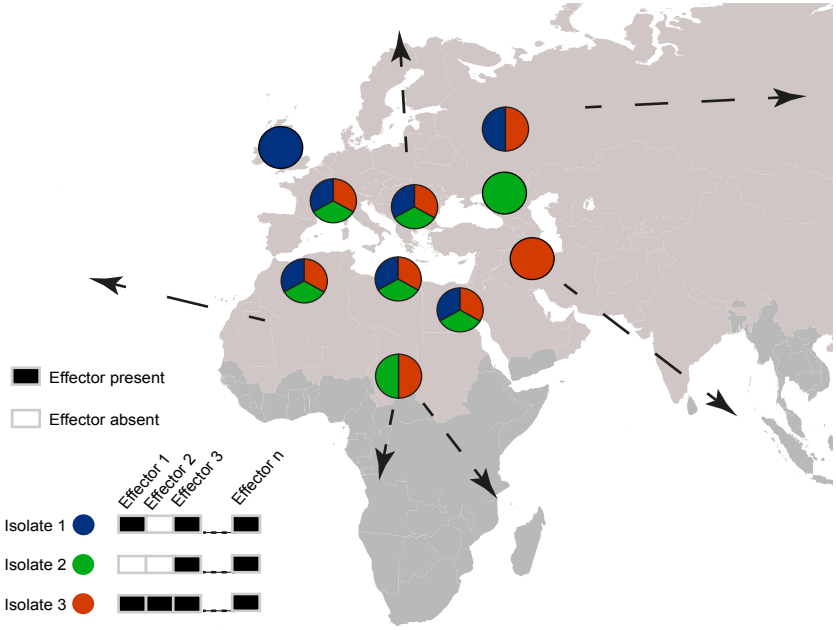


Figure 2. Monitoring of the spatial dispersal of plant pathogens. Rapid sequencing of plant pathogen genomes from field samples and subsequently identification of their effector repertories allows rapid insight into pathogen populations. Here, for example, a hypothetical pathogen population in several regions across Europe, Asia, and Africa, can be divided into three distinct lineages based on their effector repertoires, with the relative occurrence of each individual lineage within a geographic region shown by the pie chart. Sequencing of pathogen field samples and knowledge of their effector repertoires can be used to track the dispersal of pathogen lineages into new geographic regions (arrows).

ACKNOWLEDGMENTS

Work in the laboratory of BPHJT is support by the Research Council Earth and Life Sciences (ALW) of the Netherlands Organization of Scientific Research (NWO). MFS acknowledges the receipt of a VENI grant of ALW-NWO.

REFERENCES

Amselem, J., Cuomo, C. A., van Kan, J. A. L., Viaud, M., Benito, E. P., Couloux, A., Coutinho, P. M., de Vries, R. P., Dyer, P. S., Fillinger, S., Fournier, E., Gout, L., Hahn, M., Kohn, L., Lapalu, N., Plummer, K. M., Pradier, J.-M., Quévillon, E., Sharon, A., Simon, A., ten Have, A., Tudzynski, B., Tudzynski, P., Wincker, P., Andrew, M., Anthouard, V., Beever, R. E., Beffa, R., Benoit, I., Bouzid, O., Brault, B., Chen, Z., Choquer, M., Collémare, J., Cotton, P., Danchin, E. G., Da Silva, C., Gautier, A., Giraud, C., Giraud, T., Gonzalez, C., Grossetete, S., Güldener, U., Henrissat, B., Howlett, B. J., Kodira, C., Kretschmer, M., Lappartient, A., Leroch, M., Levis, C., Mauceli, E., Neuvéglise, C., Oeser, B., Pearson, M., Poulain, J., Poussereau, N., Quesneville, H., Rascle, C., Schumacher, J., Ségurens, B., Sexton, A., Silva, E., Sirven, C., Soanes, D. M., Talbot, N. J., Templeton, M., Yandava, C., Yarden, O., Zeng, Q., Rollins, J. A., Lebrun, M.-H., and Dickman, M. 2011. Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. PLoS Genetics 7:e1002230.

Ashton, P. M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., Wain, J., and O'Grady, J. 2015. MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. Nature Biotechnology 33:296-300.

Bai, Y., van der Hulst, R., Bonnema, G., Marcel, T. C., Meijer-Dekens, F., Niks, R. E., and Lindhout, P. 2005. Tomato defense to *Oidium neolycopersici*: dominant *Ol* genes confer isolate-dependent resistance via a different mechanism than recessive *ol-2*. Molecular Plant-Microbe Interaction 18:354-362.

Balesdent, M.-H., Fudal, I., Ollivier, B., Bally, P., Grandaubert, J., Eber, F., Chèvre, A.-M., Leflon, M., and Rouxel, T. 2013. The dispensable chromosome of *Leptosphaeria maculans* shelters an effector gene conferring avirulence towards *Brassica rapa*. New Phytologist 198:887-898.

Bart, R., Cohn, M., Kassen, A., McCallum, E. J., Shybut, M., Petriello, A., Krasileva, K., Dahlbeck, D., Medina, C., Alicai, T., Kumar, L., Moreira, L. M., Neto, J. R., Verdier, V., Santana, M. A., Kositcharoenkul, N., Vanderschuren, H., Gruissem, W., Bernal, A., and Staskawicz, B. J. 2012. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. Proceedings of the National Academy of Sciences of the United States of America 109:E1972-E1979.

Bos, J. I. B., Armstrong, M. R., Gilroy, E. M., Boevink, P. C., Hein, I., Taylor, R. M., Zhendong, T., Engelhardt, S., Vetukuri, R. R., Harrower, B., Dixelius, C., Bryan, G., Sadanandom, A., Whisson, S. C., Kamoun, S., and Birch, P. R. J. 2010. *Phytophthora infestans* effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. Proceedings of the National Academy of Sciences of the United States of America 107:9909-9914.

Check Hayden, E. 2015. Pint-sized DNA sequencer impresses first users. Nature 521:15-16.

Chin, C. S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E. E., Turner, S. W., and Korlach, J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nature Methods 10:563-569.

Cook, D. E., Mesarich, C. H., and Thomma, B. P. H. J. 2015. Understanding plant immunity as a surveillance system to detect invasion. Annual Review of Phytopathology 53:541-563.

Cooke, D. E. L., Cano, L. M., Raffaele, S., Bain, R. A., Cooke, L. R., Etherington, G. J., Deahl, K. L., Farrer, R. A., Gilroy, E. M., Goss, E. M., Grünwald, N. J., Hein, I., MacLean, D., McNicol, J. W., Randall, E., Oliva, R. F., Pel, M. A., Shaw, D. S., Squires, J. N., Taylor, M. C., Vleeshouwers, V. G. A. A., Birch, P. R. J., Lees, A. K., and Kamoun, S. 2012. Genome analyses of an aggressive and invasive lineage of the Irish potato famine pathogen. PLoS Pathogens 8:e1002940.

de Jonge, R., Bolton, M. D., Kombrink, A., van den Berg, G. C. M., Yadeta, K. A., and Thomma, B. P. H. J. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Research 23:1271-1282.

de Jonge, R., Peter van Esse, H., Maruthachalam, K., Bolton, M. D., Santhanam, P., Saber, M. K., Zhang, Z., Usami, T., Lievens, B., Subbarao, K. V., and Thomma, B. P. H. J. 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. Proceedings of the National Academy of Sciences of the United States of America 109:5110-5115.

Dean, R. A., Talbot, N. J., Ebbole, D. J., Farman, M. L., Mitchell, T. K., Orbach, M. J., Thon, M., Kulkarni, R., Xu, J.-R., Pan, H., Read, N. D., Lee, Y.-H., Carbone, I., Brown, D., Oh, Y. Y., Donofrio, N., Jeong, J. S., Soanes, D. M., Djonovic, S., Kolomiets, E., Rehmeyer, C., Li, W., Harding, M., Kim, S., Lebrun, M.-H., Bohnert, H., Coughlan, S., Butler, J., Calvo, S., Ma, L.-J., Nicol, R., Purcell, S., Nusbaum, C., Galagan, J. E., and Birren, B. W. 2005. The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434:980-986.

Dohm, J. C., Lottaz, C., Borodina, T., and Himmelbauer, H. 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Research 36:e105.

Dong, S., Raffaele, S., and Kamoun, S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Current Opinion in Genetics & Development 35:57-65.

Du, J., and Vleeshouwers, V. G. 2014. The do's and don'ts of effectoromics. Methods Molecular Biology 1127:257-268.

Faino, L., and Thomma, B. P. H. J. 2014. Get your high-quality low-cost genome sequence. Trends in Plant Science 19:288-291.

Faino, L., Seidl, M. F., Datema, E., van den Berg, G. C. M., Janssen, A., Wittenberg, A. H. J., and Thomma, B. P. H. J. 2015. Single-molecule real-time sequencing combined with optical mapping yields completely finished fungal genome. mBio 6:e00936-15.

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484:186-194.

Gawehns, F., Cornelissen, B. J. C., and Takken, F. L. W. 2013. The potential of effector-target genes in breeding for plant innate immunity. Microbial Biotechnology 6:223-229.

Giraldo, M. C., and Valent, B. 2013. Filamentous plant pathogen effectors in action. Nature Review Microbiology 11:800-814.

Goodwin, S., Gurtowski, J., Ethe-Sayers, S., Deshpande, P., Schatz, M. C., and McCombie, W. R. 2015. Oxford Nanopore sequencing, hybrid error correction, and *de novo* assembly of a eukaryotic genome. Genome Research 25:1750-1756.

Gout, L., Kuhn, M. L., Vincenot, L., Bernard-Samain, S., Cattolico, L., Barbetti, M., Moreno-Rico, O., Balesdent, M.-H., and Rouxel, T. 2007. Genome structure impacts molecular evolution at the AvrLm1 avirulence locus of the plant pathogen *Leptosphaeria maculans*. Environmental Microbiology 9:2978-2992.

Grandaubert, J., Bhattacharyya, A., and Stukenbrock, E. H. 2015. RNA-seq-based gene annotation and comparative genomics of four fungal grass pathogens in the genus *Zymoseptoria* identify novel orphan genes and species-specific invasions of transposable elements. G3 5:1323-1333.

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., White, O., Buell, C. R., and Wortman, J. R. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biology 9:1-22.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., Cakir, C., Carrington, J. C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M. A., Fugelstad, J., Gilroy, E. M., Gnerre, S., Green, P. J., Grenville-Briggs, L. J., Griffith, J., Grunwald, N. J., Horn, K., Horner, N. R., Hu, C. H., Huitema, E., Jeong, D. H., Jones, A. M., Jones, J. D., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzon, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., Savidor, A., Schornack, S., Schwartz, D. C., Schumann, U. D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D. J., Sykes, S., Thines, M., van de Vondervoort, P. J., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., and Stanke, M. 2015. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. Bioinformatics 32:767-769.

Holt, C., and Yandell, M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12:491.

Hubbard, A., Lewis, C., Yoshida, K., Ramirez-Gonzalez, R., de Vallavieille-Pope, C., Thomas, J., Kamoun, S., Bayles, R., Uauy, C., and Saunders, D. 2015. Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. Genome Biology 16:23.

Jiang, R. H. Y., Tripathy, S., Govers, F., and Tyler, B. M. 2008. RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. Proceedings of the National Academy of Sciences of the United States of America 105:4874-4879.

Kamper, J., Kahmann, R., Bolker, M., Ma, L.-J., Brefort, T., Saville, B. J., Banuett, F., Kronstad, J. W., Gold, S. E., Muller, O., Perlin, M. H., Wosten, H. A. B., de Vries, R., Ruiz-Herrera, J., Reynaga-Pena, C. G., Snetselaar, K., McCann, M., Perez-Martin, J., Feldbrugge, M., Basse, C. W., Steinberg, G., Ibeas, J. I., Holloman, W., Guzman, P., Farman, M., Stajich, J. E., Sentandreu, R., Gonzalez-Prieto, J. M., Kennell, J. C., Molina, L., Schirawski, J., Mendoza-Mendoza, A., Greilinger, D., Munch, K., Rossel, N., Scherer, M., Vranes, M., Ladendorf, O., Vincon, V., Fuchs, U., Sandrock, B., Meng, S., Ho, E. C. H., Cahill, M. J., Boyce, K. J., Klose, J., Klosterman, S. J., Deelstra, H. J., Ortiz-Castellanos, L., Li, W., Sanchez-Alonso, P., Schreier, P. H., Hauser-Hahn, I., Vaupel, M., Koopmann, E., Friedrich, G., Voss, H., Schluter, T., Margolis, J., Platt, D., Swimmer, C., Gnirke, A., Chen, F., Vysotskaia, V., Mannhaupt, G., Guldener, U., Munsterkotter, M., Haase, D., Oesterheld, M., Mewes, H.-W., Mauceli, E. W., DeCaprio, D., Wade, C. M., Butler, J., Young, S., Jaffe, D. B., Calvo, S., Nusbaum, C., Galagan, J., and Birren, B. W. 2006. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. Nature 444:97-101.

Kim, K. E., Peluso, P., Babayan, P., Yeadon, P. J., Yu, C., Fisher, W. W., Chin, C.-S., Rapicavoli, N. A., Rank, D. R., Li, J., Catcheside, D. E. A., Celniker, S. E., Phillippy, A. M., Bergman, C. M., and Landolin, J. M. 2014. Long-read, whole-genome shotgun sequence data for five model organisms. Scientific Data 1:140045.

Kircher, M., and Kelso, J. 2010. High-throughput DNA sequencing – concepts and limitations. BioEssays 32:524-536.

Klosterman, S. J., Subbarao, K. V., Kang, S., Veronese, P., Gold, S. E., Thomma, B. P., Chen, Z., Henrissat, B., Lee, Y. H., Park, J., Garcia-Pedrajas, M. D., Barbara, D. J., Anchieta, A., de Jonge, R., Santhanam, P., Maruthachalam, K., Atallah, Z., Amyotte, S. G., Paz, Z., Inderbitzin, P., Hayes, R. J., Heiman, D. I., Young, S., Zeng, Q., Engels, R., Galagan, J., Cuomo, C. A., Dobinson, K. F., and Ma, L. J. 2011. Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. PLoS Pathogens 7:e1002137.

Koch, A., Kumar, N., Weber, L., Keller, H., Imani, J., and Kogel, K.-H. 2013. Host-induced gene silencing of cytochrome P450 lanosterol C14α-demethylase–encoding genes confers strong resistance to *Fusarium* species. Proceedings of the National Academy of Sciences of the United States of America 110:19324-19329.

Koren, S., Schatz, M. C., Walenz, B. P., Martin, J., Howard, J. T., Ganapathy, G., Wang, Z., Rasko, D. A., McCombie, W. R., Jarvis, E. D., and Phillippy, A. M. 2012. Hybrid error correction and *de novo* assembly of single-molecule sequencing reads. Nature Biotechnology 30:693-700.

Levy-Sakin, M., and Ebenstein, Y. 2013. Beyond sequencing: optical mapping of DNA in the age of nanotechnology and nanoscopy. Current Opinion in Biotechnology 24:690-698.

Lewis, J. D., Wan, J., Ford, R., Gong, Y., Fung, P., Nahal, H., Wang, P. W., Desveaux, D., and Guttman, D. S. 2012. Quantitative interactor screening with next-generation sequencing (QIS-Seq) identifies *Arabidopsis thaliana* MLO2 as a target of the *Pseudomonas syringae* type III effector HopZ2. BMC Genomics 13:8-8.

Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., and Kahmann, R. 2015. Fungal effectors and plant susceptibility. Annual Review of Plant Biology 66:513-545.

Loman, N. J., Quick, J., and Simpson, J. T. 2015. A complete bacterial genome assembled *de novo* using only nanopore sequencing data. Nature Methods 12:733-735.

Ma, L.-J., van der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M.-J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., Houterman, P. M., Kang, S., Shim, W.-B., Woloshuk, C., Xie, X., Xu, J.-R., Antoniw, J., Baker, S. E., Bluhm, B. H., Breakspear, A., Brown, D. W., Butchko, R. A. E., Chapman, S., Coulson, R., Coutinho, P. M., Danchin, E. G. J., Diener, A., Gale, L. R., Gardiner, D. M., Goff, S., Hammond-Kosack, K. E., Hilburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C. D., Koehrsen, M., Kumar, L., Lee, Y.-H., Li, L., Manners, J. M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S.-Y., Proctor, R. H., Regev, A., Ruiz-Roldan, M. C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D. C., Turgeon, B. G., Wapinski, I., Yoder, O., Young, S., Zeng, Q., Zhou, S., Galagan, J., Cuomo, C. A., Kistler, H. C., and Rep, M. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367-373.

McDonald, B. A., and Linde, C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annual Review of Phytopathology 40:349-379.

Meijer, H. J. G., Mancuso, F. M., Espadas, G., Seidl, M. F., Chiva, C., Govers, F., and Sabidó, E. 2014. Profiling the secretome and extracellular proteome of the potato late blight pathogen *Phytophthora infestans*. Molecular & Cellular Proteomics 13:2101-2113.

Mesarich, C. H., Griffiths, S. A., van der Burgt, A., Ökmen, B., Beenen, H. G., Etalo, D. W., Joosten, M. H. A. J., and de Wit, P. J. G. M. 2014. Transcriptome sequencing uncovers the *Avr5* avirulence gene of the tomato leaf mold pathogen *Cladosporium fulvum*. Molecular Plant-Microbe Interactions 27:846-857.

Metzker, M. L. 2010. Sequencing technologies - the next generation. Nature Review Genetics 11:31-46.

Michelmore, R. W., Christopoulou, M., and Caldwell, K. S. 2013. Impacts of resistance gene genetics, function, and evolution on a durable future. Annual Review of Phytopathology 51:291-319.

Mirzadi Gohari, A., Ware, S. B., Wittenberg, A. H. J., Mehrabi, R., Ben M'Barek, S., Verstappen, E. C. P., van der Lee, T. A. J., Robert, O., Schouten, H. J., de Wit, P. P. J. G. M., and Kema, G. H. J. 2015. Effector discovery in the fungal wheat pathogen *Zymoseptoria tritici*. Molecular Plant Pathology 16:931-945.

Nowara, D., Gay, A., Lacomme, C., Shaw, J., Ridout, C., Douchkov, D., Hensel, G., Kumlehn, J., and Schweizer, P. 2010. HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen *Blumeria graminis*. The Plant Cell 22:3130-3141.

Pennisi, E. 2010. Armed and dangerous. Science 327:804-805.

Powers, J. G., Weigman, V. J., Shu, J., Pufky, J. M., Cox, D., and Hurban, P. 2013. Efficient and accurate whole genome assembly and methylome profiling of *E. coli*. BMC Genomics 14:1-18.

Quick, J., Ashton, P., Calus, S., Chatt, C., Gossain, S., Hawker, J., Nair, S., Neal, K., Nye, K., Peters, T., De Pinna, E., Robinson, E., Struthers, K., Webber, M., Catto, A., Dallman, T., Hawkey, P., and Loman, N. 2015. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of *Salmonella*. Genome Biology 16:114.

Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., Bore, J. A., Koundouno, R., Dudas, G., Mikhail, A., Ouédraogo, N., Afrough, B., Bah, A., Baum, J. H. J., Becker-Ziaja, B., Boettcher, J. P., Cabeza-Cabrerizo, M., Camino-Sánchez, Á., Carter, L. L., Doerrbecker, J., Enkirch, T., Dorival, I. G., Hetzelt, N., Hinzmann, J., Holm, T., Kafetzopoulou, L. E., Koropogui, M., Kosgey, A., Kuisma, E., Logue, C. H., Mazzarelli, A., Meisel, S., Mertens, M., Michel, J., Ngabo, D., Nitzsche, K., Pallasch, E., Patrono, L. V., Portmann, J., Repits, J. G., Rickett, N. Y., Sachse, A., Singethan, K., Vitoriano, I., Yemanaberhan, R. L., Zekeng, E. G., Racine, T., Bello, A., Sall, A. A., Faye, O., Faye, O., Magassouba, N. F., Williams, C. V., Amburgey, V., Winona, L., Davis, E., Gerlach, J., Washington, F., Monteil, V., Jourdain, M., Bererd, M., Camara, A., Somlare, H., Camara, A., Gerard, M., Bado, G., Baillet, B., Delaune, D., Nebie, K. Y., Diarra, A., Savane, Y., Pallawo, R. B., Gutierrez, G. J., Milhano, N., Roger, I., Williams, C. J., Yattara, F., Lewandowski, K., Taylor, J., Rachwal, P., J. Turner, D., Pollakis, G., Hiscox, J. A., Matthews, D. A., Shea, M. K. O., Johnston, A. M., Wilson, D., Hutley, E., Smit, E., Di Caro, A., Wölfel, R., Stoecker, K., Fleischmann, E., Gabriel, M., Weller, S. A., Koivogui, L., Diallo, B., Keïta, S., Rambaut, A., Formenty, P., Günther, S., and Carroll, M. W. 2016. Real-time, portable genome sequencing for *Ebola* surveillance. Nature 530:228-232.

Raffaele, S., and Kamoun, S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nature Reviews Microbiology 10:417-430.

Rovenich, H., Boshoven, J. C., and Thomma, B. P. H. J. 2014. Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. Current Opinion in Plant Biology 20:96-103.

Saunders, D. G. O., Win, J., Cano, L. M., Szabo, L. J., Kamoun, S., and Raffaele, S. 2012. Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. PLoS One 7:e29847.

Schirawski, J., Mannhaupt, G., Munch, K., Brefort, T., Schipper, K., Doehlemann, G., Di Stasio, M., Rossel, N., Mendoza-Mendoza, A., Pester, D., Muller, O., Winterberg, B., Meyer, E., Ghareeb, H., Wollenberg, T., Munsterkotter, M., Wong, P., Walter, M., Stukenbrock, E., Guldener, U., and Kahmann, R. 2010. Pathogenicity determinants in smut fungi revealed by genome comparison. Science 330:1546-1548.

Schmidt, S. M., Lukasiewicz, J., Farrer, R., van Dam, P., Bertoldo, C., and Rep, M. C. 2016. Comparative genomics of *Fusarium oxysporum* f. sp. *melonis* reveals the secreted protein recognized by the Fom-2 resistance gene in melon. New Phytologist 209:307-318.

Schornack, S., van Damme, M., Bozkurt, T. O., Cano, L. M., Smoker, M., Thines, M., Gaulin, E., Kamoun, S., and Huitema, E. 2010. Ancient class of translocated oomycete effectors targets the host nucleus. Proceedings of the National Academy of Sciences of the United States of America 107:17421-17426.

Seidl, M. F., and Thomma, B. P. H. J. 2014. Sex or no sex: Evolutionary adaptation occurs regardless. BioEssays 36:335-345.

Seidl, M. F., Faino, L., Shi-Kunne, X., van den Berg, G. C., Bolton, M. D., and Thomma, B. P. H. J. 2015. The genome of the saprophytic fungus *Verticillium tricorpus* reveals a complex effector repertoire resembling that of its pathogenic relatives. Molecular Plant-Microbe Interactions 28:362-373.

Selker, E. U. 2002. Repeat-induced gene silencing in fungi. Advances in Genetics 46:439-450.

Shin, S. C., Ahn, D. H., Kim, S. J., Lee, H., Oh, T.-J., Lee, J. E., and Park, H. 2013. Advantages of single-molecule real-time sequencing in high-GC content genomes. PLoS One 8:e68824.

Simpson, A. J., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bove, J. M., Briones, M. R., Bueno, M. R., Camargo, A. A., Camargo, L. E., Carraro, D. M., Carrer, H., Colauto, N. B., Colombo, C., Costa, F. F., Costa, M. C., Costa-Neto, C. M., Coutinho, L. L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A. P., Ferreira, A. J., Ferreira, V. C., Ferro, J. A., Fraga, J. S., Franca, S. C., Franco, M. C., Frohme, M., Furlan, L. R., Garnier, M., Goldman, G. H., Goldman, M. H., Gomes, S. L., Gruber, A., Ho, P. L., Hoheisel, J. D., Junqueira, M. L., Kemper, E. L., Kitajima, J. P., Krieger, J. E., Kuramae, E. E., Laigret, F., Lambais, M. R., Leite, L. C., Lemos, E. G., Lemos, M. V., Lopes, S. A., Lopes, C. R., Machado, J. A., Machado, M. A., Madeira, A. M., Madeira, H. M., Marino, C. L., Marques, M. V., Martins, E. A., Martins, E. M., Matsukuma, A. Y., Menck, C. F., Miracca, E. C., Miyaki, C. Y., Monteriro-Vitorello, C. B., Moon, D. H., Nagai, M. A., Nascimento, A. L., Netto, L. E., Nhani, A., Jr., Nobrega, F. G., Nunes, L. R., Oliveira, M. A., de Oliveira, M. C., de Oliveira, R. C., Palmieri, D. A., Paris, A., Peixoto, B. R., Pereira, G. A., Pereira, H. A., Jr., Pesquero, J. B., Quaggio, R. B., Roberto, P. G., Rodrigues, V., de, M. R. A. J., de Rosa, V. E., Jr., de Sa, R. G., Santelli, R. V., Sawasaki, H. E., da Silva, A. C., da Silva, A. M., da Silva, F. R., da Silva, W. A., Jr., da Silveira, J. F., Silvestri, M. L., Siqueira, W. J., de Souza, A. A., de Souza, A. P., Terenzi, M. F., Truffi, D., Tsai, S. M., Tsuhako, M. H., Vallada, H., Van Sluys,

M. A., Verjovski-Almeida, S., Vettore, A. L., Zago, M. A., Zatz, M., Meidanis, J., and Setubal, J. C. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* consortium of the organization for nucleotide sequencing and analysis. Nature 406:151-159.

Sperschneider, J., Dodds, P. N., Gardiner, D. M., Manners, J. M., Singh, K. B., and Taylor, J. M. 2015. Advances and challenges in computational prediction of effectors from plant pathogenic fungi. PLoS Pathogens 11:e1004806.

Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., Manners, J. M., and Taylor, J. M. 2016. EffectorP: predicting fungal effector proteins from secretomes using machine learning. New Phytologist 210:743–761.

Stergiopoulos, I., and de Wit, P. 2009. Fungal effector proteins. Annual Review of Phytopathology 47:233-263.

Stergiopoulos, I., Collemare, J., Mehrabi, R., and De Wit, P. J. G. M. 2013. Phytotoxic secondary metabolites and peptides produced by plant pathogenic *Dothideomycete* fungi. FEMS Microbiology Reviews 37:67-93.

Stukenbrock, E. H., and McDonald, B. A. 2008. The origins of plant pathogens in agro-ecosystems. Annual Review of Phytopathology 46:75-100.

Thomma, B. P. H. J., Seidl, M. F., Shi-Kunne, X., Cook, D. E., Bolton, M. D., van Kan, J. A. L., and Faino, L. 2016. Mind the gap; seven reasons to close fragmented genome assemblies. Fungal Genetics and Biology 90:24-30.

Torto, T. A., Li, S., Styer, A., Huitema, E., Testa, A., Gow, N. A., van West, P., and Kamoun, S. 2003. EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. Genome Research 13:1675-1685.

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H., Aerts, A., Arredondo, F. D., Baxter, L., Bensasson, D., Beynon, J. L., Chapman, J., Damasceno, C. M., Dorrance, A. E., Dou, D., Dickerman, A. W., Dubchak, I. L., Garbelotto, M., Gijzen, M., Gordon, S. G., Govers, F., Grunwald, N. J., Huang, W., Ivors, K. L., Jones, R. W., Kamoun, S., Krampis, K., Lamour, K. H., Lee, M. K., McDonald, W. H., Medina, M., Meijer, H. J., Nordberg, E. K., Maclean, D. J., Ospina-Giraldo, M. D., Morris, P. F., Phuntumart, V., Putnam, N. H., Rash, S., Rose, J. K., Sakihama, Y., Salamov, A. A., Savidor, A., Scheuring, C. F., Smith, B. M., Sobral, B. W., Terry, A., Torto-Alalibo, T. A., Win, J., Xu, Z., Zhang, H., Grigoriev, I. V., Rokhsar, D. S., and Boore, J. L. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313:1261-1266.

van Schie, C. C., and Takken, F. L. 2014. Susceptibility genes 101: how to be a good host. Annual Review of Phytopathology 52:551-581.

Vleeshouwers, V. G. A. A., Rietman, H., Krenek, P., Champouret, N., Young, C., Oh, S.-K., Wang, M., Bouwmeester, K., Vosman, B., Visser, R. G. F., Jacobsen, E., Govers, F., Kamoun, S., and Van der Vossen, E. A. G. 2008. Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora Infestans* avirulence genes. PLoS One 3:e2875.

von Heijne, G. 1990. The signal peptide. The Journal of Membrane Biology 115:195-201.

Whisson, S. C., Boevink, P. C., Moleleki, L., Avrova, A. O., Morales, J. G., Gilroy, E. M., Armstrong, M. R., Grouffaud, S., van West, P., Chapman, S., Hein, I., Toth, I. K., Pritchard, L., and Birch, P. R. J. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. Nature 450:115-118.

Williams, A. H., Sharma, M., Thatcher, L. F., Azam, S., Hane, J. K., Sperschneider, J., Kidd, B. N., Anderson, J. P., Ghosh, R., Garg, G., Lichtenzveig, J., Kistler, H. C., Shea, T., Young, S., Buck, S.-A. G., Kamphuis, L. G., Saxena, R., Pande, S., Ma, L.-J., Varshney, R. K., and Singh, K. B. 2016. Comparative genomics and prediction of conditionally dispensable sequences in legume–infecting *Fusarium oxysporum formae speciales* facilitates identification of candidate effectors. BMC genomics 17:1-24.

Yandell, M., and Ence, D. 2012. A beginner's guide to eukaryotic genome annotation. Nature Reviews Genetics 13:329-342.

Chapter 3

Verticillium dahliae strains that infect the same host plant display highly divergent effector catalogs

Hesham A.Y. Gibriel[1,2,$], Jinling Li[1,$], Longfu Zhu[3], Michael F. Seidl[1,#], Bart P.H.J. Thomma[1,#]

[1] Laboratory of Phytopathology, Wageningen University & Research, 6708 PB Wageningen, The Netherlands.
[2] Current address: School of Agriculture and Food Science, University College Dublin, Dublin, Ireland.
[3] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei 430070, China.

[$] These authors contributed equally
[#] These authors also contributed equally to this work

To be submitted for publication

ABSTRACT

Effectors are proteins secreted by pathogens to support colonization of host plants, often by deregulating host immunity. Effector genes are often localized within dynamic lineage-specific (LS) genomic regions, allowing rapid evolution of effector catalogues. It is thought that such localization permits pathogens to be competitive in the co-evolutionary arms races with their hosts. For a broad host-range pathogen such as *Verticillium dahliae* it is unclear to what extent single members of their total effector repertoires contribute to disease development on multiple hosts. Here, we determined the core and LS effector repertoires of a collection of *V. dahliae* strains, as well as the ability of these strains to infect a range of plant species comprising tomato, cotton, *Nicotiana benthamiana*, Arabidopsis, and sunflower to assess whether the presence of particular LS effectors correlates with the ability to infect particular plant species. Surprisingly, we found that *V. dahliae* strains that are able to infect the same host plant harbor highly divergent LS effector repertoires. Furthermore, we observed differential *V. dahliae* core effector gene expression between host plants. Our data suggest that different *V. dahliae* lineages utilise divergent effector catalogs to colonize the same host plant, suggesting considerable redundancy among the activities of effector catalogs between lineages.

INTRODUCTION

Plant pathogens cause devastating diseases on crop plants, threatening food security worldwide (Fisher et al. 2012; Pennisi 2010). In order to establish their infection, pathogens secrete effector molecules that can modulate host physiology, often by deregulating host immune responses (Cook et al. 2015; Jones and Dangl 2006). However, in turn, effectors may become recognized by plant immune receptors, leading to the activation of immune responses and attempted arrest of pathogen invasion (Cook et al. 2015; Jones and Dangl 2006). Thus, pathogens need to continuously evolve their effector catalogues by modifying or purging existing effectors that became recognized, or by acquiring novel effectors to suppress effector-triggered immune responses (Cook et al. 2015; Jones and Dangl 2006).

Genomes of plant pathogens are often thought to have evolved two distinct compartments; one comprising gene-rich, repeat-poor genomic regions that contain core genes that mediate general physiology, and one comprising gene-poor, plastic, repeat-rich genomic regions that contain effector genes and other pathogenicity-related genes (Dong et al. 2015; Raffaele and Kamoun 2012). The plastic genomic regions are either embedded within core chromosomes or reside on separate chromosomes that are often referred to as conditionally dispensable chromosomes (CDCs) (Dong et al. 2015; Raffaele and Kamoun 2012). For instance, effector genes of the tomato-pathogen *Fusarium oxysporum* f. sp. *lycopersici*, known as secreted in xylem (*SIX*) genes, are located on dispensable chromosomes (Schmidt et al. 2013). In contrast, all known effectors of the fungal wheat pathogen *Zymoseptoria tritici* are located on core chromosomes, while no recognizable effector genes reside on dispensable chromosomes (Kema et al. 2018; Marshall et al. 2011; Meile et al. 2018). Core chromosomes also carry effector genes in other fungal plant pathogens, such as the fungal smut pathogen *Ustilago maydis* and the fungal tomato pathogen *Cladosporium fulvum* (Hemetsberger et al. 2015; Stergiopoulos et al. 2010). A genome compartmentalization with physically separated effector-containing regions is often referred to as a "two-speed" genome organization because it is thought that gene-rich, repeat-poor genomic regions evolve slowly, while gene-poor, repeat-rich genomic regions evolve quicker (Croll and McDonald 2012; Raffaele and Kamoun 2012). Accordingly, the occurrence of effector genes within plastic genomic regions allows rapid evolution of effector catalogues and permits pathogens to be competitive in the co-evolution with hosts and evade their immune systems (Dong et al. 2015; Raffaele and Kamoun 2012).

Verticillium dahliae is a soil-borne fungal plant pathogen that is able to infect a broad range of plant species, including crops such as tomato, potato, lettuce, and cotton (Fradin and Thomma 2006; Inderbitzin et al. 2011). The fungus infects plants through their roots and subsequently colonizes the water-conducting xylem vessels, leading to vascular wilt disease (Fradin and Thomma 2006). Comparative genomics between closely related *V. dahliae* strains revealed that they carry highly dynamic, repeat rich, lineage-specific (LS)

regions that are only present in a subset of *V. dahliae* strains, and that account for up to 4 Mb of the ~35 Mb genome (de Jonge et al. 2013; Faino et al. 2016). These LS regions are enriched for *in planta*-induced effector genes that contribute to fungal virulence (de Jonge et al. 2013). However, effector genes are not only found in LS regions, as also the core genome harbors effector genes, such as those encoding a family of necrosis and ethylene-inducing-like proteins (NLPs) some of which were found to induce cell death in dicotyledonous plants (de Jonge et al. 2011; Santhanam et al. 2013). Similarly, also a family of lysin motif (LysM) effectors is encoded in the core genome, various homologs of which have been reported to enhance virulence by suppression of chitin-triggered immunity in other fungal pathogens (de Jonge et al. 2010; Kombrink et al. 2017; Marshall et al. 2011; Mentlak et al. 2012; Takahara et al. 2016). However, only a single LysM effector that is encoded in an LS region of *V. dahliae* strain VdLs17, and is thus not identified in the genomes of other *V. dahliae* strains, was found to contribute to virulence by suppression of chitin-triggered immunity, whereas no role in virulence could be attributed to any of the core LysM effectors (Kombrink et al. 2017).

Whereas *V. dahliae* is characterized by its generally broad host range, differential pathogenicity among hosts occurs for individual strains (Bhat and Subbarao 1999). In this study, we analysed the genomes of a collection of *V. dahliae* strains and assessed their core and LS effector catalogues in relation to their host ranges. To this end, we selected a set of strains that are well-adapted to cause disease on tomato (*Solanum lycopersicum*), cotton (*Gossypium hirsutum*), Australian tobacco (*Nicotiana benthamiana*), Arabidopsis (*Arabidopsis thaliana*), and sunflower (*Helianthus annuus*) and determined their core and LS effector catalogues as well as their *in planta* expression profiles.

MATERIALS AND METHODS

V. dahliae strains and plant inoculations

In total, 21 *V. dahliae* strains collected at different geographical locations were used in this study (Table S1). All strains were grown on potato dextrose agar (PDA; Oxoid, Basingstoke, UK) at 22°C, and conidiospores were collected from 10-day-old plates and washed with tap water. Disease assays were performed on sunflower (*Helianthus annuus* L. cv. Tutti), cotton (*Gossypium hirsutum* cv. Simian 3), tomato (*Solanum lycopersicum* cv. Moneymaker), *Nicotiana benthamiana* and *Arabidopsis thaliana* (Col-0) plants using the root-dipping inoculation method as previously described (Fradin et al. 2009; Song et al. 2017). Briefly, two-week-old (Arabidopsis, tomato, cotton, sunflower) or three-week-old (*N. benthamiana*) seedlings were carefully uprooted and the roots were rinsed in water. Subsequently, the roots were dipped for eight minutes in a suspension of 10^6 conidiospores/mL of water. Control plants were treated similarly by root dipping in tap water without conidiospores. Disease symptoms were scored up to 21 (tomato, *N. benthamiana*, Arabidopsis), 28 (cotton) or 45 (sunflower) days post inoculation (dpi).

Genome sequencing and assembly

The genome sequences of 13 *V. dahliae* strains were previously determined, nine of which were previously sequenced using Illumina HiSeq 2000 (de Jonge et al. 2013; de Jonge et al. 2012) and four that were sequenced using long-read PacBio Single-Molecule Real-Time (SMRT) sequencing technology (Table S1) (Faino et al. 2015). Additionally, eight *V. dahliae* strains (Table S1) were newly sequenced in this study. To this end, genomic DNA of *V. dahliae* strains was obtained from conidiospores that were harvested from ten-day-old cultures grown on potato dextrose agar as described previously (de Jonge et al. 2012). Library preparation (~500 bp insert size) and genomic sequencing (100 bp paired-end reads) were performed at the Beijing Genome Institute (BGI, Hong Kong). All *V. dahliae* strains that were sequenced with short-read sequencing technology (17 strains) were assembled with the A5 pipeline (default parameters) that automates data cleaning, error correction, assembly, and quality control (Tritt et al. 2012). Genome assembly statistics for all 21 *V. dahliae* strains were calculated using QUAST (Gurevich et al. 2013). Repeats were identified using RepeatModeler (version 1.0.8) (default parameters) (Smit and Hubley 2010) and masked using RepeatMasker (version 4.0.7) (default parameters) (Smit et al. 2016).

Gene prediction and annotation

Previously generated gene annotations of *V. dahliae* strains JR2 (Faino et al. 2015) and CQ2 (unpublished data) were used in this study. For the remaining 19 *V. dahliae* strains, gene annotation was performed using the Maker2 pipeline (Holt and Yandell 2011) that combines *ab initio* protein-coding gene evidence from SNAP (Korf 2004), Augustus (Stanke and Waack 2003), and GeneMark-HMM (Lukashin and Borodovsky 1998). Additionally, Maker2 was provided with the previously generated reference gene annotation of *V. dahliae* strain JR2 (Faino et al. 2015), gene annotation of *V. dahliae* strain CQ2 (unpublished data), and protein homologs of 260 predicted fungal proteomes obtained from the UniProt database (Apweiler et al. 2004).

Effector profiling

We determined core and LS regions of each *V. dahliae* strain. For LS regions, pairwise whole-genome alignments of the 21 *V. dahliae* strains were performed using NUCmer (version 3.1) (--maxmatch), which is part of the MUMer package (Kurtz et al. 2004), and LS regions (here defined as genomic regions that are shared by <19 *V. dahliae* strains) were extracted. Subsequently, core regions (regions shared by ≥19 *V. dahliae* strains) were determined. Genes localized within core and LS regions were extracted using BEDtools intersect (Quinlan and Hall 2010).

To identify candidate effectors, N-terminal signal peptides were first predicted with SignalP (version 4.1) (Petersen et al. 2011). Subsequently, the machine-learning approach applied in EffectorP (version 1.0) (default parameters) was used (Sperschneider et al. 2016). Effector genes localized within core and LS regions were extracted using BEDtools intersect (Quinlan and Hall 2010). Sequence similarity between predicted LS effectors

was established by an all-vs.-all analyses using BLASTp (E-value cutoff 1e-5) (Altschul et al. 1990). Clustering of LS effector sequences into different families was performed using MCL (default options) (Li et al. 2003) and visualized using the R package pheatmap (Kolde 2015).

Assessment of gene expression

To assess gene expression levels, two RNA-seq datasets were used. The first RNA-seq dataset was previously generated from *V. dahliae* strain JR2-infecting *Nicotiana benthamiana* plants at 4, 8, 12, and 16 dpi (de Jonge et al. 2012; Faino et al. 2014). The second RNA-seq dataset was obtained from *V. dahliae* strain V991 infecting *Gossypium hirsutum* (cotton) plants at 6, 9, 12, and 15 dpi (L. Zhu, unpublished data). Mapping of RNA-seq datasets to the corresponding genomes was performed using STAR (version 2.5.3) (--runThreadN 16) (Dobin et al. 2013), and gene expression levels were determined using RSEM (version 1.2.3) (calculate-expression command) (default parameters) (Li and Dewey 2011), by calculating transcripts per million (TPM) for each gene in each sample.

RESULTS

Verticillium dahliae pathogenicity on a panel of potential host plants

To evaluate the pathogenicity of a collection of *V. dahliae* strains on a panel of potential host plant species, we conducted inoculation experiments with 21 strains on the Solanaceae crop plant tomato and model plant *N. benthamiana*, the Malvaceae crop plant cotton, the Asteraceae crop plant sunflower, and the Brassicaceae model plant Arabidopsis. Despite the fact that *V. dahliae* is often considered a broad host range pathogen, there is no individual *V. dahliae* strain in this collection that is able to cause disease on all tested plant species (Table 1). All isolates are pathogenic on Arabidopsis (Figure 1) and on *N. benthamiana* (Figure 2), albeit that the severity of disease symptoms induced by different strains vary considerably (Figure 1 & Figure 2). Most strains were also found to cause disease on cotton, with the exception of strains 2009-605 and V152 that are non-pathogenic on this host species (Figure 3). Interestingly, *V. dahliae* strains CQ2, 463 and ST100 cause severe defoliation, while the other pathogenic strains induce mild to moderate disease symptoms that include wilting, stunting and chlorosis in the absence of defoliation (Figure 3). Fewer *V. dahliae* strains are able to cause disease on tomato (Figure 4). Besides several strains that cause defoliation on cotton, the tomato non-pathogenic strains also include several non-defoliators on cotton like Vd39 and 85S (Figure 3 & Figure 4). Interestingly, whereas strain V152 that is non-pathogenic on cotton also fails to cause disease on tomato, strain 2009-605 that is non-pathogenic on cotton is able to cause wilt disease on tomato (Figure 4). Strikingly, except for *V. dahliae* strain 85S that induces clear wilt disease symptoms on sunflower plants, including stunting, chlorosis and necrosis, all other strains fail to cause disease on sunflower (Figure 5).

Thus, differential pathogenicity occurs within the collection of *V. dahliae* strains tested here.

Table 1. Inoculation experiments with a collection of *V. dahliae* strains on a collection of potential host plants.

Strain	Arabidopsis	*N. benthamiana*	Cotton	Tomato	Sunflower
ST100	+	+	+	-	-
463	+	+	+	-	-
CQ2	+	+	+	-	-
Vd39	+	+	+	-	-
85S	+	+	+	-	+
JKG8	+	+	+	+	-
DVD-S94	+	+	+	+	-
DVD3	+	+	+	+	-
DVD31	+	+	+	+	-
DVD161	+	+	+	+	-
DVD-S29	+	+	+	+	-
ST14.01	+	+	+	+	-
CBS38166	+	+	+	+	-
DVD-S26	+	+	+	+	-
Vandijk	+	+	+	-	-
ST16.01	+	+	+	-	-
2009-605	+	+	-	+	-
V152	+	+	-	-	-
V52	+	+	+	+	-
VdLs17	+	+	+	+	-
JR2	+	+	+	+	-

'+'= pathogenic; '-'= non-pathogenic. Disease symptoms were scored up to 21 days post inoculation (dpi) (tomato, *N. benthamiana*, Arabidopsis), 28 dpi (cotton) or 45 dpi (sunflower). The inoculation experiments were executed twice with similar results.

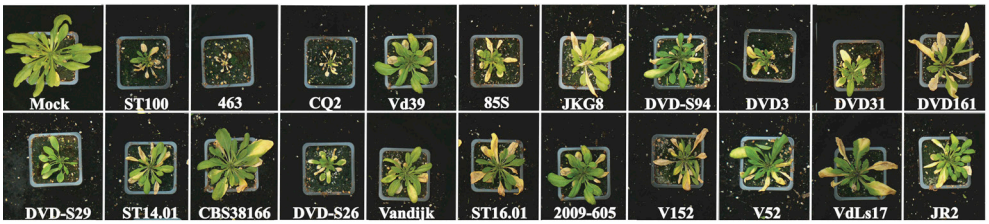


Figure 1. Phenotypes of *Arabidopsis thaliana* inoculated with *V. dahliae* strains. Typical appearance of *A. thaliana* (Col-0) plants upon mock-inoculation or inoculation with a collection of *V. dahliae* strains. All inoculated *V. dahliae* strains are pathogenic on *A. thaliana*. Note that disease symptoms range from stunting, wilting to severe tissue necrosis. Pictures show representative plants at 21 days after inoculation taken from one of two independent inoculation experiments.

Figure 2. Phenotypes of *Nicotiana benthamiana* inoculated with *V. dahliae* strains. Typical appearance of *N. benthamiana* plants upon mock-inoculation or inoculation with a collection of *V. dahliae* strains. All inoculated *V. dahliae* strains are pathogenic on *N. benthamiana*. Note that disease symptoms include stunting, wilting and severe tissue necrosis. Pictures show representative plants at 21 days after inoculation taken from one of two independent inoculation experiments.

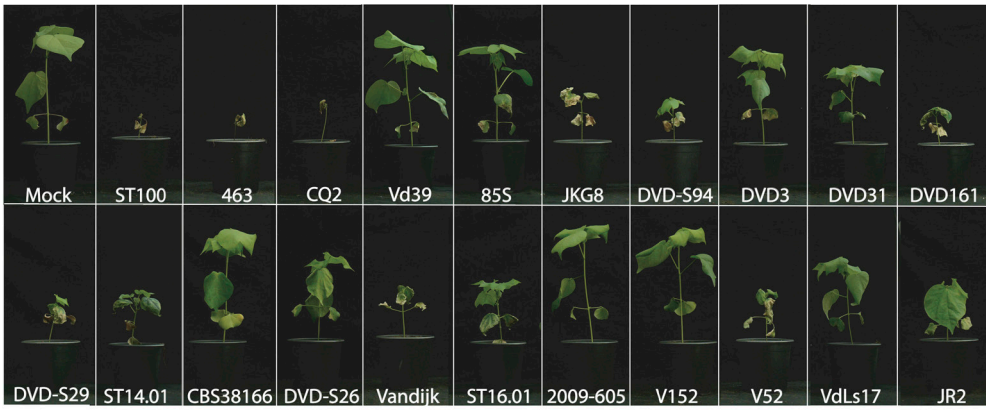


Figure 3. Phenotypes of cotton plants inoculated with *V. dahliae* strains. Typical appearance of cotton (cv. Simian 3) plants upon mock-inoculation or inoculation with a collection of *V.*

dahliae strains. *V. dahliae* strains display differential pathogenticity on cotton plants. Note that several *V. dahliae* strains cause defoliation symptoms, while others induce wilting, stunting but not defoliation. Pictures show representative plants at 28 days after inoculation taken from one of two independent inoculation experiments.

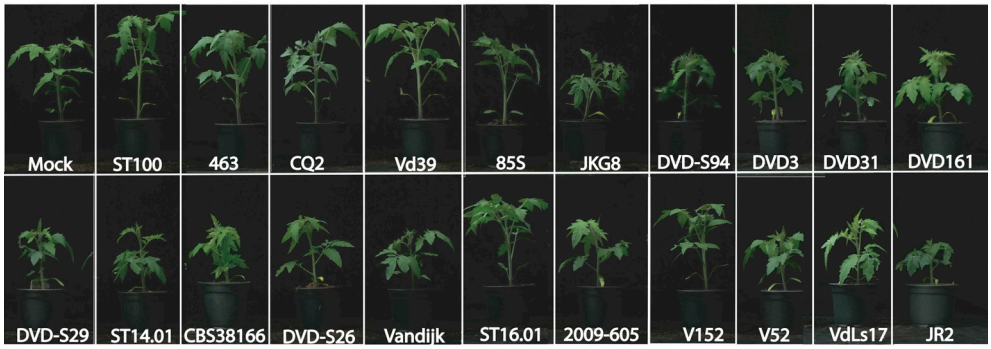


Figure 4. Phenotypes of tomato plants inoculated with *V. dahliae* strains. Typical appearance of tomato (cv. Moneymaker) plants upon mock-inoculation or inoculation with a collection of *V. dahliae* strains. *V. dahliae* strains display differential pathogenticity on tomato plants. Note that pathogenic strains induce clear stunting and significant reduction in canopy area development on inoculated plants. Pictures show representative plants at 21 days after inoculation taken from one of two independent inoculation experiments.

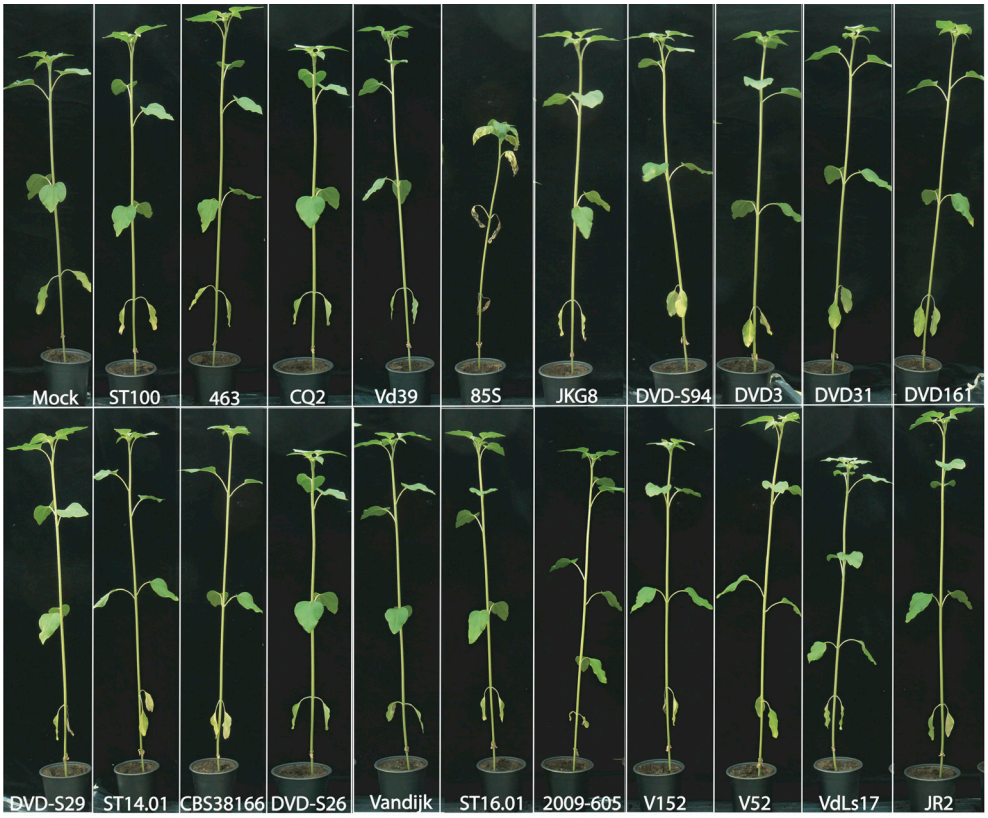


Figure 5. Phenotypes of sunflower plants inoculated with *V. dahliae* strains. Typical appearance of sunflower (cv. Tutti) plants upon mock-inoculation or inoculation with a collection of *V. dahliae* strains. *V. dahliae* strains display differential pathogenticity on sunflower plants. Note that besides the stunting, the plant inoculated with pathogenic strain 85S also displays chlorosis and wilting symptoms. Pictures show representative plants at 45 days after inoculation taken from one of two independent inoculation experiments.

Genome assemblies and annotations of a collection of *V. dahliae* strains

The genome sequences of 13 *V. dahliae* strains (Table S1) were obtained from previous studies, nine of which were determined using the Illumina HiSeq 2000 platform (de Jonge et al. 2013; de Jonge et al. 2012), and four were sequenced using long-read PacBio Single-Molecule Real-Time (SMRT) sequencing technology (Faino et al. 2015). In this study, we sequenced the genomes of eight additional *V. dahliae* strains (Table S1) using the Illumina HiSeq 2000 platform, yielding ~1.1 Gb of paired-end (PE) library-derived reads (500 bp insert size; 100 bp read length) per strain. As the genomes of *V. dahliae* strains that were previously sequenced with Illumina technology showed a reduced N50 size of about 35.55 kb (de Jonge et al. 2012), all the Illumina sequenced genomes were re-assembled in this study. The short reads of the 17 *V. dahliae* strains

that were sequenced with Illumina technology were assembled into ~34 Mb, with the largest assembly of 35.90 Mb for *V. dahliae* strain Vd39, and the smallest assembly of 33.14 Mb for *V. dahliae* strain DVD-S29 (Table S2). All assemblies comprised between 1,000 and 4,188 scaffolds with an N50 of ~50 kb, except for *V. dahliae* strains 463 and V52 that were assembled in 4,188 and 3,419 scaffolds with an N50 of 17.74 kb and 21.80 kb, respectively (Table S2).

To assess the completeness of the assemblies, the Benchmarking Universal Single-Copy Orthologs (BUSCO) software was used, which uses a set of 1,315 core Ascomycota genes as queries (Simão et al. 2015). The BUSCO scores amounted to ~91% for all the assemblies, except for *V. dahliae* strains 463 and v52 that resulted in 75% and 78.6%, respectively (Table S2). For the PacBio sequenced strains, BUSCO scored 99.30% for 85S and 97.50% for CQ2, while the BUSCO scores for the gapless genome assemblies of JR2 and VdLs17 (Faino et al. 2015) amounted to 99.40% and 98.90%, respectively (Table S2).

Repetitive elements are strong drivers of genome evolution in plant pathogens (Seidl and Thomma 2017). Thus, the amounts and types of repetitive elements in the genomes of *V. dahliae* strains were predicted by combining *de novo* and known repetitive elements with RepeatMasker (Smit et al. 2016). The repeat content within the *V. dahliae* genomes varied between 6.64% (2.26 Mb) for *V. dahliae* strain 2009-605 and 13.43% (4.83 Mb) for *V. dahliae* strain 85S (Table S3). Out of all the annotated repetitive elements, different repeat families were identified, which included long terminal repeats (LTRs) (2 Mb, ~5.7%), long interspersed nuclear elements (LINEs) (40 kb, ~0.11%), and short interspersed nuclear elements (SINEs) (2.9 kb, ~0.01%) (Table S3).

We subsequently inferred gene annotations for the various *V. dahliae* strains. For *V. dahliae* strains JR2 and CQ2, a previously determined gene annotation was used (Faino et al. 2015). The completeness of gene annotation for both strains was assessed using BUSCO that scored only a low score for CQ2 (72%) compared to JR2 (90.80%). Thus, gene annotations were inferred for all 20 *V. dahliae* strains, except for *V. dahliae* strain JR2. The Maker2 pipeline (Holt and Yandell 2011) was used that combines *de novo*, homology-based, and previous gene annotations for *V. dahliae* strains CQ2 (B. Thomma and J. Li, unpublished data) and JR2 (Faino et al. 2015), and protein homologs of 260 predicted fungal proteomes. The number of genes varied from 10,461 for *V*. *dahliae* strain VanDijk to 11,341 for *V. dahliae* v52 (Table 2).

Identification of core and lineage-specific (LS) effector catalogs

Initially, the secretomes for each of the *V. dahliae* strains were predicted, identifying between 1,002 secreted proteins for *V. dahliae* strain 463 and 1,108 proteins for *V. dahliae* strain ST16 (Table 2). Subsequently, the machine-learning algorithm of EffectorP (Sperschneider et al. 2016) was used, which identified between 169 effectors for *V. dahliae* strain CQ2 and 212 effectors for *V. dahliae* strain JR2 (Table 2).

Table 2. Summary of genome sizes, core and lineage-specific (LS) regions, number of genes, and number of effector genes for each strain.

Strain	Genome size (Mb)							Number of genes						EffectorP		
	Total	Repeats	%Repeats	Core	%Core	LS	%LS	Total	Secreted	Core	LS	Core/ secreted	LS/ secreted	Total	Core	LS
2009-605	34.06	2.26	6.64	32.16	94.43	1.90	5.57	10893	1105	9707	1186	1020	85	190	170	20
463	34.03	3.56	10.46	32.66	95.96	1.38	4.04	11188	1002	10671	517	967	35	204	190	14
85S	35.93	4.83	13.43	34.11	94.93	1.82	5.07	10580	1062	9706	874	995	67	174	168	6
CBS38166	34.03	2.83	8.32	32.59	95.76	1.44	4.24	10676	1024	9823	853	969	55	180	168	12
DVD161	33.47	2.61	7.81	32.31	96.53	1.16	3.47	10624	1035	9872	752	980	55	195	172	23
DVD31	33.58	2.85	8.50	32.48	96.72	1.10	3.28	10642	1034	9941	701	981	53	185	164	21
DVD3	34.42	3.47	10.10	33.04	96.00	1.37	4.00	10741	1035	9934	807	982	53	191	172	19
DVD-S26	34.74	2.82	8.11	32.59	93.82	2.15	6.18	10977	1057	9871	1106	979	78	196	178	18
DVD-S29	33.14	2.55	7.69	32.07	96.77	1.07	3.23	10563	1020	9866	697	978	42	193	178	15
DVD-S94	34.42	3.31	9.61	32.98	95.80	1.44	4.20	10684	1057	9827	857	990	67	185	165	20
JKG8	33.85	2.75	8.13	32.56	96.21	1.28	3.79	10565	1074	9647	918	1006	68	170	156	14
ST100	34.92	3.66	10.49	33.28	95.32	1.64	4.68	10697	1041	9789	908	972	69	189	170	19
ST14.01	34.48	3.39	9.83	32.92	95.49	1.55	4.51	10660	1039	9716	944	976	63	175	162	13
ST16.01	34.22	2.31	6.74	31.95	93.37	2.27	6.63	11005	1108	9687	1318	1017	91	195	174	21
V152	33.98	2.79	8.22	32.21	94.82	1.76	5.18	11137	1059	10103	1034	988	71	195	175	20
V52	33.54	2.41	7.19	31.83	94.91	1.71	5.09	11341	1030	10648	693	980	50	207	188	19
VanDijk	33.17	2.34	7.06	31.97	96.39	1.20	3.61	10461	1068	9591	870	999	69	180	163	17
Vd39	35.90	4.45	12.39	33.69	93.84	2.21	6.16	10662	1074	9579	1083	988	86	171	159	12
CQ2	35.82	4.15	11.60	34.03	95.00	1.79	5.00	10528	1064	9628	900	994	70	169	165	4
JR2	36.15	4.26	11.77	33.68	93.15	2.47	6.85	11432	1067	10328	1104	1001	66	212	200	12
VdLs17	35.97	4.20	11.68	33.52	93.19	2.45	6.81	10856	1090	9688	1168	1000	90	179	165	14

Subsequently, we determined the core genome, here defined as regions that are shared by ≥19 *V. dahliae* strains, and LS regions, here defined as regions that are shared by <19 *V. dahliae* strains, for all *V. dahliae* strains. The core regions of all *V. dahliae* strains comprise 32.79 Mb (93-97%) of the genome, while LS regions comprise between 1.06 and 2.47 Mb (3-7%) (Table 2). On average, the core regions of *V. dahliae* strains harbor 9,886 genes, comprising 988 genes that encode secreted proteins, of which 171 were classified as effectors based on EffectorP (Table 2). The LS regions of *V. dahliae* strains harbor between 517 genes for *V. dahliae* strain 463 and 1,318 genes for *V. dahliae* strain ST16 (Table 2). Of these LS genes, 35 genes encode secreted proteins for *V. dahliae* strain 463 and 91 for *V. dahliae* strain ST16, of which ~15 genes were classified as effectors for each *V. dahliae* strain (Table 2). We tested these predictions on the previously identified LS effector gene of *V. dahliae* strain JR2, namely *Ave1*, which was successfully identified as a LS effector gene (de Jonge et al. 2012) (Figure S1).

To assess general characteristics of core and LS effector genes, features such as their distance to transposable elements (TEs), gene length, inter-genic length, and expression were determined. For all strains, we observed that LS effector genes are shorter in length than core effector genes (Figure 6A, B; Figure S2). The average inter-genic length of LS effector genes is slightly longer (1476 bp) compared to the average inter-genic length of core effector genes (1232 bp) (Figure 6A, B; Figure S3). Moreover, LS effector genes localize closer to TEs than core effector genes, even though this trend is not significant for *V. dahliae* strain JR2 (Figure 6A, B; Figure S4). Finally, LS effector genes of *V. dahliae* strain JR2 were found to be significantly higher expressed *in planta* on *N. benthamiana* than core effector genes (Figure 7A), although no such difference was found between LS and core effector genes of *V. dahliae* strain CQ2 *in planta* on cotton (Figure 7B).

***V. dahliae* strains that infect the same host plant harbor divergent effector repertoires**

It has previously been shown that *V. dahliae* strains display differential capacity to infect particular host plants (Bhat and Subbarao 1999). We furthermore showed that especially LS effectors contribute to *V. dahliae* pathogenicity on individual plant hosts (de Jonge et al. 2013; de Jonge et al. 2012; Kombrink et al. 2017). Collectively, this suggests that *V. dahliae* strains may harbor an array of specialized effectors that only function on particular host plants. Therefore, we assessed whether the presence of particular LS effectors in the various *V. dahliae* strains correlates with the ability to infect particular hosts. In total, we predicted 333 LS effectors over the various strains (Table 2) that were clustered into 110 families that are either shared by sub-groups of *V. dahliae* strains or are strain-specific (Figure 8). Of these effectors, the previously identified LS effector gene *Ave1* is shared by sub-groups of *V. dahliae* strains (de Jonge et al. 2013), while the *Sun1* effector gene of *V. dahliae* strain 85S is strain specific (Chapter 4) (Figure 8). Notably, in contrast to previous observations, we found that the LS effector genes *XLOC_00170*, *XLOC_008951*, and *XLOC_009059* were only present in *V. dahliae* strain JR2 and not in additional *V. dahliae* strains (de Jonge et al. 2013).

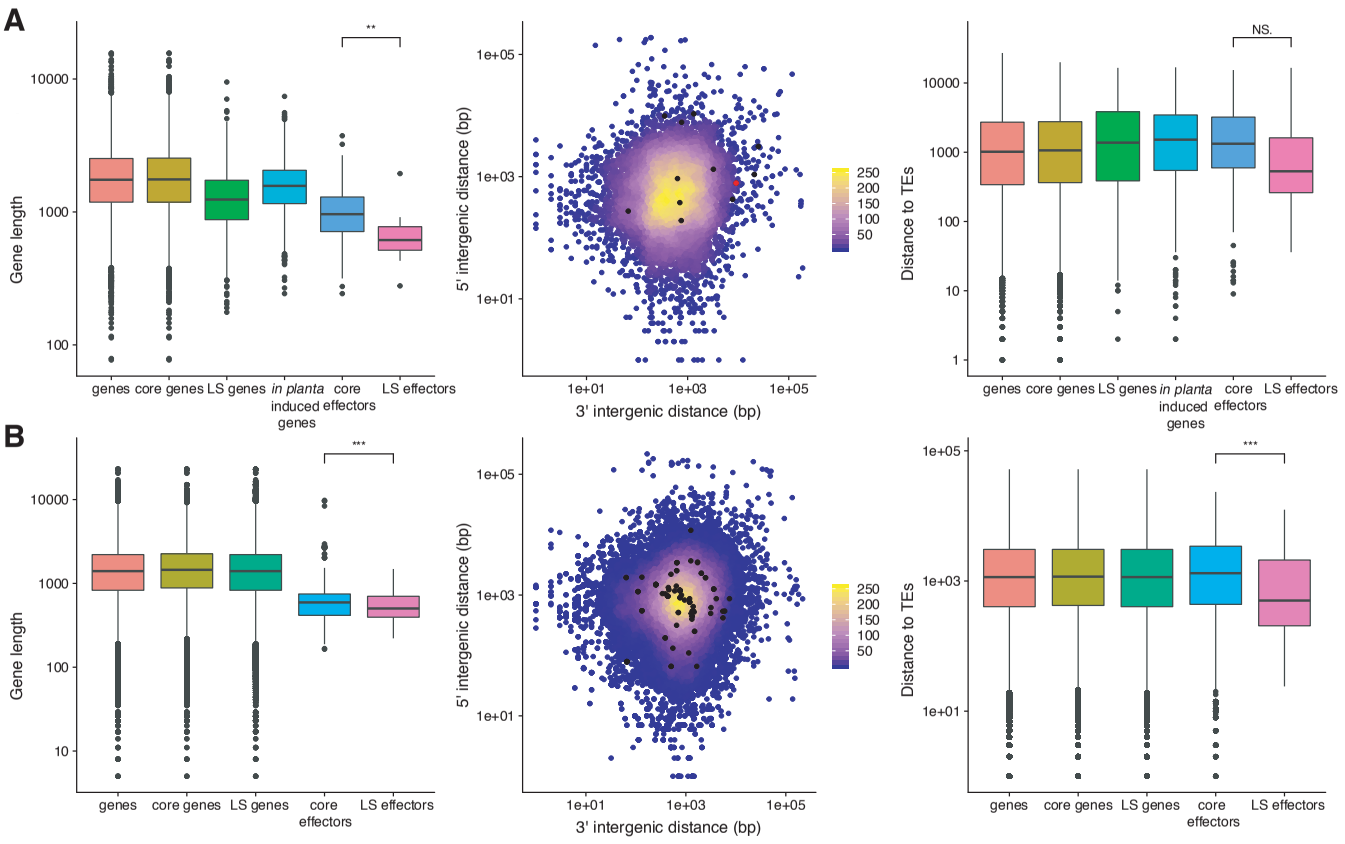


Figure 6. Characteristics of core and LS effector genes. Gene length, two-dimensional density plot of 5'- and 3'-flanking inter-genic regions, where LS effectors are indicated in (black) and Ave1 effector in (red) color, and distance to closest transposable element (TE) for core and LS effector genes of *V. dahliae* strain JR2 **(A)** vs all *V. dahliae* strains **(B)** are displayed.

Thus, we searched the 333 LS effector genes against the *V. dahliae* genome assemblies using BLAST (tblastn), which revealed that multiple (candidate) effector genes, including the previous identified ones, were absent from the gene annotation, highlighting the challenge of computational effector gene identification (Gibriel et al. 2016). Intriguingly, we observed highly dissimilar LS effector catalogs among *V. dahlia*e strains that are able to infect the same host plant (Figure 8). Of the 333 LS effectors, not even a single one is shared among all strains that infect the same host plant (Figure 8). Even more strikingly, *V. dahliae* strains that infect the same host plant do not cluster based on their LS effector repertoires (Figure 8). Thus, *V. dahliae* strains harbor highly divergent LS effector catalogs, the composition of which does not correlate with the host plant they are able to infect.

Variability in core effector transcription profiles

In many plant pathogens, core effector genes are *in planta* highly-induced and play essential roles on a multitude of hosts (Guyon et al. 2014; Hemetsberger et al. 2015; Santhanam et al. 2013; Yin et al. 2017). Thus, we assessed the expression of *V. dahliae* core effector genes during invasion of different plant species. First, we mapped RNA-seq datasets from *N. benthamiana* plants colonized by *V. dahliae* strain JR2 (de Jonge et al. 2012) against the reference genome sequence of *V. dahliae* strain JR2 (Faino et al. 2015), and *G. hirsutum* (cotton) plants colonized by *V. dahliae* strain V991 (LF Zhu, unpublished data) against its closely related *V. dahliae* strain CQ2. Subsequently, RNAseq reads overlapping shared core effector genes of *V. dahliae* JR2 and *V. dahliae* strain CQ2 were quantified. We observed that the transcription profiles can be clustered into: 1) transcribed effector genes on both hosts (those with log10 TPM value>0 in both hosts), 2) differentially transcribed effector genes between the two hosts (those with log10 TPM value of 0 in one host and >0 in the other host), and 3) non-transcribed effector genes on either hosts (those with TPM value of 0 in both hosts) (Figure 9A). Of the 165 shared core effector genes between JR2 and CQ2, 61 effector genes were transcribed during cotton as well as *N. benthamiana* colonization, whereas 19 effector genes were only transcribed on cotton, and 41 effector genes were transcribed on *N. benthamiana* (Figure 9B). Additionally, we identified 44 effector genes that were non-transcribed in both strains (Figure 9A). Thus, differential *V. dahliae* core effector gene expression is observed on different host plants. Overall, we are not able to link the composition or the expression of effector gene catalogs to the ability to infect particular host plants.

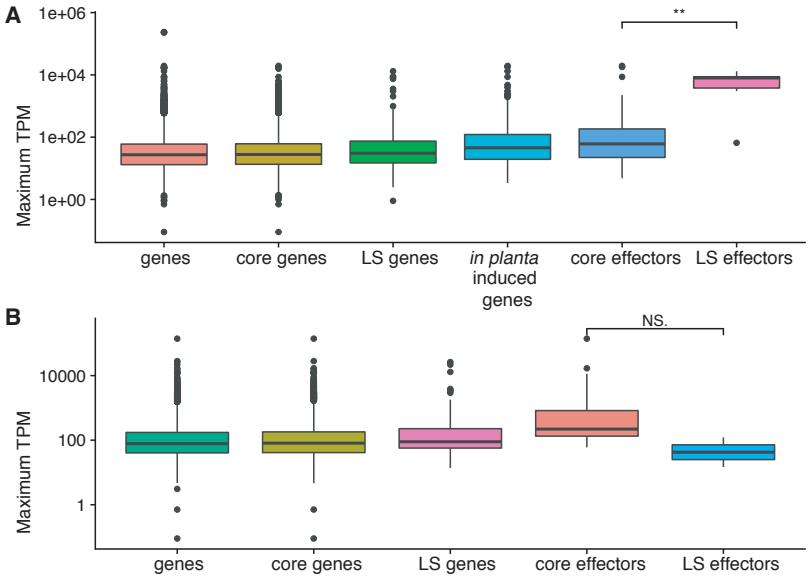


Figure 7. Maximum TPM values of core and LS effector genes over four time points during *N. benthamiana* and cotton infection by *V. dahliae* strains JR2 and V911. A) Maximum TPM values per time point (4, 8, 12, and 16) of core and LS effector genes during *N. benthamiana* infection by *V. dahliae* strain JR2. **B)** Maximum TPM values per time point (6, 9, 12, and 15) of core and LS effector genes during cotton infection by *V. dahliae* strain V911.

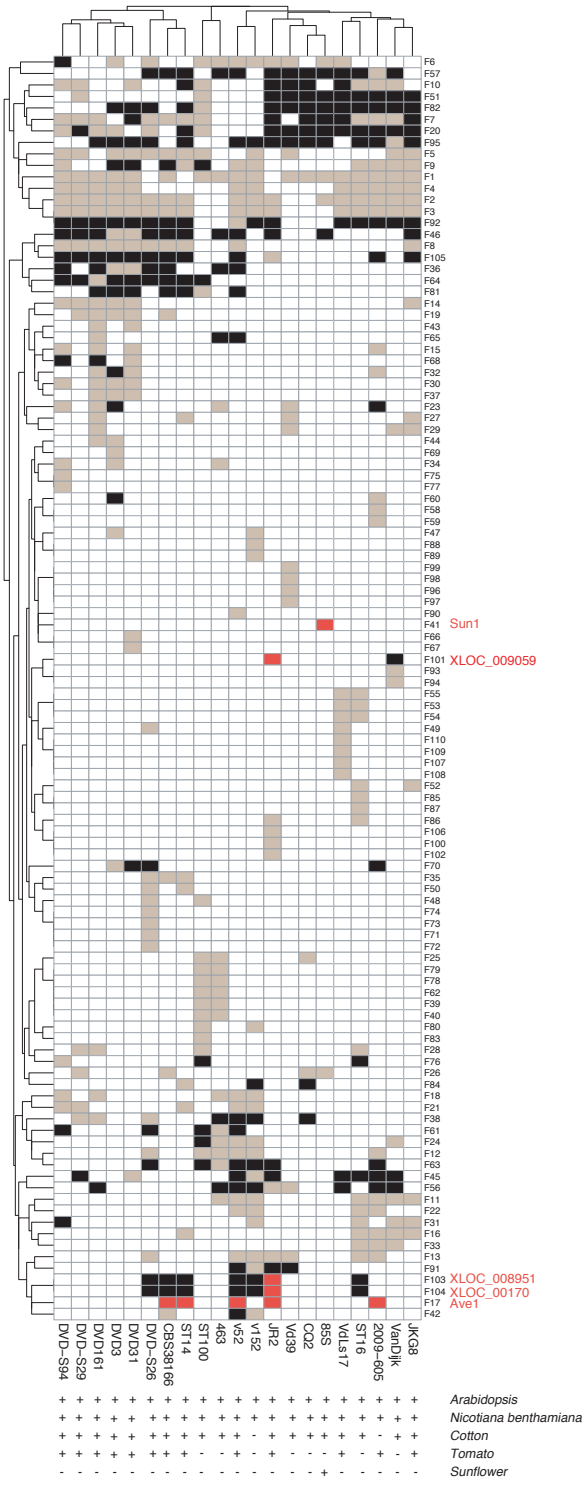


Figure 8. *Verticillium dahliae* strains that infect the same host plant harbor highly divergent LS effector repertoires. LS effectors were clustered into 110 families and their presence (brown) and absence (white) between *V. dahliae* strains is shown. Previously described LS effectors are colored in red (de Jonge et al. 2013; Chapter 4). Black colors indicate the absence of LS effectors from the predicted gene annotation for *V. dahliae* strains and the presence in *V. dahliae* genomes based on BLAST (tblastn).

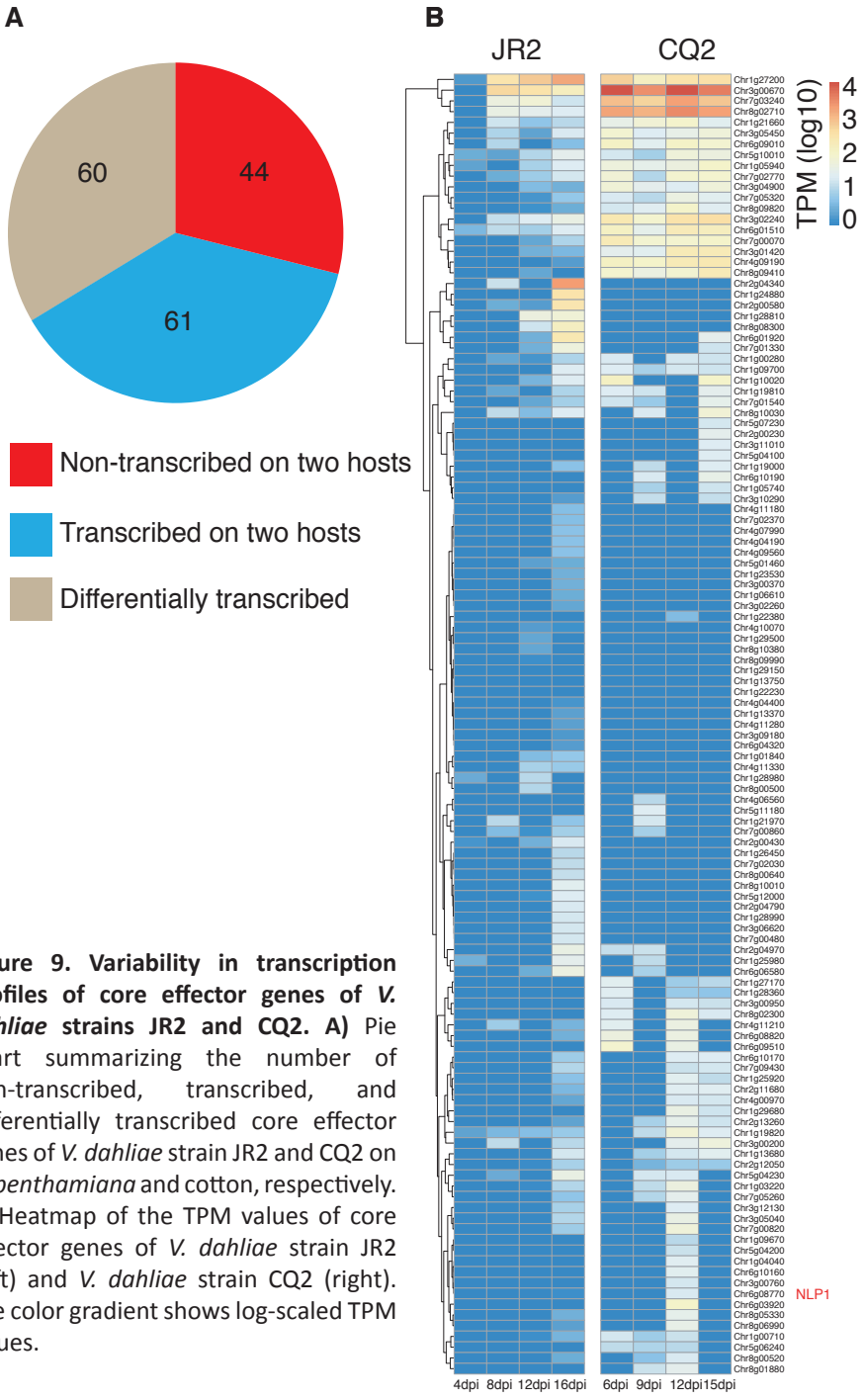


Figure 9. Variability in transcription profiles of core effector genes of *V. dahliae* strains JR2 and CQ2. A) Pie chart summarizing the number of non-transcribed, transcribed, and differentially transcribed core effector genes of *V. dahliae* strain JR2 and CQ2 on *N. benthamiana* and cotton, respectively. **B)** Heatmap of the TPM values of core effector genes of *V. dahliae* strain JR2 (left) and *V. dahliae* strain CQ2 (right). The color gradient shows log-scaled TPM values.

DISCUSSION

For broad host-range pathogens such as *V. dahliae* it is unclear to what extent co-evolution of the pathogen with multiple hosts occurs simultaneously, and what implications this has for their effector repertoires. Whereas it is conceivable that broad host-range pathogens employ core effectors that are active on a multitude of hosts because they target broadly conserved general physiological processes in these hosts, they may also harbor an array of specialized effectors that only exert their activity on particular host species. Presence of such specialized effectors is suggested by the observation that different *V. dahliae* strains, despite their general ability to infect a wide array of host plants, generally display differential capacity to infect particular hosts (Bhat and Subbarao 1999). Therefore, in this study we investigated whether the absence or presence of particular effectors correlates with the ability to infect particular hosts. First, we predicted the LS effector gene repertoires of *V. dahliae* strains and identified 333 LS effector genes, of which the previously identified LS effector genes *Ave1* of *V. dahliae* strain JR2 (de Jonge et al. 2013) and the *Sun1* effector gene of *V. dahliae* strain 85S (Chapter 4) were successfully predicted (Figure 8). As improper gene annotation could hamper effector gene discovery (Gibriel et al. 2016), we searched the 333 effector genes against *V. dahliae* genomes using BLAST (tblastn) and successfully identified the previously described effector genes *XLOC_00170, XLOC_008951,* and *XLOC_009059* of *V. dahliae* strain JR2, which are shared by a subset of *V. dahliae* strains (de Jonge et al. 2013) (Figure 8). Intriguingly, we observed that *V. dahliae* strains that are able to infect the same host plant, in this case focused on tomato, cotton, *N. benthamiana*, Arabidopsis, and sunflower, harbor highly divergent LS effector catalogs and not even a single LS effector is shared by the strains that are able to infect either of these host species (Figure 8). This strongly suggests that different strains infect the same host plant by utilizing different effector compositions. If one assumes that different strains of *V. dahliae* must target the same host physiological processes in order to establish the infection, this implies that there must be a significant degree of redundancy among the various effectors that occur in the different strains. Interestingly, fungi of the ascomycete species *Fusarium oxysporum* display a similar infection biology as *V. dahliae*, being soil-borne pathogens that colonize the xylem tissues of their host plants to cause vascular wilt disease and with a largely overlapping host range (Ma et al. 2013). However, whereas individual strains of *V. dahliae* are characterized by their generally broad host range, strains of *F. oxysporum* are generally host-specific and are therefore assigned to *formae speciales* (Ma et al. 2013; Ma et al. 2010). Comparative genomics revealed that *Fusarium* strains able to infect various host plants carry highly overlapping LS effector repertoires (Coleman et al. 2009; Ma et al. 2010; van Dam et al. 2016). Moreover, strains that infect the same host plant cluster based on their LS effector repertoires (Van Dam et al. 2016). Thus, in contrast to *F. oxysporum* strains that harbor highly overlapping LS effectors based on their host range (Coleman et al. 2009; Ma et al. 2010; van Dam et al. 2016), *V. dahliae* strains evolved highly divergent LS effector catalogs, the composition of which does not correlate with the host plant they are able to infect.

Plant pathogens harbor an array of *in planta* highly-induced core effector genes that play essential roles on a multitude of hosts (Guyon et al. 2014; Hemetsberger et al. 2015; Santhanam et al. 2013; Yin et al. 2017). We quantified RNAseq reads (RNAseq reads of *V. dahliae* JR2-infecting *N. benthamiana* plants and RNAseq reads of *V. dahliae* strain V991-infecting cotton) overlapping core effector genes of JR2 and CQ2 (a strains closely related to V991). Subsequently, we categorized the effectors in three groups based on their expression profiles. We identified a group of effector genes that are transcribed in both strains (Figure 9B), suggesting that this group of effectors contributes to *V. dahliae* colonization of *N. benthamiana* and cotton. Additionally, we observed a group of effector genes that are highly transcribed in *V. dahliae* strain JR2 on *N. benthamiana* but only lowly transcribed in strain CQ2 on cotton, and vice versa (Figure 9B), suggesting that they differentially contribute to virulence on these two host plants. Furthermore, we identified a group of effector genes that are not transcribed in either strains (Figure 9A), suggesting that this group of core effectors do not play a role in virulence on *N. benthamiana* or on cotton. Nevertheless, it may well be that this group of effector genes is transcribed during *V. dahliae* colonization of other host plants. Surprisingly, we observed that a member of the family of necrosis- and ethylene-inducing-like proteins (NLPs), namely *NLP1*, is not expressed on *N. benthamiana* and is lowly expressed on cotton (Figure 9B), whereas we previously found based on real-time PCR that this effector gene is transcribed in *V. dahliae* strain JR2 on *N. benthamiana* and on tomato, but also in *V. dahliae* strain V592 on cotton (Santhanam et al. 2013; Zhou et al. 2012). Likely, the low expression of *NLP1* expression based on our RNAseq data may be due to the low amount of fungal RNAseq reads among plant-derived ones, as only 0.05% of the reads could be mapped to the *V. dahliae* genome (Faino et al. 2014). Thus, it needs to be taken into account that only highly expressed fungal genes are identified, and it might be worthwhile to confirm whether the group of non-transcribed effector genes still contains lowly or moderately expressed genes based on real-time PCR.

Localization of pathogen effector genes within dynamic genomic regions allows pathogens to rapidly evolve to evade plant immunity once an effector gets recognized by the host (Dong et al. 2015; Raffaele and Kamoun 2012). When we compared the core and LS effector repertoires between the analysed *V. dahliae* strains, we observed no remarkable differences in the number of core and LS effectors (Table 2). Nevertheless, by comparing features of core and LS effectors we observed that LS effectors are typically shorter in length, localize significantly closer to TEs, and have slightly longer inter-genic lengths when compared with core effectors (Figure 6). Consistent with this observation, a recent study that analysed the genomes of nine different *Verticillium* species showed that species-specific genes displayed significantly shorter gene lengths and longer inter-genic lengths when compared with genes that are conserved across the various species within the genus (Shi-Kunne et al. 2017). Similarly, LS genes of the fungal wheat pathogen *Z. tritici* were frequently found to be shorter, closer to TEs, and have longer inter-genic lengths, when compared with core genes (Haueisen et al. 2017; Plissonneau et al. 2018; Plissonneau et al. 2016). It has been previously suggested

that the localization of effector genes in close proximity to TEs and within gene-poor regions mediate rapid evolution of effector catalogs (Raffaele and Kamoun 2012; Seidl and Thomma 2017). Thus, the localization of LS effector genes of *V. dahliae* strains in close proximity to TEs may mediate accelerated evolution of effector catalogs.

In conclusion, our data demonstrate the extensive variability within the effector repertoires of the broad host range pathogen *V. dahliae*. We have demonstrated that LS effectors are highly divergent among *V. dahliae* strains that infect the same host plant, and core effector genes are differentially expressed between hosts. The variability within LS and core effector genes of *V. dahliae* strains may lead to rapid immunity evasion, which may allow pathogen strains to be competitive in the co-evolution with their multiple hosts.

ACKNOWLEDGEMENTS

J Li acknowledges a PhD fellowship from China Scholarship Council (CSC). Work in the laboratories of M.F.S and B.P.H.J.T. and is supported by the Research Council Earth and Life Sciences (ALW) of the Netherlands Organization of Scientific Research (NWO).

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. Journal of Molecular Biology 215.

Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M. J., Natale, D. A., O'Donovan, C., Redaschi, N., and Yeh, L.-S. L. 2004. Uniprot: the universal protein knowledgebase. Nucleic Acids Research 32:D115-D119.

Bhat, R. G., and Subbarao, K. V. 1999. Host range specificity in *Verticillium dahliae*. Phytopathology 89:1218-1225.

Coleman, J. J., Rounsley, S. D., Rodriguez-Carres, M., Kuo, A., Wasmann, C. C., Grimwood, J., Schmutz, J., Taga, M., White, G. J., Zhou, S., Schwartz, D. C., Freitag, M., Ma, L.-j., Danchin, E. G. J., Henrissat, B., Coutinho, P. M., Nelson, D. R., Straney, D., Napoli, C. A., Barker, B. M., Gribskov, M., Rep, M., Kroken, S., Molnár, I., Rensing, C., Kennell, J. C., Zamora, J., Farman, M. L., Selker, E. U., Salamov, A., Shapiro, H., Pangilinan, J., Lindquist, E., Lamers, C., Grigoriev, I. V., Geiser, D. M., Covert, S. F., Temporini, E., and VanEtten, H. D. 2009. The genome of *Nectria haematococca*: Contribution of supernumerary chromosomes to gene expansion. PLoS Genetics 5:e1000618.

Cook, D. E., Mesarich, C. H., and Thomma, B. P. H. J. 2015. Understanding plant immunity as a surveillance system to detect invasion. Annual Review of Phytopathology 53:541-563.

Croll, D., and McDonald, B. A. 2012. The accessory genome as a cradle for adaptive evolution in pathogens. PLoS Pathogens 8:e1002608.

de Jonge, R., Bolton, M. D., Kombrink, A., van den Berg, G. C. M., Yadeta, K. A., and Thomma, B. P. H. J. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Research 23:1271-1282.

de Jonge, R., Bolton, M. D., and Thomma, B. P. H. J. 2011. How filamentous pathogens co-opt plants: the ins and outs of fungal effectors. Current Opinion in Plant Biology 14:400-406.

de Jonge, R., Peter van Esse, H., Maruthachalam, K., Bolton, M. D., Santhanam, P., Saber, M. K., Zhang, Z., Usami, T., Lievens, B., Subbarao, K. V., and Thomma, B. P. H. J. 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. Proceedings of the National Academy of Sciences of the United States of America 109:5110-5115.

de Jonge, R., van Esse, H. P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., van der Krol, S., Shibuya, N., Joosten, M. H., and Thomma, B. P. H. J. 2010. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. Science 329:953-955.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.

Dong, S., Raffaele, S., and Kamoun, S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Current Opinion in Genetics & Development 35:57-65.

Faino, L., de Jonge, R., and Thomma, B. P. H. J. 2014. The transcriptome of *Verticillium dahliae*-infected *Nicotiana benthamiana* determined by deep RNA sequencing. Plant Signaling & Behavior 7:1065-1069.

Faino L, Seidl MF, Datema E, van den Berg GC, Janssen A, Wittenberg AH, Thomma B. P. H. J. 2015. Single-molecule real-time sequencing combined with optical mapping yields completely finished fungal genome. mBio 6: e00936-00915.

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484:186-194.

Fradin, E. F., and Thomma, B. P. H. J. 2006. Physiology and molecular aspects of Verticillium wilt diseases caused by *V. dahliae* and *V. albo-atrum*. Molecular Plant Pathology 7:71-86.

Fradin, E. F., Zhang, Z., Juarez Ayala, J. C., Castroverde, C. D. M., Nazar, R. N., Robb, J., Liu, C.-M., and Thomma, B. P. H. J. 2009. Genetic dissection of Verticillium wilt resistance mediated by tomato Ve1. Plant Physiology 150:320-332.

Gibriel, H. A. Y., Thomma, B. P. H. J., and Seidl, M. F. 2016. The age of effectors: genome-based discovery and applications. Phytopathology 106: 1206–1212.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072-1075.

Guyon, K., Balagué, C., Roby, D., and Raffaele, S. 2014. Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. BMC Genomics 15:336.

Haueisen, J., Moeller, M., Eschenbrenner, C. J., Grandaubert, J., Seybold, H., Adamiak, H., and Stukenbrock, E. H. 2017. Extremely flexible infection programs in a fungal plant pathogen. bioRxiv: https://doi.org/10.1101/229997.

Hemetsberger, C., Mueller, A. N., Matei, A., Herrberger, C., Hensel, G., Kumlehn, J., Mishra, B., Sharma, R., Thines, M., Hückelhoven, R., and Doehlemann, G. 2015. The fungal core effector Pep1 is conserved across smuts of dicots and monocots. New Phytologist 206:1116-1126.

Holt, C., and Yandell, M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12:491.

Inderbitzin, P., Bostock, R. M., Davis, R. M., Usami, T., Platt, H. W., and Subbarao, K. V. 2011. Phylogenetics and taxonomy of the fungal vascular wilt pathogen *Verticillium*, with the descriptions of five new species. PLoS One 6:e28341.

Jones, J. D. G., and Dangl, J. L. 2006. The plant immune system. Nature 444:323-329.

Kema, G. H. J., Mirzadi Gohari, A., Aouini, L., Gibriel, H. A. Y., Ware, S. B., van den Bosch, F., Manning-Smith, R., Alonso-Chavez, V., Helps, J., Ben M'Barek, S., Mehrabi, R., Diaz-Trujillo, C., Zamani, E., Schouten, H. J., van der Lee, T. A. J., Waalwijk, C., de Waard, M. A., De Wit, P. J. G. M., Verstappen, E. C. P., Thomma, B. P. H. J., Meijer, H. J. G., and Seidl, M. F. 2018. Stress and sexual reproduction affect the dynamics of the wheat pathogen effector AvrStb6 and strobilurin resistance. Nature Genetics 23:678.

Kolde, R. 2015. pheatmap: Pretty Heatmaps. R package version 1.0. 8.

Kombrink, A., Rovenich, H., Shi-Kunne, X., Rojas-Padilla, E., van den Berg, G. C. M., Domazakis, E., de Jonge, R., Valkenburg, D.-J., Sánchez-Vallet, A., Seidl, M. F., and Thomma, B. P. H. J. 2017. *Verticillium dahliae* LysM effectors differentially contribute to virulence on plant hosts. Molecular Plant Pathology 18:596-608.

Korf, I. 2004. Gene finding in novel genomes. BMC Bioinformatics 5: 59.

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S. L. 2004. Versatile and open software for comparing large genomes. Genome Biology 5: R12.

Li, B., and Dewey, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323.

Li, L., Stoeckert, C. J., and Roos, D. S. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Research 13:2178-2189.

Lukashin, A. V., and Borodovsky, M. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Research 26.

Ma, L.-J., Geiser, D. M., Proctor, R. H., Rooney, A. P., O'Donnell, K., Trail, F., Gardiner, D. M., Manners, J. M., and Kazan, K. 2013. *Fusarium* pathogenomics. Annual Review of Microbiology 67:399-416.

Ma, L. J., Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M. J., Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., and Henrissat, B. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464.

Marshall, R., Kombrink, A., Motteram, J., Loza-Reyes, E., Lucas, J., Hammond-Kosack, K. E., Thomma, B. P. H. J., and Rudd, J. J. 2011. 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 Physiology 156:756-769.

Meile, L., Croll, D., Brunner, P. C., Plissonneau, C., Hartmann, F. E., McDonald, B. A., and Sánchez-Vallet, A. 2018. A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance to septoria tritici blotch. New Phytologist 65:512.

Mentlak, T. A., Kombrink, A., Shinya, T., Ryder, L. S., Otomo, I., Saitoh, H., Terauchi, R., Nishizawa, Y., Shibuya, N., Thomma, B. P. H. J., and Talbot, N. J. 2012. Effector-mediated suppression of chitin-triggered immunity by *Magnaporthe oryzae* is necessary for rice blast disease. The Plant Cell 24:322-335.

Pennisi, E. 2010. Armed and dangerous. Science 327:804-805.

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods 8:785-786.

Plissonneau, C., Hartmann, F. E., and Croll, D. 2018. Pangenome analyses of the wheat pathogen *Zymoseptoria tritici* reveal the structural basis of a highly plastic eukaryotic genome. BMC Biology 16:673.

Plissonneau, C., Stürchler, A., and Croll, D. 2016. The evolution of orphan regions in genomes of a fungal pathogen of wheat. mBio 7.

Quinlan, A. R., and Hall, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26.

Raffaele, S., and Kamoun, S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nature Review Microbiology 10:417-430.

Santhanam, P., van Esse, H. P., Albert, I., Faino, L., Nürnberger, T., and Thomma, B. P. H. J. 2013. Evidence for functional diversification within a fungal NEP1-like protein family. Molecular Plant Microbe Interactions 26:278-286.

Schmidt, S. M., Houterman, P. M., Schreiver, I., Ma, L., Amyotte, S., Chellappan, B., Boeren, S., Takken, F. L. W., and Rep, M. 2013. MITEs in the promoters of effector genes allow prediction of novel virulence genes in *Fusarium oxysporum*. BMC Genomics 14:1-21.

Seidl, M. F., and Thomma, B. 2017a. Transposable elements direct the coevolution between plants and microbes. Trends in Genetics 33:842-851.

Shi-Kunne, X., Faino, L., van den Berg, G. C. M., Thomma, B. P. H. J., and Seidl, M. F. 2017. Evolution within the fungal genus *Verticillium* is characterized by chromosomal rearrangement and gene loss. Environmental Microbiology 20:1362–1373.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210-3212.

Smit, A., and Hubley, R. 2010. RepeatModeler Open-1.0. Repeat Masker Website.
http://www.repeatmasker.org/Repeat-Modeler.html.

Smit, A., Hubley, R., and Green, P. 2016. RepeatMasker Open-4.0. Available from: http://www.repeatmasker.org

Song, Y., Liu, L., Wang, Y., Valkenburg, D.-J., Zhang, X., Zhu, L., and Thomma, B. P. H. J. 2017. Transfer of tomato immune receptor Ve1 confers Ave1-dependent *Verticillium* resistance in tobacco and cotton. Plant Biotechnology Journal 16:638-648.

Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., Manners, J. M., and Taylor, J. M. 2016. EffectorP: predicting fungal effector proteins from secretomes using machine learning. New Phytologist 210:743-761.

Stanke, M., and Waack, S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics 19.

Stergiopoulos, I., van den Burg, H. A., Okmen, B., Beenen, H. G., van Liere, S., Kema, G. H., and de Wit, P. J. 2010. Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. Proceedings of the National Academy of Sciences of the United States of America 107:7610-7615.

Takahara, H., Hacquard, S., Kombrink, A., Hughes, H. B., Halder, V., Robin, G. P., Hiruma, K., Neumann, U., Shinya, T., Kombrink, E., Shibuya, N., Thomma, B. P. H. J., and O'Connell, R. J. 2016. *Colletotrichum higginsianum* extracellular LysM proteins play dual roles in appressorial function and suppression of chitin-triggered plant immunity. New Phytologist 211:1323-1337.

Tritt, A., Eisen, J. A., Facciotti, M. T., and Darling, A. E. 2012. An integrated pipeline for *de novo* assembly of microbial genomes. PLoS One 7:e42304.

van Dam, P., Fokkens, L., Schmidt, S. M., Linmans, J. H. J., Kistler, H. C., Ma, L.-J., and Rep, M. 2016. Effector profiles distinguish *formae speciales* of *Fusarium oxysporum*. Environmental Microbiology 18:4087-4102.

Yin, J., Gu, B., Huang, G., Tian, Y., Quan, J., Lindqvist-Kreuze, H., and Shan, W. 2017. Conserved RXLR effector genes of *Phytophthora infestans* expressed at the early stage of potato infection are suppressive to host defense. Frontiers in Plant Sciences 8.

Zhou, B. J., Jia, P. S., Gao, F., and Guo, H. S. 2012. Molecular characterization and functional analysis of a necrosis- and ethylene-inducing, protein-encoding gene family from *Verticillium dahliae*. Molecular Plant-Microbe Interactions 25:964-975.

SUPPLEMENTAL INFORMATION

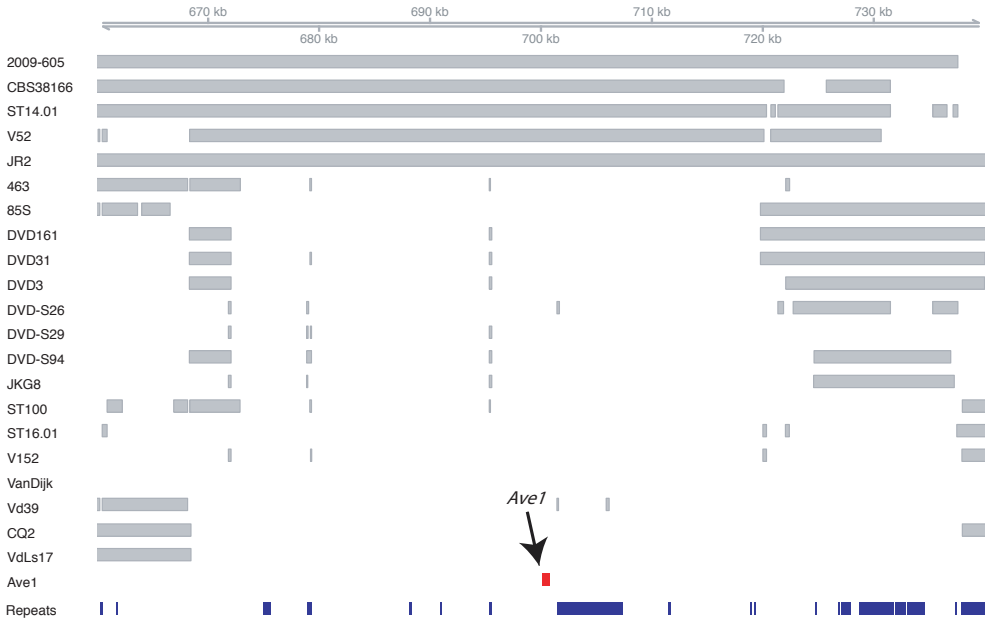


Figure S1. *V. dahliae* effector gene *Ave1* is localized within a lineage-specific region. A highly variable, repeat-rich (blue), lineage-specific (LS) region, which harbors the *Ave1* effector gene (red) of *V. dahliae* strain JR2 is shown. This LS effector gene is only present in a subset of *V. dahliae* strains. Grey bars indicate the genome alignments of *V. dahliae* strains to the reference strain JR2. This effector gene is only present in strains 2009-605, CBS38166, ST14.01, V52, and JR2, but absent in the other strains.

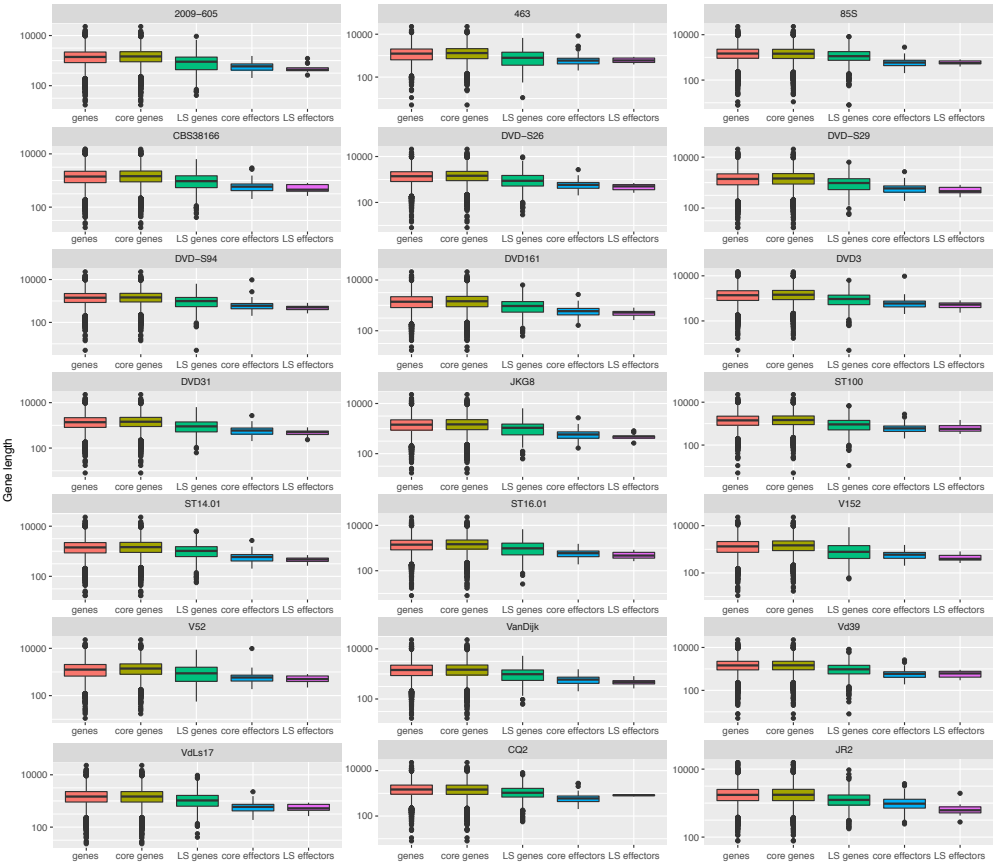


Figure S2. Length of genes and effector genes of *V. dahliae* strains. Gene length (log10) is shown for all genes, core genes, LS genes, core effectors, and LS effectors for all *V. dahliae* strains.

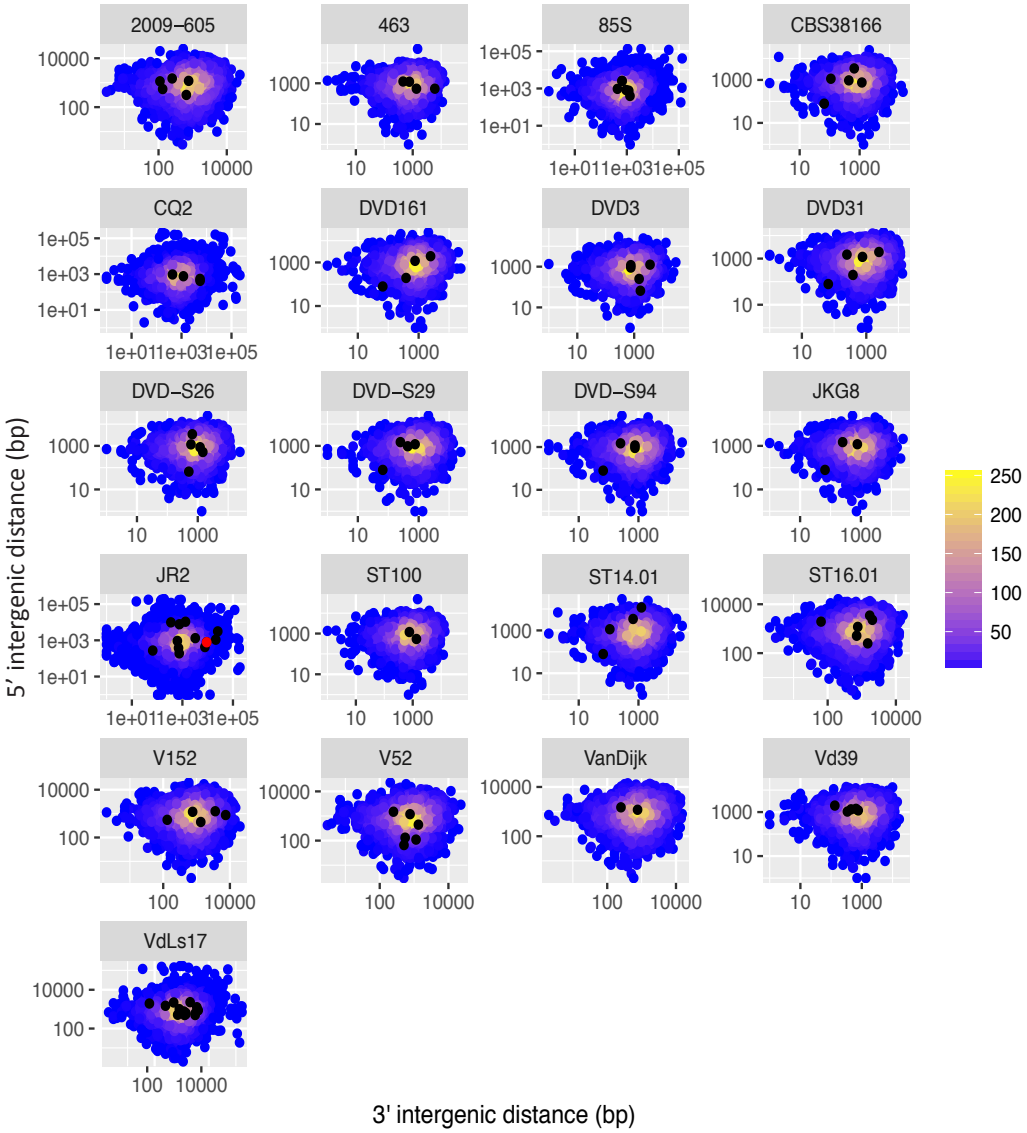


Figure S3. 5′ and 3′ inter-genic length of genes and effector genes of *V. dahliae* strains. Inter-genic length (bp) is shown for all genes (blue) and LS effector genes (black) for all *V. dahliae* strains. *Ave1* effector gene of *V. dahliae* strain JR2 is highlighted in red.

Figure S4. Distance to TEs of genes and effector gnenes of *V. dahliae* strains. Distance to TEs (log10) is shown for all genes, core genes, LS genes, core effectors, and LS effectors for all *V. dahliae* strains.

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

Strain	Sequencing platform	Reference	Originating host	Geographical location
CQ2	PacBio	(Depotter et al. 2018)	Cotton	China
85S	PacBio	(Depotter et al. 2018)	Sunflower	France
VdLs17	PacBio	(Faino et al. 2015)	Lettuce	Ca, USA
JR2	PacBio	(Faino et al. 2015)	Tomato	ON, Canada
CBS38166	Illumina HiSeq 2000	(de Jonge et al. 2012)	Tomato	QC, Canada
ST14.01	Illumina HiSeq 2000	(de Jonge et al. 2012)	Pistachio	CA, USA
ST100	Illumina HiSeq 2000	(de Jonge et al. 2013)	Soil	Belgium
DVD3	Illumina HiSeq 2000	(de Jonge et al. 2012)	Potato	Canada
DVD31	Illumina HiSeq 2000	(de Jonge et al. 2012)	Tomato	Canada
DVD161	Illumina HiSeq 2000	(de Jonge et al. 2012)	Potato	ON, Canada
DVD-S26	Illumina HiSeq 2000	(de Jonge et al. 2012)	Soil	Canada
DVD-S29	Illumina HiSeq 2000	(de Jonge et al. 2012)	Soil	Canada
DVD-S94	Illumina HiSeq 2000	(de Jonge et al. 2012)	Soil	Canada
JKG8	Illumina HiSeq 2000	This study	Potato	The Netherlands
2009-605	Illumina HiSeq 2000	This study	Bell pepper	Ukraine
463	Illumina HiSeq 2000	This study	Cotton	Mexico
ST16.01	Illumina HiSeq 2000	This study	Cotton	Syria
V152	Illumina HiSeq 2000	This study	Oak	Hungry
V52	Illumina HiSeq 2000	This study	Pepper	Austria
Vd39	Illumina HiSeq 2000	This study	Sunflower	Germany
VanDijk	Illumina HiSeq 2000	This study	Chrysanthemum	The Netherlands

Table S2. Assembly statistics for the various *Verticillium dahliae* genomes.

Strain	Genome size (Mb)	# scaffolds (>= 0 bp)	# scaffolds (>= 1000 bp)	GC (%)	N50 (Kb)	# N's per 100 kbp	BUSCO (%)
2009-605	34.06	1931	1521	54.76	55.11	123.45	95.4
463	34.03	4188	3562	53.47	17.74	244.06	75
85S	35.93	40	40	53.55	3176.09	0	99.3
CBS38166	34.03	2092	1727	54.24	45.08	316.06	91.6
DVD161	33.47	2155	1819	54.4	41.42	330.79	90.8
DVD31	33.58	2429	2064	54.14	35.75	336.71	89.2
DVD3	34.42	1921	1693	53.62	42.77	378.65	88.9
DVD-S26	34.74	2275	1894	54.2	43.54	367.67	92.2
DVD-S29	33.14	2226	1811	54.56	42.73	306.89	91.6
DVD-S94	34.42	1730	1494	53.92	53.17	287	91.9
JKG8	33.85	1840	1458	54.5	56.35	172.12	95.6
ST100	34.92	2137	1756	53.63	49.89	315.39	93
ST14.01	34.48	1571	1336	54.03	61.81	221.84	95
ST16.01	34.22	1821	1454	54.9	57.46	121.71	95.6
V152	33.98	2539	2174	54.28	32.47	146.98	87.6
V52	33.54	3419	2920	54.39	21.80	162.37	78.6
VanDijk	33.17	1769	1389	54.71	61.05	109.26	95.7
Vd39	35.90	1579	1222	53.55	94.44	222.52	98.7
CQ2	35.82	17	17	53.26	3754.19	0	97.5
JR2	36.15	8	8	53.89	4168.63	0	99.4
VdLs17	35.97	8	8	54	5894.01	0	98.9

Table S3. Summary of transposable elements of the various *Verticillium dahliae* strains.

	SINEs		LINEs		LTRs	
Strain	length occupied (bb)	percentage of sequence (%)	length occupied (bb)	percentage of sequence (%)	length occupied (bb)	percentage of sequence (%)
2009-605	0	0	0	0	1403090	4.12
463	0	0	43660	0.13	1601065	4.7
85S	2289.00	0.01	57668.00	0.16	3082210.00	8.58
CBS38166	0.00	0.00	29731.00	0.09	1625629.00	4.78
DVD161	1994.00	0.01	0.00	0.00	1871766.00	5.59
DVD31	15326.00	0.05	14424.00	0.04	1718276.00	5.12
DVD3	1923.00	0.01	17323.00	0.05	2262138.00	6.57
DVD-S26	0.00	0.00	34185.00	0.10	1847224.00	5.32
DVD-S29	22393.00	0.01	7832.00	0.02	1297199.00	3.91
DVD-S94	0.00	0.00	23539.00	0.07	2082451.00	6.05
JKG8	2294.00	0.01	0.00	0.00	1344573.00	3.97
ST100	1720.00	0.01	50597.00	0.14	2693973.00	7.71
ST14.01	0.00	0.00	30264.00	0.09	1610719.00	4.67
ST16.01	1582.00	0.00	50167.00	0.15	1009974.00	2.95
V152	0.00	0.00	40614.00	0.12	1660694.00	4.89
V52	3965.00	0.01	28553.00	0.09	1442720.00	4.30
VanDijk	1954.00	0.01	0	0	1337583.00	4.03
Vd39	1582.00	0.00	71392.00	0.20	3201873.00	8.92
CQ2	3965.00	0.01	54386.00	0.15	3156817.00	8.81
JR2	1954.00	0.01	98739.00	0.27	2841008.00	7.86
VdLs17	0.00	0.00	235025.00	0.65	2875975.00	7.99

Chapter 4

The *Verticillium dahliae* effector Sun1 quantitatively contributes to virulence on sunflower

Jinling Li[1$], Hesham Gibriel[1$], Hélène Missonnier[2], Luigi Faino[1], Xiaoqian Shi-Kunne[1], Michael F. Seidl[1], Alban Jacques[2#], and Bart P.H.J. Thomma[1#]

[1] Laboratory of Phytopathology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
[2] Department of Physiologie, Pathologie et Génétique Végétales (PPGV), Université de Toulouse, INP- PURPAN, Toulouse, France, Syngenta France S.A.S., Saint-Sauveur, France.

[$] These authors contributed equally to this work
[#] These authors contributed equally to this work

To be submitted for publication

ABSTRACT

Plant pathogens employ effector molecules to manipulate host physiology and establish themselves within their hosts. *Verticillium dahliae* is a highly destructive xylem-colonizing fungal pathogen that causes vascular wilt disease on diverse crops, such as tomato (*Solanum lycopersicum*), cotton (*Gossypium hirsutum*), and sunflower (*Helianthus annuus*). In this study, we show that only particular *V. dahliae* strains cause vascular wilt disease on sunflower. Based on comparative genomics of a sunflower pathogenic strain with a diversity of non-pathogenic strains, we identified two candidate effector genes that occur in the sunflower-pathogenic strain 85S and that are highly expressed during host colonization. Intriguingly, these two candidate effector genes appeared to be identical copies that arose by a segmental duplication. Here, we show that this duplicated effector gene quantitatively contributes to *V. dahliae* virulence on sunflower but not on *Nicotiana benthamiana* or *Arabidopsis thaliana*.

INTRODUCTION

Plants and microbes engage in a diverse array of symbiotic relationships, ranging from pathogenic to mutualistic. Pathogenic interactions between plants and microbial pathogens have been described as ongoing arms races in which plants try to halt microbial ingress while pathogens attempt to continue symbiosis (Jones and Dangl 2006; Thomma et al. 2011; Cook et al. 2015). In these arms races, plants have developed various types of immune receptors to detect invading pathogens through sensing various pathogen-derived or pathogen-induced molecular patterns, so-called invasion patterns, that betray microbial invasion to activate appropriate immune responses (Cook et al., 2015; Rodriguez-Moreno et al., 2017). In turn, to enable a parasitic life on their hosts, microbial pathogens secrete so-called effector molecules to deregulate host immune responses and support successful host infection (Jones and Dangl, 2006; Rovenich et al., 2014; Cook et al., 2015). Whereas most extensively studied effectors are proteinaceous molecules, other types of microbially secreted molecules, such as secondary metabolites and small RNAs, have also been described as non-canonical effectors (Wang et al., 2015; Rodriguez-Moreno et al., 2018). Interestingly, ongoing research has revealed that effectors are not exclusively secreted by plant pathogens but other types of symbionts such as endophytes and mutualists similarly employ effectors to establish host interactions (Rovenich et al., 2014). For instance, the mutualistic fungus *Laccaria bicolor* secretes the effector protein MiSSP7 to interact with jasmonic acid (JA) signalling components of the host and facilitate symbiosis by suppressing host immunity (Plett et al., 2014). Moreover, as all symbionts establish themselves in environments that comprise other microbes including antagonists, effector molecules may also act in self-defense and competition with microbiome co-inhabitants (Rovenich et al., 2014; Snelders et al., 2018).

Vascular wilts caused by soil-borne fungal species of the *Verticillium* genus are among the most devastating plant diseases worldwide with an estimated annual loss of €3 billion in the 20 most affected hosts (Depotter et al., 2016). Within the *Verticillium* genus, *V. dahliae* is the most notorious pathogenic species that has the ability to infect over 200 plant species including high-value crops such as tomato (*Solanum lycopersicum*), cotton (*Gossypium hirsutum*), and sunflower (*Helianthus annuus*) (Inderbitzin and Subbarao, 2014). *V. dahliae* invades plant hosts via roots and then colonizes the water-conducting xylem vessels, which disrupts water transport and causes the characteristic wilting symptoms (Fradin and Thomma, 2006; Klosterman et al., 2011). Verticillium wilt disease management is notoriously challenging since conventional fungicides are generally ineffective to eradicate the pathogen once it has entered the xylem tissues (Fradin and Thomma, 2006; Klosterman et al., 2009). Moreover, due to the broad host range of the pathogen and long viability of the resting structures in the soil, cultivation practices such as crop rotation do not result in efficient disease management (Pegg and Brad, 2002; Fradin and Thomma, 2006). Although genetic resistance has been considered as a preferred strategy for disease control, only few Verticillium wilt resistance genes have

been identified so far (Schaible et al., 1951; Simko et al., 2004; Fradin and Thomma, 2006). For example, in tomato, immune receptor Ve1 confers resistance to race 1 *V. dahliae* and *Verticillium albo-atrum* strains through recognition of the race 1-specific effector protein Ave1 (Fradin et al., 2009; de Jonge et al., 2012). Transfer of tomato Ve1 into other crop species like tobacco and cotton can provide resistance against race 1 *Verticillium* strains (Song et al., 2018). The high economic impact of Verticillium wilt disease, combined with the absence of curative treatments, substantiates the need for developing novel disease control strategies. Recently, host-induced gene silencing (HIGS), which involves host expression of double-stranded RNAs to target and silence essential pathogen genes to confer host resistance, has been developed for controlling Verticillium wilt disease in cotton and tomato (Song and Thomma, 2016; Zhang et al., 2016).

To design novel control strategies to combat Verticillium wilt disease, a thorough understanding of molecular mechanisms underlying pathogenesis of *V. dahliae* is of fundamental importance (Fradin and Thomma, 2006). Over the past years, forward genetic approaches such as random mutagenesis and proteomic techniques have been performed to identify potential pathogenicity and virulence factors of *V. dahliae* (El-Bebany et al., 2010; Santhanam, 2014; Chen et al., 2016; Zhang et al., 2017). The ease and low cost of present-day genome sequencing (Faino and Thomma, 2014; Gibriel et al., 2016) have made it possible to sequence multiple *V. dahliae* isolates from various host niches. Using comparative population genomics, we previously identified the *V. dahliae* race-specific effector Ave1 that is crucial for fungal aggressiveness during tomato colonization (de Jonge et al., 2012). Subsequently, Ave1 was demonstrated to contribute to fungal virulence not only on tomato plants that lack the *Ve1* resistance gene, but also on tobacco, Arabidopsis, as well as on cotton (de Jonge et al., 2012; Song et al., 2018). More recently, we applied comparative genomics to identify the *V. dahliae* defoliating (D) pathotype-specific effector (named D effector) that is responsible for defoliation symptoms on cotton and olive (Li et al. unpublished). Interestingly, we found that the D effector acts as a pathogenicity determinant of *V. dahliae* on diverse host species (Li et al. unpublished). Using a similar approach, we uncovered a *V. dahliae* effector (named Tom1) that mediates pathogenicity of *V. dahliae* on tomato (Li et al. unpublished).

Sunflower is a worldwide planted oil crop that is known to maintain stable yields across a wide variety of environmental conditions (Kane and Rieseberg, 2007). However, Verticillium wilt disease constitutes an important constraint for the production of sunflower (Yao et al., 2011; Guo et al., 2017). Despite huge yield losses caused by *V. dahliae*, so far relatively little is known about how the pathogen causes disease on sunflower and no effectors have been characterized for facilitating fungal aggressiveness on sunflower. In this study, comparative genome analysis of a sunflower pathogenic strain with multiple non-pathogenic strains was performed to identify *V. dahliae* effector gene candidates that are essential for virulence on sunflower.

MATERIALS AND METHODS

Fungal strains and plant materials

Verticillium dahliae strains (Table S1) were grown on potato dextrose agar (PDA) at 22°C. Sunflower (*H. annuus* cv. Tutti), *Nicotiana benthamiana*, cotton (*G. hirsutum* cv. Simian3), tomato (*S. lycopersicum* cv. Moneymaker), and *Arabidopsis thaliana* (Col-0) plants were grown under controlled greenhouse conditions (Unifarm, Wageningen, The Netherlands) at 21°C/19°C during 16/8 hours day/night periods, respectively, with 70% relative humidity and 100 W/m^2 supplemental light when the light intensity dropped below 150 W/m^2.

V. dahliae genome sequences and phylogenetic tree construction

Genomes of the four *V. dahliae* strains (JR2, VdLs17, CQ2, and 85S) were sequenced using Single-Molecule Real-Time (SMRT) sequencing technology (Faino et al., 2015; Depotter et al., 2018) (Table S1). Additionally, genomes of *V. dahliae* strains V574 and BP2 were sequenced using the Illumina HiSeq 2000 (Li et al. unpublished) (Table S1). A phylogenetic tree of the *V. dahliae* strains was generated with REALPHY (version 1.12) (Bertels et al., 2014) using Bowtie2 (Langmead and Salzberg, 2012) to map genomic reads against the reference *V. dahliae* strain JR2. A maximum likelihood phylogenetic tree was inferred using RAxML (version 8.2.8) with the GTRGAMMA model and 500 bootstrap replicates (Stamatakis, 2014).

V. dahliae strain 85S effector identification

In order to identify effector genes of *V. dahliae* strain 85S that mediate aggressiveness on sunflower plants, we mapped DNA reads from non-pathogenic strains to the genome assembly of *V. dahliae* strain 85S using BWA (BWA-mem algorithm) (v0.7.12) (Li and Durbin, 2010). Read coverage mapping of all *V. dahliae* strains over the 85S reference genome was calculated in 100 bp windows using BEDTools coverage (v2.25) (Quinlan and Hall, 2010). Genomic regions were considered present if the breadth of coverage ≥50%, while those with breadth of coverage <50% were considered absent. lineage-specific (LS) genomic regions, here defined as genomic regions that are only present in *V. dahliae* strain 85S and absent in all other non-pathogenic *V. dahliae* strains, were determined and genes localized within these LS regions were extracted using BEDtools intersect (v2.25) (Quinlan and Hall, 2010).

A gene annotation for *V. dahliae* strain 85S was generated using the Maker2 pipeline (Holt and Yandell, 2011). To identify potential effector genes of strain 85S, N-terminal signal peptides were first predicted with SignalP (version 4.1) (Petersen et al., 2011). Subsequently, the machine-learning approach applied in EffectorP (version 1.0) (default parameters) was used (Sperschneider et al., 2016). Effector genes localized within LS regions were extracted using BEDtools intersect (Quinlan and Hall, 2010). Syntenies between effector gene copies was analysed using NUCmer (version 3.1) (--maxmatch),

which is part of the MUMer package (Kurtz et al., 2004), and visualized using R package genoPlotR (v0.8.6) (Guy et al., 2010).

Gene expression analysis

To determine expression of *Sun1 in planta*, 12-day-old sunflower seedlings were root-inoculated with conidiospores of *V. dahliae* strain 85S as described previously (Fradin et al., 2009). Stems of inoculated sunflower plants were harvested at 9, 16, 24, 32 and 40 days post inoculation (dpi), flash frozen in liquid nitrogen and stored at -80°C for total RNA extraction. Total RNA extraction and cDNA synthesis were performed as described previously (Song et al., 2016). Quantitative real time PCR was preformed to detect the expression of *Sun1* using primers Sun1-F(RT-PCR) and Sun1-R(RT-PCR), whereas the *V. dahliae GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) gene served as an endogenous control (Table S2).

Generation of *Sun1* deletion strains

To generate single *Sun1* gene deletion constructs, sequence stretches of approximately 1.1 kb upstream and 1.2 kb downstream of the *Sun1* coding sequence were amplified from genomic DNA of *V. dahliae* strain 85S using primer pairs SKO-Sun1-LBF/LBR and SKO-Sun1-RBF/RBR (Table S2). The amplicons were cloned into the pRF-HU2 vector as described previously (Frandsen et al., 2008) and the resulting deletion construct was transformed into strain 85S via *Agrobacterium tumefaciens*-mediated transformation as described earlier (Santhanam, 2012). Putative deletion transformants were selected on PDA supplemented with cefotaxime (200 µg/mL) and hygromycin (50 µg/mL) and homologous gene replacement was verified by PCR analysis using outsider primer F and outsider primer R (Table S2).

To generate *Sun1* double deletion mutants, sequence stretches of approximately 1.3 kb upstream and 1.1 kb downstream of the *Sun1* coding sequence were amplified using primer pairs DKO-Sun1-LBF/LBR and DKO-Sun1-RBF/RBR (Table S2). The amplified products were cloned into vector pRF-NU2. Next, the gene replacement construct was transformed into a *Sun1* single deletion mutant. Putative double deletion transformants were selected on PDA supplemented with nourseothricin (15 µg/mL) and hygromycin (50 µg/mL) and confirmed by PCR analysis using outsider primer F and outsider primer R (Table S2).

Pathogen inoculations and fungal recovery assays

Inoculations were performed on sunflower, cotton, tomato, *N. benthamiana*, and *A. thaliana* plants as previously described (Fradin et al., 2009). Disease symptoms were scored up to 21 (tomato, *A. thaliana*, and *N. benthamiana*), 28 (cotton), or 45 (sunflower) days post inoculation (dpi). Plants were photographed and Image J was used to determine the canopy area while a rectilinear scale was used to measure plant height. Fungal biomass in *A. thaliana*, *N. benthamiana*, and sunflower were determined as previously described (Song et al., 2018). Stems of five inoculated plants were

harvested at 21 dpi (*A. thaliana* and *N. benthamiana*), or 45 dpi (sunflower) for genomic DNA extraction. Real-time quantitative PCR (qPCR) was conducted to quantify fungal colonization (Ellendorff et al., 2009) using the fungus-specific primers ITS-F and ITS-R (Table S2). Primers targeting the sunflower *elongation factor 1-alpha* (*EF-1α*) gene, *A. thaliana RuBisCo*, and *N. benthamiana RuBisCo* gene were used as endogenous plant controls as described earlier (Table S2) (Song et al., 2018).

Fungal recovery assays were conducted to detect *V. dahliae* strains *in planta* as previously described (Fradin et al., 2009). Stem sections were harvested at 45 days post inoculation (dpi), surface-sterilized and sliced into small discs. For each *V. dahliae* strain, 12 to 15 stem slices from five pooled plants transferred onto PDA supplemented with chloramphenicol (35 µg/mL) and incubated at 22°C for 7 days.

RESULTS

Verticillium dahliae inoculations on sunflower

Various strains of *V. dahliae* have been sequenced in our laboratory but their capacity to infect sunflower remained unknown (Klosterman et al., 2011; de Jonge et al., 2013; Thomma unpublished data). Therefore, in addition to the sunflower-pathogenic *V. dahliae* strain 85S, five strains of which the genome has been sequenced were tested for their ability to infect sunflower. These comprise JR2 and VdLs17, for which a gapless genome assembly has previously been generated (Faino et al., 2015) and that are known to cause Verticillium wilt on tomato and lettuce, respectively, and three *V. dahliae* strains that are phylogenetically distant from JR2 and VdLs17, namely the cotton-pathogenic strains CQ2, BP2 and V574 (Figure S1). Interestingly, in contrast to strain 85S that induced clear stunting, chlorosis and wilting on inoculated sunflower plants, JR2, CQ2, VdLs17, BP2 and V574 failed to cause visible disease symptoms, suggesting that these strains do not have the capacity to infect sunflower (Figure 1A-B). This finding was further corroborated by fungal recovery assays, as plating of stem sections of the inoculated plants on agar medium resulted in fungal outgrowth from all sections of 85S-inoculated plants while no fungal growth was recovered from stem sections of plants inoculated with any of the other strains (Figure 1C). Thus, except for *V. dahliae* strain 85S that can be classified as a sunflower pathogenic strain, all other strains (JR2, CQ2, VdLs17, BP2 and V574) were classified as non-pathogenic on sunflower.

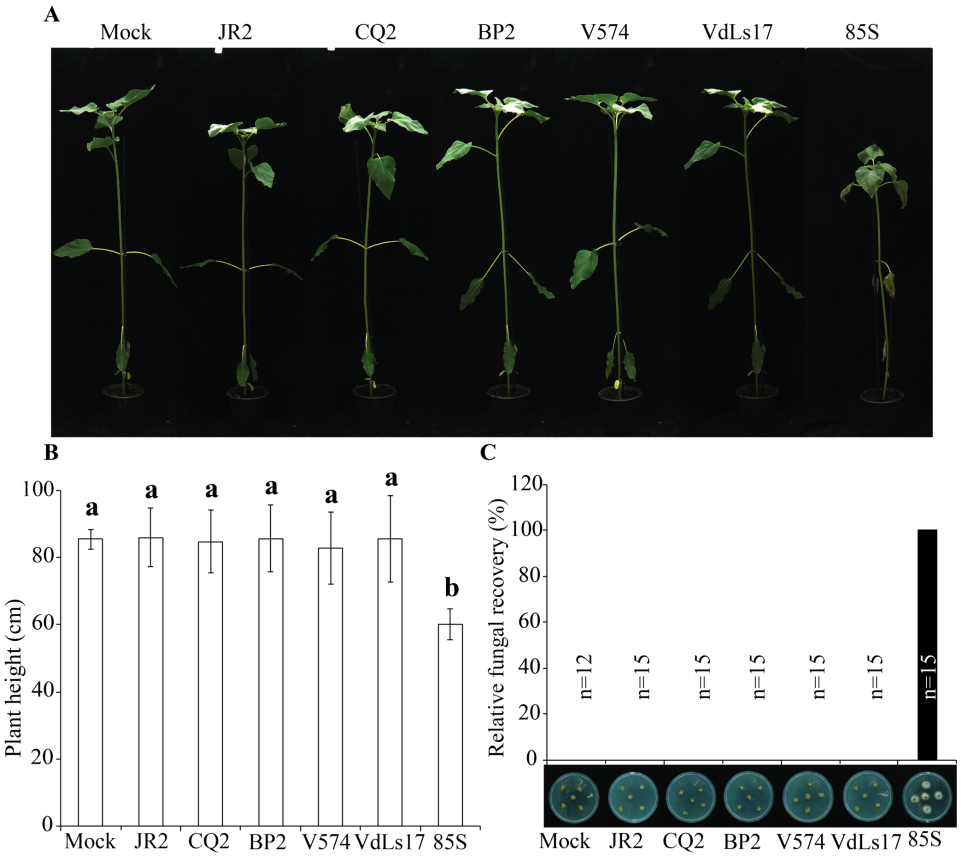


Figure 1. Phenotypes of sunflower plants inoculated with sequenced *Verticillium dahliae* strains.

(A) Side view of sunflower (cv. Tutti) plants at 45 days after mock-inoculation or inoculation with *V. dahliae* strains JR2, CQ2, BP2, V574, VdLs17 and 85S.Besides the stunting, the plant inoculated with strain 85S also displays chlorosis and wilting symptoms. **(B)** Quantification of *V. dahliae*-induced plant stunting at 45 days post inoculation (dpi). Bars represent the average height of five plants with standard deviation. Different letter labels indicate statistically significant differences (Student's *t*-test; $P < 0.05$). **(C)** Fungal outgrowth at 10 days after plating of sunflower stem sections harvested at 45 days post inoculation (dpi) with *V. dahliae* strains JR2, CQ2, BP2, V574, VdLs17 and 85S. The bar graph shows the percentage of stem section slices from which fungal outgrowth was observed. Inoculation experiments were performed with five plants for each fungal strain and independently repeated twice with similar results.

Comparative genomics identifies the effector gene *Sun1*

We aimed to identify effectors that mediate aggressiveness on sunflower plants by comparative genomics of pathogenic strain 85S and the non-pathogenic strains JR2, VdLs17, CQ2, BP2, and V574. To this end, the genomes of *V. dahliae* strains 85S, CQ2, VdLs17, as well as JR2 were sequenced with PacBio technology (Faino et al., 2015; Depotter et al., 2018). In addition, the genomes of BP2 and V574 were sequenced by using the Illumina HiSeq 2000 (Li et al. unpublished).

Gene annotation for *V. dahliae* strain 85S using the Maker2 pipeline (Holt and Yandell, 2011) yielded a total of 10,580 protein-coding genes. Of these genes, 1,062 genes were predicted to encode putative secreted proteins. Subsequently, we used the machine-learning approach applied in EffectorP (Sperschneider et al., 2016) to identify 174 candidate effectors for *V. dahliae* strain 85S. To identify 85S-specific effector genes, a reference-based alignment strategy was used. All reads of the non-pathogenic *V. dahliae* strains were mapped onto the genome assembly of strain 85S and lineage specific (LS) regions for which no synteny occurred in any of the non-pathogenic strains were extracted. This comparative analysis revealed 6,924 LS sequences of 100 bp in size, collectively comprising 159 genes including five that encode secreted proteins. Of these five genes, two were classified as effector genes based on EffectorP. Remarkably, both effector genes were found to encode a putative effector of 125 amino acids and share 100% sequence identity, despite the observation that they are located on two different contigs, namely unitig_4 (292.60 kb) and unitig_14 (693.13 kb) (Figure 2A). Moreover, further alignment revealed that both contigs are highly syntenic (99% identity) (Figure 2B), suggesting that they arose from a recent segmental duplication. We named the putative effector genes *Sun1-a* and *Sun1-b*, as candidates for mediating virulence on **sun**flower. Similar to most fungal effectors reported so far (Stergiopoulos and de Wit, 2009; de Jonge et al., 2011), no Sun1 homologs were found in other species.

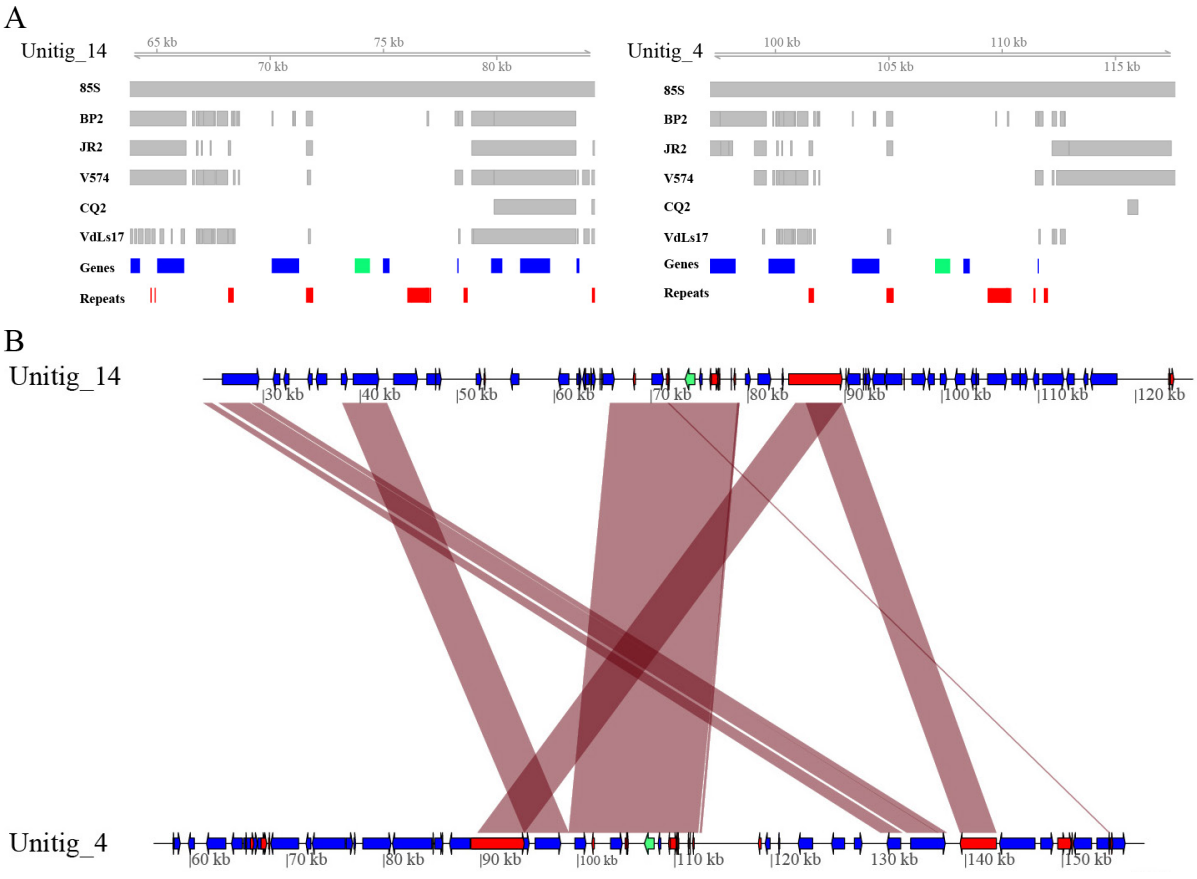


Figure 2. Comparative genomics of *V. dahliae* strains identifies the *Sun1* effector gene.
(A) The alignment of reads derived from non-pathogenic *V. dahliae* strains to the sunflower-pathogenic *V. dahliae* strain 85S reveals two LS regions of ~4 kb on contigs unitig_14 and unitig_4. Note, besides other genes (blue) these two LS regions harbor the effector

gene *Sun1* (green) that is flanked by repeats (red), and solely present in the strain 85S. **(B)** Synteny analyses of unitig_4 and unitig_14 harboring the *Sun1* effector gene. A region encompassing 50 kb upstream and downstream of *Sun1* effector gene is shown. Highly syntenic regions (99% identity) upstream and downstream of this effector gene on unitig_14 and unitig_4 are indicated with red ribbons.

To further assess the potential of Sun1 as a virulence effector on sunflower, the transcript level of *Sun1* during infection of *V. dahliae* strain 85S on sunflower was investigated using real-time PCR in a time course experiment up to 40 days post inoculation (dpi). This analysis revealed that *Sun1* transcript levels increased significantly at 9 days post-inoculation (dpi) and remained high at later time points (Figure 3).

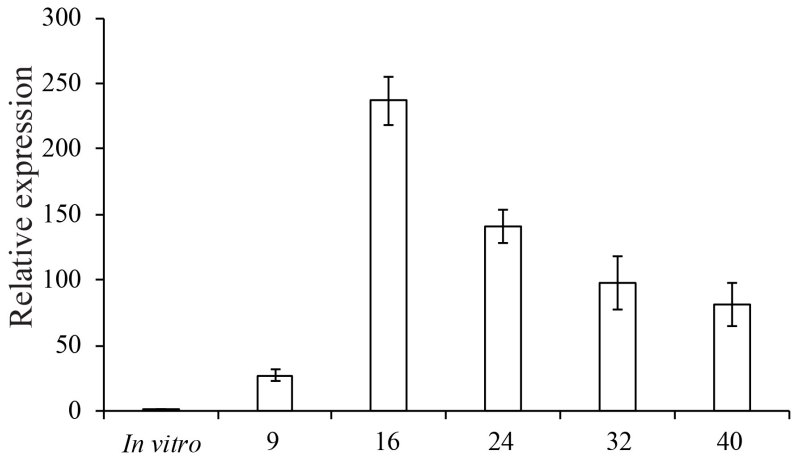


Figure 3. *Sun1* is highly induced during infection of *Verticillium dahliae* on sunflower. Twelve day-old sunflower (cv. Tutti) seedlings were root-inoculated with *V. dahliae* strain 85S and stems were harvested at 9, 16, 24, 32 and 40 days post inoculation (dpi). After RNA isolation and cDNA synthesis, real-time PCR was performed to determine the relative expression level of *Sun1* using the *V. dahliae GAPDH* gene as a reference. Expression of *Sun1 in vitro* (potato dextrose agar) was set to 1.

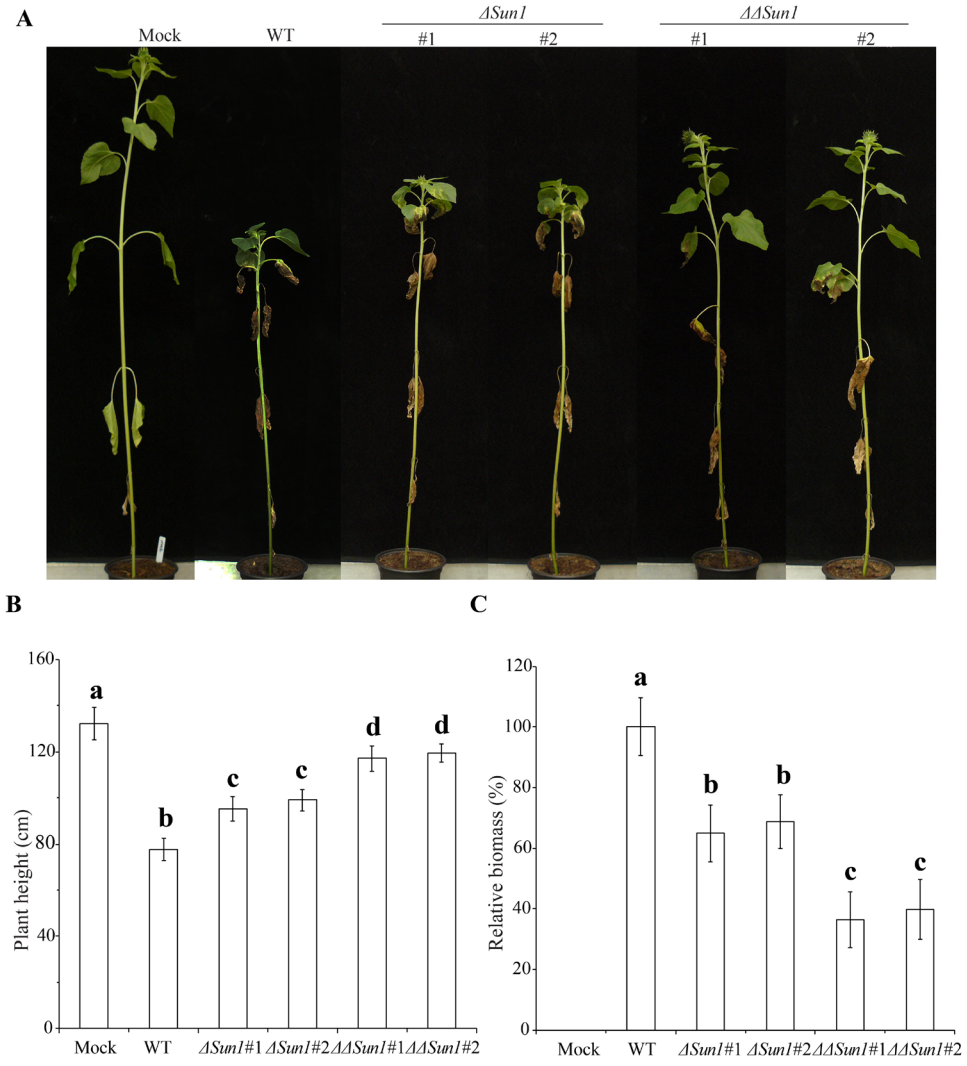


Figure 4. *Sun1* is required for full virulence on sunflower.
(A) Typical phenotype of sunflower (cv. Tutti) upon mock-inoculation or inoculation with wild-type strain 85S (WT), two *Sun1* single deletion strains (*ΔSun1*#1 and *ΔSun1*#2) and two *Sun1* double deletion strains (*ΔΔSun1*#1 and *ΔΔSun1*#2). Photographs were taken at 45 days post inoculation (dpi). **(B)** Quantification of *V. dahliae*-induced plant stunting at 45 dpi. Bars represent the average height of five plants with standard deviation. **(C)** Fungal biomass as determined with real-time PCR at 45 dpi. Bars represent *V. dahliae ITS* levels relative to sunflower *elongation factor 1-α* levels (for equilibration) with standard deviation in a sample of five pooled plants. The fungal biomass in sunflower plants upon inoculation with

the wild-type strain 85S is set to 100%. Different letter labels indicate statistically significant differences (Student's *t*-test; $P < 0.05$). Inoculation experiments were performed with five plants for each fungal strain and independently repeated twice with similar results.

The Sun1 effector quantitatively contributes to virulence on sunflower

To further assess the contribution of the Sun1 effector to virulence of *V. dahliae* strain 85S on sunflower, we generated deletion mutants for both effector gene copies through homologous recombination (Figure S2A). To this end, we first generated single gene copy deletion mutants, followed by deletion of the second *Sun1* gene copy. Putative single and double gene deletion mutants were confirmed by PCR (Figure S2B). Thereby, we also confirmed that the predicted presence of the two gene copies is genuine and not the consequence of an assembly artefact (Figure S2B). Two independent single (*ΔSun1*#1 and *ΔSun1*#2) and two independent double (*ΔΔSun1*#1 and *ΔΔSun1*#2) gene deletion mutants were selected for subsequent inoculation assays.

Single deletion (*ΔSun1*) mutants exhibited markedly reduced virulence on sunflower when compared with wild-type strain 85S at 45 days post inoculation (dpi) (Figure 4A), demonstrated by significantly reduced stunting and compromised fungal colonization of the host plants (Figure 4B-C). Interestingly, sunflower plants that were inoculated with double deletion (*ΔΔSun1*) mutants showed a similar disease phenotype as plants upon inoculation with *ΔSun1* mutants during relatively early stages of infection up until 24 dpi, as no significant difference in disease symptoms and plant height reduction was observed. However, from 32 dpi onwards, sunflower plants that were inoculated with *ΔΔSun1* mutants showed significantly less wilting symptoms and stunting than plants inoculated with *ΔSun1* mutants (Figure 4B). Moreover, real-time PCR quantification of fungal biomass demonstrated that *ΔΔSun1* mutants accumulated significantly less fungal biomass during infection than *ΔSun1* mutants and the wild type strain (Figure 4C). Taken together, these data demonstrate that the Sun1 effector quantitatively contributes to virulence on sunflower plants.

The Sun1 effector is dispensable for virulence on *Nicotiana benthamiana* and *Arabidopsis thaliana*

To investigate whether the observed role of *Sun1* in virulence is confined to sunflower only or also extends to other host species, we first tested the virulence of the *V. dahliae* strain 85S on the Solanaceous crop tomato, the model plant *N. benthamiana*, the Malvaceae crop cotton and the Brassicaceous model plant *A. thaliana*. While 85S failed to cause wilt disease on tomato (Figure S3) and caused only mild wilt disease symptom on cotton (Figure S4), clear symptoms of disease were observed on *A. thaliana* and *N. benthamiana* upon inoculation. Next, we examined the virulence of *Sun1* deletion mutants (*ΔSun1* and *ΔΔSun1*) on *A. thaliana* and *N. benthamiana* to evaluate the contribution of this effector to virulence on these two hosts. However, unlike our observations on sunflower, deletion of *Sun1* did not result in significantly compromised virulence on *N. benthamiana*, as similar Verticillium wilt symptoms, including wilting

and stunting, were observed for *ΔSun1* and *ΔΔSun1* mutants as for the wild-type strain 85S at 3 weeks post inoculation (Figure 5A, C). Real time PCR quantification of fungal biomass confirmed that all strains colonized *N. benthamiana* plants to a similar extent (Figure 5D). Similarly, Sun1 effector was also not found to contribute to virulence on *A. thaliana* (Figure 5B, E-F). Therefore, we conclude that the Sun1 effector does not play a general role in virulence, but its contribution appears to be confined to virulence on sunflower.

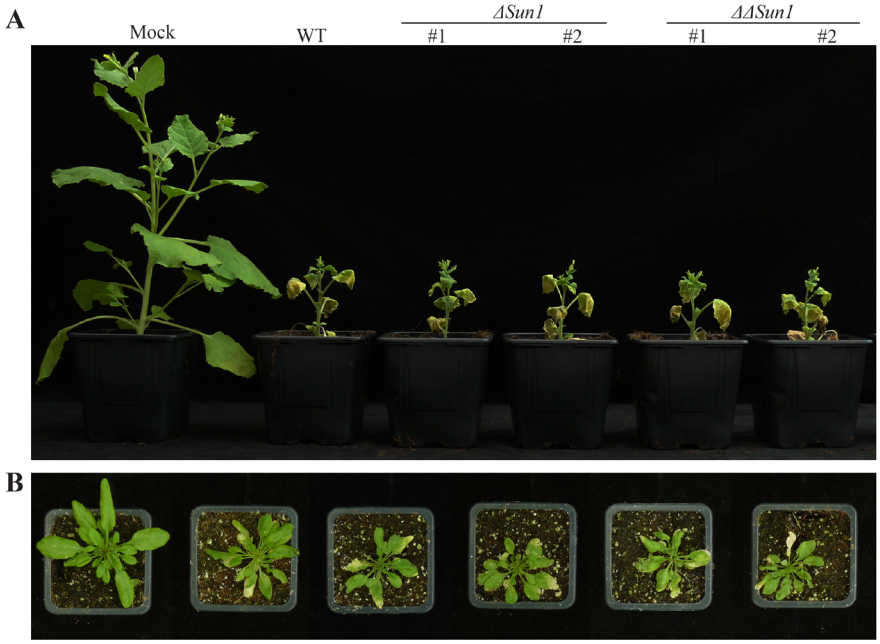


Figure 5. *Sun1* is dispensable for virulence on *Nicotiana benthamiana* and *Arabidopsis thaliana*.

(A) Typical phenotype of *N. benthamiana* plants that were mock-inoculated or inoculated with wild type strain 85S (WT), two *Sun1* single deletion strains (*ΔSun1*#1 and *ΔSun1*#2) and two *Sun1* double deletion strains (*ΔΔSun1*#1 and *ΔΔSun1*#2). **(B)** Typical phenotype of *A. thaliana* (Col-0) plants that were mock-inoculated or inoculated with indicated fungal strains in panel A at 21 (dpi).

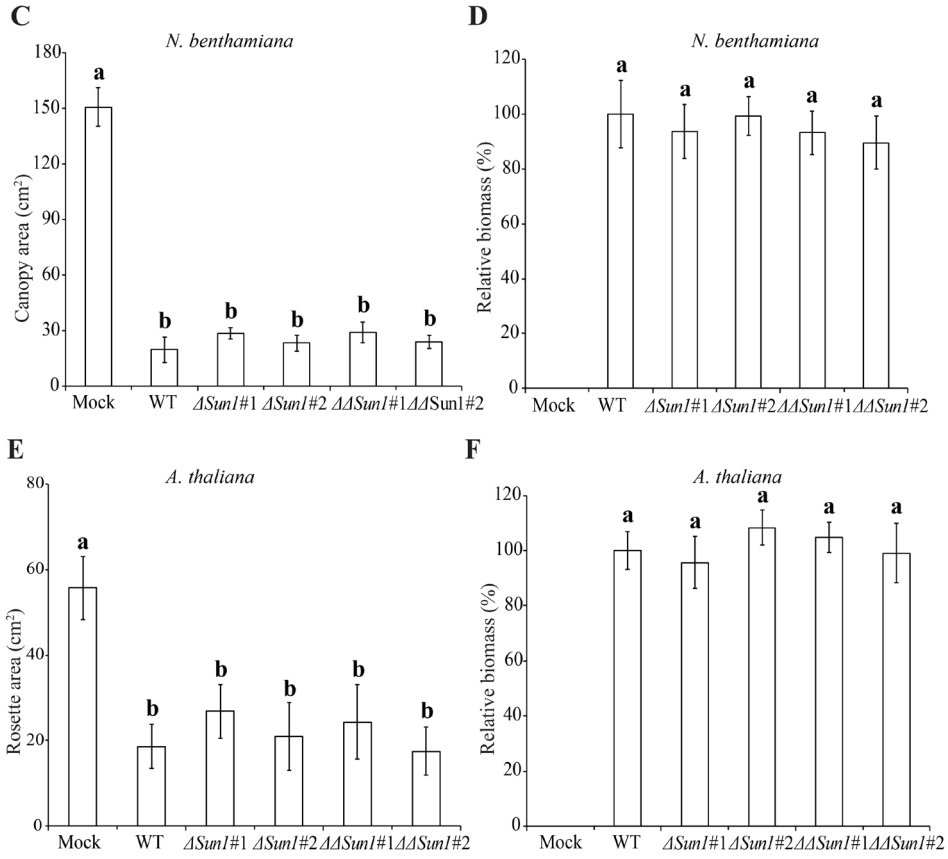


***Figure 5 continued. Sun1* is dispensable for virulence on *Nicotiana benthamiana* and *Arabidopsis thaliana.* (C)** Quantification of the canopy area of *N. benthamiana* at 21 dpi. Bars represent the average canopy area of five plants with standard deviation. Different letter labels indicate statistically significant differences (Student's *t*-test; $P < 0.05$). **(D)** Fungal biomass as determined with real-time PCR at 21 dpi. Bars represent *V. dahliae ITS* levels relative to *N. benthamiana RuBisCo* levels (for equilibration) with standard deviation in a sample of five pooled plants. The fungal biomass in *N. benthamiana* plants upon inoculation with the wild-type strain 85S is set to 100%. The same letter labels indicate no statistically significant differences (Student's *t*-test; $P < 0.05$). **(E)** Quantification of the rosette area of five *A. thaliana* plants at 21 dpi. Bars represent the average rosette area of five plants with standard deviation. Different letter labels indicate statistically significant differences (Student's *t*-test; $P < 0.05$). **(F)** Fungal biomass as determined with real-time PCR at 21 dpi. Bars represent *V. dahliae ITS* levels relative to *A. thaliana RuBisCo* levels (for equilibration) with standard deviation in a sample of five pooled plants. The fungal biomass in *A. thaliana* plants upon inoculation with the wild-type strain 85S is set to 100%. The same letter labels

indicate no statistically significant differences (Student's *t*-test; $P < 0.05$). Photographs were taken at 21 days post inoculation (dpi). Inoculation experiments were performed with five plants for each fungal strain and independently repeated twice with similar results.

DISCUSSION

While *V. dahliae* strains are typically characterized by their broad host range, pathogenic isolates still display differential host specificity as individual isolates only infect a limited number of plant species (Bhat and Subbarao, 1999). Over the years, a number of studies have suggested that the host range of a given pathogen is largely governed by the presence of host-specific virulence genes that contribute to disease establishment (van der Does and Rep, 2007). Like many filamentous pathogens, *V. dahliae* also employs effector molecules as virulence factors to establish disease (Klimes et al., 2015). To date, a number of *V. dahliae* effectors have been reported to contribute to disease establishment on various hosts. For example, the effector proteins Ave1 and NLP1 (necrosis-and ethylene-inducing-like protein 1) have been shown to function as virulence factors on multiple host plants, such as tomato, *N. benthamiana*, as well as *A. thaliana* (de Jonge et al., 2012; Song et al., 2018; Santhanam et al., 2013). The recently identified D effector was shown to act as a pathogenicity factor of *V. dahliae* defoliating (D) pathotype strains on cotton, olive, *N. benthamiana* and *A. thaliana* (Li et al. unpublished). In contrast, other *V. dahliae* effectors may promote fungal virulence only on particular host species or genotypes. For instance, the chitin-binding lysin motif (LysM) effector Vd2LyM was reported to contribute to fungal virulence on tomato, but not on *N. benthamiana* or *A. thaliana* (Kombrink et al., 2017). Similarly, the effector NLP2 is required for full virulence of *V. dahliae* strain JR2 on tomato and *A. thaliana*, but not on *N. benthamiana* (Santhanam et al., 2013). In addition, the Tom1 effector was found to function as a pathogenicity effector that determines host specificity of *V. dahliae* on tomato only (Li et al. unpublished). Similarly, in this study we observed that the Sun1 effector quantitatively contributes to fungal aggressiveness on sunflower (Figure 4), but not on *N. benthamiana* or *A. thaliana* (Figure 5). Thus, we anticipate that the Sun1 effector may specifically facilitate fungal virulence on sunflower, but future experiments on other hosts are needed to confirm this hypothesis.

Various mechanisms have been described that can facilitate the development of effector gene repertoires in pathogenic microbes, such as genome hybridization (Stukenbrock et al., 2012), gene duplication (Dutheil et al., 2016), and horizontal gene transfer (HGT) (Friesen et al., 2006; de Jonge et al., 2012). It has been suggested that the well-characterized *V. dahliae* LS effector Ave1 has been acquired from plants via HGT (de Jonge et al., 2012). Genomic comparisons of multiple *V. dahliae* strains revealed that all *V. dahliae* strains carry highly variable LS genomic regions, accounting for 1-5 Mb of their ~35-Mb genome, that are unique or shared by only a sub-set of *V. dahliae* isolates (de Jonge et al., 2013; Faino et al., 2015, 2016). Interestingly, numerous *in-planta*-

induced effector genes reside in LS genomic regions that largely consist of segmental genomic duplications, suggesting that gene duplications may play an important role in the emergence of effector genes in *V. dahliae* (Jonge et al., 2013; Faino et al., 2016). In the present study, we show that two identical *Sun1* effector genes arose by a segmental genomic duplication event and the high level of similarity between flanking sequences of the *Sun1* effector gene suggests that this duplication occurred rather recently (Figure 2). In line with this finding, two exact copies of the *D* effector gene in *V. dahliae* D pathotype strains have emerged by a recent segmental duplication as well (Li et al. unpublished). The relevance of the occurrence of two identical effector gene copies and their impact on fungal adaption remains unknown at this point. Possibly, the emergence of two copies of the effector gene is relevant to maintain fungal aggressiveness on host plants. In this light it is interesting to note that the functionality of the two copies is not redundant, as they both quantitatively contribute to virulence and, consequently, deletion of a single copy markedly affects fungal virulence.

Recent duplications of effector genes can be subjected to subsequent evolutionary diversification, leading to novel or altered functionality of one of the two gene copies (Sanchez-Vallet et al., 2018). Thus, effector genes of filamentous pathogens that arise from gene duplication events typically evolve in a so called “duplication-divergence” pattern: following a gene duplication event, one gene copy diverges to some extent due to functional redundancy and evolves a distinct function (Plissonneau et al., 2017). For example, gene duplications followed by sequence divergence were proposed to be responsible for the generation of novel effector genes in the smut fungus *Ustilago maydis* (Dutheil et al., 2016). Similarly, a large number of effector genes of the oomycete pathogen *Phytophthora sojae* underwent sequence diversification after gene duplication (Shen et al., 2013). In addition to experiencing functional diversification, the recent duplications of effector genes may also be subject to differential loss of the duplicated gene copies (Dong et al., 2015; Pedersen et al., 2012). Frequent effector gene losses after segmental duplications have been proposed to occur in the powdery mildew fungus *Blumeria graminis*, which contribute to the diversity of the effector repertoires of the pathogen (Wicker et al., 2013; Menardo et al., 2017). Possibly, selective forces from host immune systems contribute to this process. However, it is unclear whether either of the two copies of the *Sun1* effector gene will experience sequence divergence or gene loss over time.

ACKNOWLEDGEMENTS

J.L. acknowledges receipt of a PhD fellowship from the China Scholarship Council (CSC). MFS acknowledges the receipt of a VENI grant of ALW-NWO. B.P.H.J.T. is supported by the Research Council for Earth and Life sciences (ALW) of the Netherlands Organization for Scientific Research (NWO).

REFERENCES

Bhat RG, Subbarao KV. 1999. Host range specificity in *Verticillium dahliae*. Phytopathology 89: 1218-1225.

Chen JY, Xiao HL, Gui YJ, Zhang DD, Li L, Bao YM, Dai XF. 2016. Characterization of the *Verticillium dahliae* exoproteome involves in pathogenicity from cotton-containing medium. Frontiers in Microbiology 7: 1709.

Cook DE, Mesarich CH, Thomma BPHJ. 2015. Understanding plant immunity as a surveillance system todetect invasion. Annual Review of Phytopathology 53: 541-563.

de Jonge R, Bolton MD, Thomma BPHJ. 2011. How filamentous pathogens co-opt plants: the ins and outs of fungal effectors. Current Opinion in Plant Biology 14: 400-406.

de Jonge R, Bolton MD, Kombrink A, van den Berg GC, Yadeta KA, Thomma BPHJ. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Research 23: 1271-1282.

de Jonge R, van Esse HP, Maruthachalam K, Bolton MD, Santhanam P, Saber MK, Zhang Z, Usami T, Lievens B, Subbarao KV, Thomma BPHJ. 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. Proceedings of the National Academy of Sciences of the United States of America 109: 5110-5115.

Depotter JR, Deketelaere S, Inderbitzin P, Tiedemann AV, Höfte M, Subbarao KV, Wood TA and Thomma BPHJ. 2016. *Verticillium longisporum*, the invisible threat to oilseed rape and other brassicaceous plant hosts. Molecular Plant Pathology 17: 1004-1016.

Depotter JRL, Shi-Kunne X, Missonnier H, Liu T, Faino L, van den Berg GCM, Wood TA, Zhang B, Jacques A, Seidl MF, Thomma BPHJ. 2018. Dynamic virulence-related regions of the fungal plant pathogen *Verticillium dahliae* display remarkably enhanced sequence conservation. bioRxiv doi: https://doi.org/10.1101/277558.

Dong S, Raffaele S, Kamoun S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Current Opinion in Genetics and Development 35: 57-65.

Dutheil JY, Mannhaupt G, Schweizer G, MK Sieber C, Münsterkötter M, Güldener U, Schirawski J, Kahmann RA. 2016. A tale of genome compartmentalization: the evolution of virulence clusters in smut fungi. Genome Biology and Evolution 8: 681-704.

Ellendorff U, Fradin EF, De Jonge R, Thomma BPHJ. 2009. RNA silencing is required for Arabidopsis defence against *Verticillium* wilt disease. Journal of Experimental Botany 60: 591-602.

El-Bebany AF, Rampitsch C, Daayf F. 2010. Proteomic analysis of the phytopathogenic soilborne fungus *Verticillium dahliae* reveals differential protein expression in isolates that differ in aggressiveness. Proteomics 10: 289-303.

Faino L, Seidl MF, Datema E, van den Berg GC, Janssen A, Wittenberg AH, Thomma BPHJ. 2015. Single-molecule real-time sequencing combined with optical mapping yields completely finished fungal genome. mBio 6: e00936-00915.

Faino L, Seidl MF, Shi-Kunne X, Pauper M, van den Berg GC, Wittenberg AH, Thomma BPHJ. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen.Genome Research 26: 1091-1100.

Faino L, Thomma BPHJ. 2014. Get your high-quality low-cost genome sequence. Trends in Plant Science 19: 288-291.

Fradin EF, Thomma BPHJ. 2006. Physiology and molecular aspects of *Verticillium* wilt diseases caused by *V. dahliae* and *V. albo-atrum*. Molecular Plant Pathology 7: 71-86.

Fradin EF, Zhang Z, Ayala JCJ, Castroverde CD, Nazar RN, Robb J, Liu C-M, Thomma BPHJ. 2009. Genetic dissection of Verticillium wilt resistance mediated by tomato Ve1. Plant Physiology 150: 320-332.

Frandsen RJ, Andersson JA, Kristensen MB, Giese H. 2008. Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi. BMC Molecular Biology 9: 70.

Friesen TL, Stukenbrock EH, Liu Z, Meinhardt S, Ling H, Faris JD, Rasmussen JB, Solomon PS, McDonald BA, Oliver RP. 2006. Emergence of a new disease as a result of interspecific virulence gene transfer. Nature Genetics 38: 953-956.

Gibriel, H. A. Y., Thomma, B. P. H. J., and Seidl, M. F. 2016. The age of effectors: genome-based discovery and applications. Phytopathology 106: 1206–1212.

Guo S, Zuo Y, Zhang Y, Wu C, Su W, Jin W, Yu H, An Y, Li Q. 2017. Large-scale transcriptome comparison of sunflower genes responsive to *Verticillium dahliae*. BMC Genomics 18: 42.

Guy L, Kultima JR, Andersson SG. 2010. genoPlotR: comparative gene and genome visualization in R. Bioinformatics 26: 2334-2335.

Holt C, Yandell M (2011) MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12: 491.

Inderbitzin P, Subbarao KV. 2014. Verticillium systematics and evolution: how confusion impedes Verticillium wilt management and how to resolve it. Phytopathology 104: 564–574.

Jones JD, Dangl JL. 2006. The plant immune system. Nature 444: 323-329.

Kane NC, Rieseberg LH. 2007. Selective sweeps reveal candidate genes for adaptation to drought and salt tolerance in common sunflower, *Helianthus annuus*. Genetics 175: 1823-1834.

Klimes A, Dobinson KF, Klosterman SJ, Thomma BPHJ. 2015. Genomics spurs rapid advances in ourunderstanding of the basic biology of vascular wilt pathogens in the genus *Verticillium*. Annual Review of Phytopathology 53: 181-198.

Klosterman SJ, Atallah ZK, Vallad GE, Subbarao KV. 2009. Diversity, pathogenicity, and management of *Verticillium* species. Annual Review of Phytopathology 47: 39-62.

Klosterman SJ, Subbarao KV, Kang S, Veronese P, Gold SE, Thomma BPHJ, Chen Z, Henrissat B, Lee YH, Park J. 2011. Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. PLoS Pathogens 7: e1002137.

Kombrink A, Rovenich H, Shi-Kunne X, Rojas-Padilla E, van den Berg G, Domazakis E, De Jonge R,Valkenburg DJ, Sánchez-Vallet A, Seidl MF, Thomma BPHJ. 2017. *Verticillium dahliae* LysM effectors differentially contribute to virulence on plant hosts. Molecular Plant Pathology 18: 596-608.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. Genome Biology 5: R12.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26: 589-595.

Menardo F, Praz CR, Wicker T, Keller B. 2017. Rapid turnover of effectors in grass powdery mildew (*Blumeria graminis*). BMC Evolutionary Biology 17: 223.

Pedersen C, Ver Loren van Themaat E, McGuffin LJ, Abbott JC, Burgis TA, Barton G, Bindschedler LV, Lu X, Maekawa T, Wessling R, Cramer R, Thordal-Christensen H, Panstruga R, Spanu PD. 2012. Structure and evolution of barley powdery mildew effector candidates. BMC Genomics 13: 694.

Pegg GF, Brady BL. 2002. Verticillium wilts. Wallingord, Oxfordhire: CABI Publishing.

Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods 8: 785-786.

Plett JM, Daguerre Y, Wittulsky S, Vayssières A, Deveau A, Melton SJ, Kohler A, Morrell-Falvey JL, Brun A, Veneault-Fourrey C, Martin F .2014. Effector MiSSP7 of the mutualistic fungus *Laccaria bicolor* stabilizes the Populus JAZ6 protein and represses jasmonic acid (JA) responsive genes. Proceedings of the National Academy of Sciences of the United States of America 111: 8299-8304.

Plissonneau C, Benevenuto J, Mohd-Assaad N, Fouché S, Hartmann FE, Croll D. 2017. Using population and comparative genomics to understand the genetic basis of effector-driven fungal pathogen evolution. Frontiers in Plant Sciences 8: 119.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26: 841-842.

Rovenich H, Boshoven JC, Thomma BPHJ. 2014. Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. Current Opinion in Plant Biology 20: 96-103.

Rodriguez-Moreno L, Ebert MK, Bolton MD, Thomma BPHJ. 2018. Tools of the crook-infection strategies of fungal plant pathogens. The Plant Journal 93: 664-674.

Sánchez-Vallet A, Fouché S, Fudal I, Hartmann FE, Soyer JL, Tellier A, Croll D. 2018. The genome biology of effector gene evolution in filamentous plant pathogens. Annual Review of Phytopathology 56: 1.

Santhanam P. 2012. Random insertional mutagenesis in fungal genomes to identify virulence factors. Plant Fungal Pathogens Springer (Methods and Protocols) 835: 509-517.

Santhanam P. 2014. Insertional mutagenesis in the vascular wilt pathogen *Verticillium dahliae*. Wageningen University PhD thesis.

Santhanam P, van Esse HP, Albert I, Faino L, Nürnberger T, Thomma BPHJ. 2013. Evidence forfunctional diversification within a fungal NEP1-like protein family. Molecular Plant-Microbe Interactions 26: 278-286.

Schaible L, Cannon OS, Waddoups V. 1951. Inheritance of resistance to Verticillium wilt in a tomato cross. Phytopathology 41: 986-990.

Shen D, Liu T, Ye W, Liu L, Liu P, Wu Y, Wang Y, Dou D. 2013. Gene duplication and fragment recombination drive functional diversification of a superfamily of cytoplasmic effectors in *Phytophthora sojae*. PLoS One 8: e70036.

Simko I, Costanzo S, Haynes K, Christ B, Jones R. 2004. Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. Theoretical and Applied Genetics 108: 217-224.

Snelders NC, Kettles GJ, Rudd JJ, Thomma BPHJ. 2018. Plant pathogen effector proteins as manipulators of host microbiomes? Molecular Plant Pathology 19: 257-259.

Song Y, Liu L, Wang Y, Valkenburg DJ, Zhang X, Zhu L, Thomma BPHJ. 2018. Transfer of tomato immune receptor Ve1 confers Ave1-dependent *Verticillium* resistance in tobacco and cotton. Plant Biotechnology Journal 16: 638-648.

Song Y, Thomma BPHJ. 2016. Host-induced gene silencing compromises Verticillium wilt in tomato and Arabidopsis. Molecular Plant Pathology 19: 77-89.

Sperschneider J, Gardiner DM, Dodds PN, Tini F, Covarelli L, Singh KB, Manners JM, Taylor JM. 2016. EffectorP: predicting fungal effector proteins from secretomes using machine learning. New Phytologist 210: 743-761.

Stamatakis, A. 2014. Raxml version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312-1313.

Stergiopoulos I, de Wit PJ. 2009. Fungal effector proteins. Annual Review of Phytopathology 47: 233-263.

Stukenbrock EH, Christiansen FB, Hansen TT, Dutheil JY, Schierup MH. 2012. Fusion of two divergent fungal individuals led to the recent emergence of a unique widespread pathogen species. Proceedings of the National Academy of Sciences of the United States of America 109: 10954–10959.

Thomma BP, Nürnberger T, Joosten MH. 2011. Of PAMPs and effectors: the blurred PTI-ETI dichotomy. Plant Cell 23: 4-15.

van der Does HC, Rep M. 2007. Virulence genes and the evolution of host specificity in plant-pathogenic fungi. Molecular Plant-Microbe Interactions 20: 1175-1182.

Wang M, Weiberg A, Jin H. 2015. Pathogen small RNAs: a new class of effectors for pathogen attacks. Molecular Plant Pathology 16: 219-223.

Wicker T, Oberhaensli S, Parlange F, Buchmann JP, Shatalina M, Roffler S, Ben-David R, Doležel J, Šimková H, Schulze-Lefert P, Spanu PD, Bruggmann R, Amselem J, Quesneville H, Ver Loren van Themaat E, Paape T, Shimizu KK, Keller B. 2013. The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. Nature Genetics 45: 1092-1096.

Yao Z, Rashid KY, Adam LR, Daayf F. 2011. *Verticillium dahliae*'s VdNEP acts both as a plant defense elicitor and a pathogenicity factor in the interaction with *Helianthus annuus*. Canadian Journal of Plant Pathology 33: 375–388.

Zhang L, Ni H, Du X, Wang S, Ma XW, Nürnberger T, Guo HS, Hua C. 2017. The *Verticillium*-specific protein VdSCP7 localizes to the plant nucleus and modulates immunity to fungal infections. New Phytologist 215: 368-381.

Zhang T, Jin Y, Zhao JH, Gao F, Zhou BJ, Fang YY, Guo HS. 2016. Host-induced gene silencing of the target gene in fungal cells confers effective resistance to the cotton wilt disease pathogen *Verticillium dahliae*. Molecular Plant 9: 939-942.

SUPPLEMENTAL INFORMATION

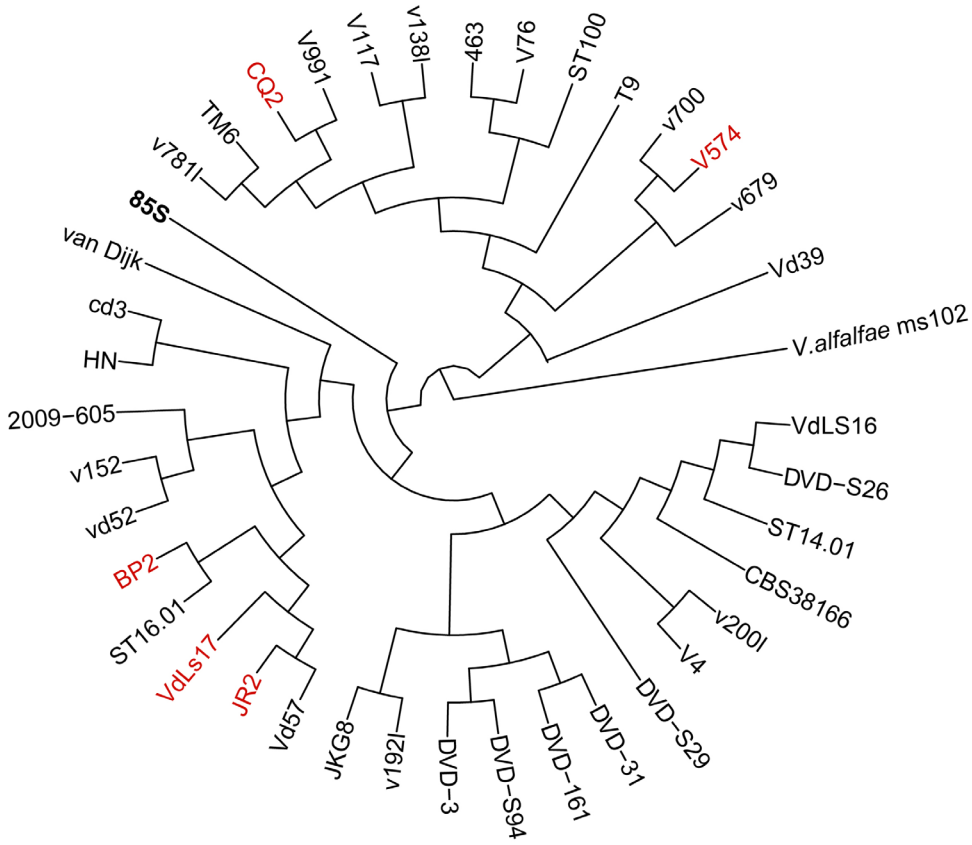


Figure S1. Phylogenetic tree of *V. dahliae* strains.
Sunflower-pathogenic and non-pathogenic strain 85S is shown in bold black. Strains that were selected for phenotypic characterization in this study are shown in red font. Phylogenetic relationships between sequenced *V. dahliae* strains are inferred using RealPhy (Bertels et al., 2014). *V. alfalfa* strain ms102 was used to root the tree.

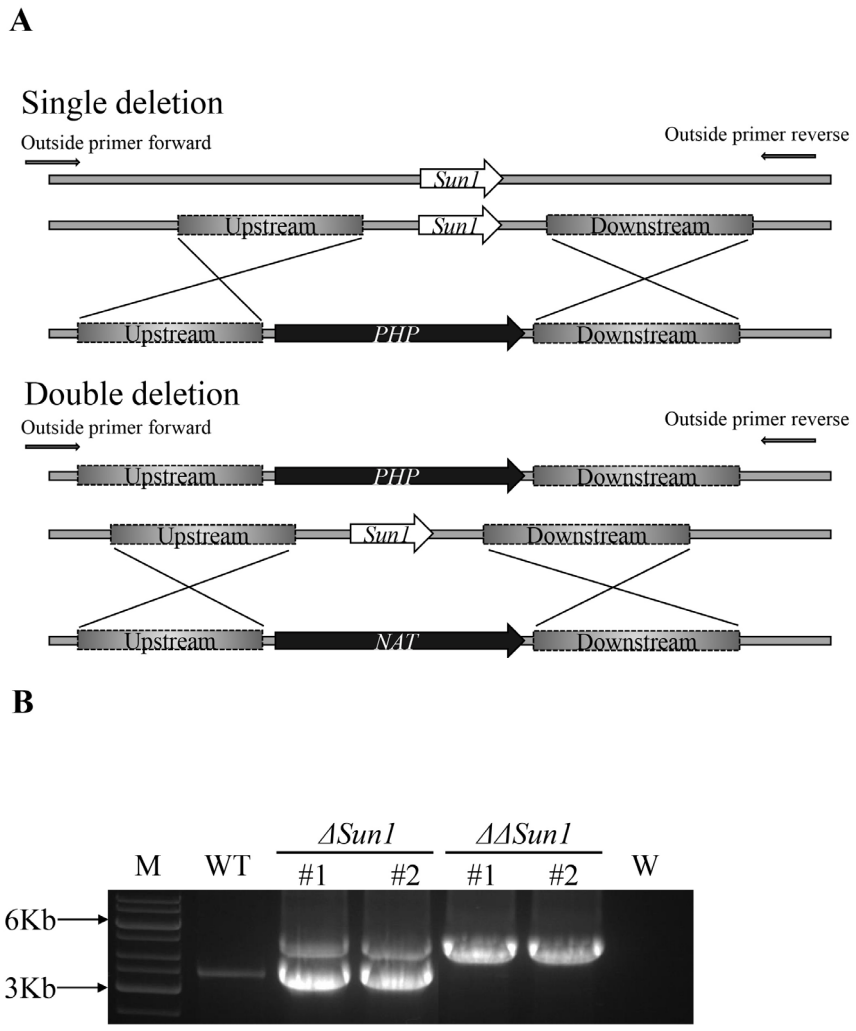


Figure S2. Construction and verification of *Sun1* single deletion and double deletion mutants.

(A) Schematic representation of the homologous recombination events to establish targeted replacement of *Sun1* with phosphotransferase *(HPH)* and the nourseothricin resistance gene cassette (NAT). **(B)** Verification of *Sun1* single deletion and double deletion strains by PCR. Amplicons generated with outside primers indicated in panel A are shown for wild-type strain 85S (WT), two *Sun1* single deletion strains (*ΔSun1*#1 and *ΔSun1*#2) and two *Sun1* double deletion strains (*ΔΔSun1*#1 and *ΔΔSun1*#2). Water was used as negative control (W).

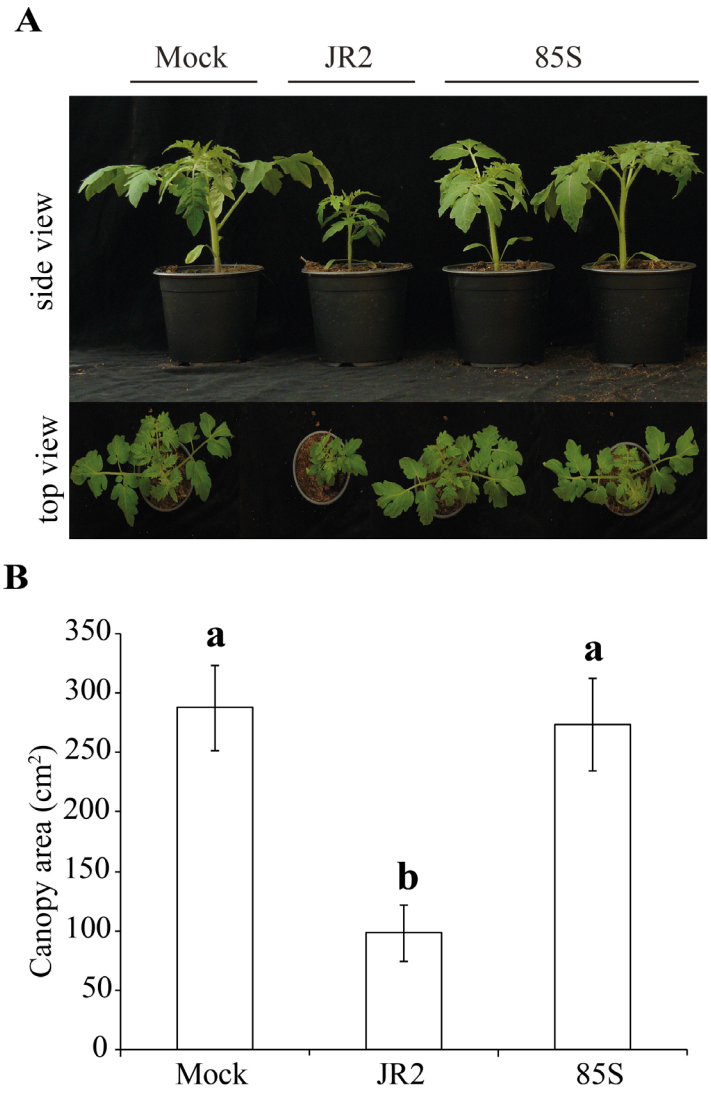


Figure S3. Phenotypes of tomato plants inoculated with *V. dahliae* strain 85S.
(A) Typical phenotype of tomato(cv. Moneymaker) plants upon mock-inoculation or inoculation with *V. dahliae* strains 85S and JR2 (as positive inoculation control) at 21 days post inoculation (dpi). **(B)** Quantification of the canopy area of tomato plants at 21 dpi. Bars represent average of canopy area of five plants with standard deviation. Different letters indicate statistically significant differences (Student's *t*-test; $P < 0.05$). Inoculation experiments were performed with five plants for each fungal strain and independently repeated twice with similar results.

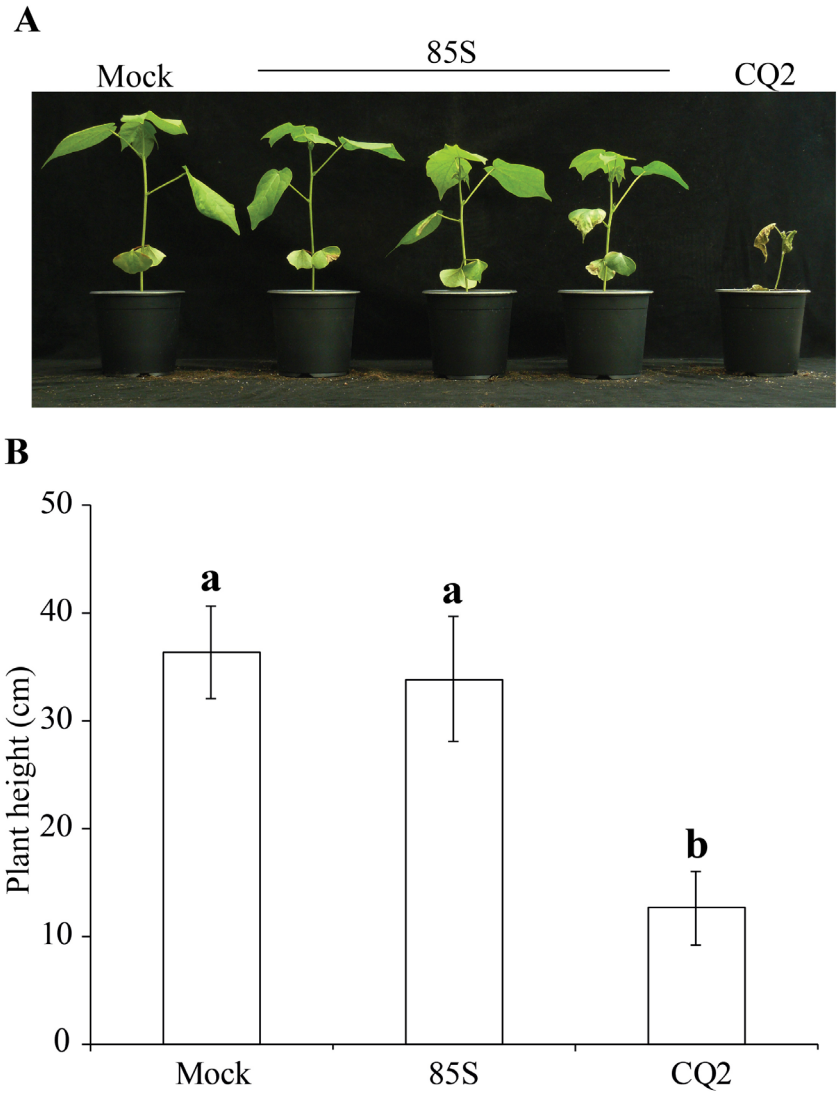


Figure S4. Phenotypes of cotton plants inoculated with *V. dahliae* strain 85S.
(A) Typical phenotype of cotton (cv. Simian3) plants upon mock-inoculation or inoculation with *V. dahliae* strains 85S and CQ2 (as positive inoculation control) at 28 days post inoculation (dpi). **(B)** Quantification of the plant height of cotton plants at 28 dpi. Bars represent average of height of five plants with standard deviation. Different letters indicate statistically significant differences (Student's *t*-test; $P < 0.05$). Inoculation experiments were performed with five plants for each fungal strain and independently repeated twice with similar results.

Table S1. *V. dahliae* strains used in this study.

Strain	Sequencing platform	Reference	Origin	Geographical location
HN	Illumina	Xu et al., 2012	Cotton	China
cd3	Illumina	Xu et al., 2012	Cotton	China
VdLs17	PacBio	Faino et al., 2015	Lettuce	USA
JR2	PacBio	Faino et al., 2015	Tomato	Canada
Vd57	Illumina	This study	Strawberry	Germany
V152	Illumina	Kombrink et al., 2017	Oak	Hungary
Vd52	Illumina	Kombrink et al., 2017	Pepper	Austria
van Dijk	Illumina	Kombrink et al., 2017	Chrysanthemum	The Netherlands
BP2	Illumina	Zhang et al., 2012	Cotton	China
ST16.01	Illumina	This study	Cotton	Syria
2009-605	Illumina	This study	Bell pepper	Ukraine
V4	Illumina	Keykhasaber, 2017	Olive	Spain
V200I	Illumina	This study	Strawberry	Germany
CBS38166	Illumina	de Jonge et al., 2012	Tomato	Canada
DVD-S26	Illumina	de Jonge et al., 2012	Soil	Canada
VdLS16	Illumina	de Jonge et al., 2012	Lettuce	USA
ST14.01	Illumina	de Jonge et al., 2012	Pistachio	USA
DVD-S29	Illumina	de Jonge et al., 2012	Soil	Canada
DVD-31	Illumina	de Jonge et al., 2012	Tomato	Canada
DVD-161	Illumina	de Jonge et al., 2012	Tomato	Canada
DVD-S94	Illumina	de Jonge et al., 2012	Soil	Canada
DVD-3	Illumina	de Jonge et al., 2012	Potato	Canada
V192I	Illumina	This study	Cotton	Spain
JKG8	Illumina	Kombrink et al., 2017	Potato	The Netherlands
85S	PacBio	This study	Sunflower	France
Vd39	Illumina	This study	Sunflower	Germany
V574	Illumina	Milgroom et al., 2014	Artichoke	Spain
v700	Illumina	Milgroom et al., 2014	Artichoke	Spain
v679	Illumina	Milgroom et al., 2014	Artichoke	Spain
T9	Illumina	Keykhasaber, 2017	Cotton	USA
V781I	Illumina	This study	Olive	Spain

Table S1. continued

Strain	Sequencing platform	Reference	Origin	Geographical location
V138I	Illumina	This study	Cotton	Spain
TM6	Illumina	Keykhasaber, 2017	Cotton	China
V117	Illumina	Keykhasaber, 2017	Olive	Spain
V991	Illumina	Zhang et al., 2012	Cotton	China
CQ2	PacBio	This study	Cotton	China
ST100	Illumina	de Jonge et al., 2012	Soil	Belgium
V76	Illumina	This study	Cotton	Mexico
463	Illumina	This study	Cotton	Mexico

Table S2. Primers used in this study.

Primer name	Oligonucleotide sequence (5'→3')[a]	Description
ITS-F	AAAGTTTTAATGGTTCGCTAAGA	Verticillium ribosomal internal transcribed spacer region (ITS), fungal biomass
ITS-R	CTTGGTCATTTAGAGGAAGTAA	Verticillium ribosomal internal transcribed spacer region (ITS), fungal biomass
SKO-Sun1-LF	GGTCTTAAUCAGATACCGATTATTGATCCTCGAC	For single *Sun1* deletion generation
SKO-Sun1-LR	GGCATTAAUCGTTAAGAGTTCATAGGCGAAGTTA	For single *Sun1* deletion generation
SKO-Sun1-RF	GGACTTAAUCTCGAAATTACAGAGCTTGCTATGA	For single *Sun1* deletion generation
SKO-Sun1-RB	GGGTTTAAUACTTGGCTATTTCTTCGTCTTTAGG	For single *Sun1* deletion generation
DKO-Sun1-LF	GGTCTTAAUTAGATTGTGTCCTGTGCAAGATATG	For double *Sun1* deletion generation
DKO-Sun1-LR	GGCATTAAUAGGGTTAACGTACATTATCAGCATG	For double *Sun1* deletion generation
DKO-Sun1-RF	GGACTTAAUGATCCCCTTGTCCATTATCTAGTGA	For double *Sun1* deletion generation
DKO-Sun1-RB	GGGTTTAAUCCGGGACAAAGGAAGGTTAATATAC	For double *Sun1* deletion generation

Table S2. *continued*

Primer name	Oligonucleotide sequence (5'→3')[a]	Description
outside primer-F	GCCTCACAACCAATCCACAG	Verification of *Sun1* deletion mutants
outside primer-R	ACATCGCCTCAGAGTCACAA	Verification of *Sun1* deletion mutants
Sun1-F (RT)	CTCATACTCTCCTCCGGTTCAT	*Sun1*, RT-PCR
Sun1-R (RT)	TTGTACCATCTCCACACGTTAAGTA	*Sun1*, RT-PCR
VdGAPDH-F	CGAGTCCACTGGTGTCTTCA	*V. dahliae GAPDH*, RT-PCR
VdGAPDH-F	CCCTCAACGATGGTGAACTT	*V. dahliae GAPDH*, RT-PCR
Ha-ELF-1α-F	ACCAAATCAATGAGCCCAAG	Sunflower *elongation Factor 1-α* (Ha-EF-1α), fungal biomass
Ha-ELF-1α-R	GAGACTCGTGGTGCATCTCA	Sunflower *elongation Factor 1-α* (Ha-EF-1α), fungal biomass
AtRubisco-F	GCAAGTGTTGGGTTCAAAGCTGGTG	Arabidopsis *Rubisco*, fungal biomass
AtRubisco-R	CCAGGTTGAGGAGTTACTCGGAATGCTG	Arabidopsis *Rubisco*, fungal biomass
NbRubisco-F	TCCGGGTATTAGGAAAAGCGT	*N. benthamiana Rubisco*, fungal biomass
NbRubisco-R	CCCAAGATCTGGGTCAGAGC	*N. benthamiana Rubisco*, fungal biomass

[a] USER cloning sites present in primer sequence are underlined, RT-PCR, real-time PCR.

Chapter 5

Signatures of adaptation to bread or durum wheat in a global collection of *Zymoseptoria tritici* isolates

Hesham A.Y. Gibriel[1,2], Guillaume Robert-Siegwald[3], Anne Genissel[3], Els C.P. Verstappen[4], L.C. Paul Keizer[4], Steven B. Goodwin[5,] Marc H. Lebrun[3], Gert H.J. Kema[1,4#] and Michael F. Seidl[1#*]

1 Laboratory of Phytopathology, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
2 Current address: School of Agriculture and Food Science, University College Dublin, Dublin, Ireland.
3 UMR BIOGER, Institut National de la Recherche Agronomique (INRA), Thiverval-Grignon, France.
4 Wageningen Plant Research, Wageningen University and Research, Wageningen, the Netherlands.
5 USDA–Agricultural Research Service, Purdue University, West Lafayette, Indiana, United States of America.

# These authors contributed equally

To be submitted for publication

ABSTRACT

Zymoseptoria tritici is a sexually reproducing fungal pathogen and is a major threat for wheat production in Europe and worldwide. The fungus exhibits host specificity, as it generally infects either bread wheat (BW; *Triticum aestivum*) or durum wheat (DW; *Triticum turgidum*), as well as cultivar specificity. By combining whole-genome sequencing of a worldwide collection of 136 *Z. tritici* isolates and phenotyping assays on a set of BW and DW cultivars, we observed extensive genetic diversity with moderate genetic differentiation among the geographical groups. Isolates originating from the Middle East, wheat's centre of origin, displayed increased genetic diversity when compared with populations from other geographical locations. We show large-scale structural genome variations and confirm the presence of 13 core chromosomes harbouring most of the identified effector genes, and structurally variable accessory chromosomes with higher SNP rates and presence/absence polymorphisms among effector genes. Profiling these effectors revealed a group of 183 conserved effectors and a group of 88 effectors that show presence/absence polymorphisms that were significantly closer to transposable elements than conserved ones. Assessment of genome-wide differences between BW and DW isolates revealed four effector genes affected by non-synonymous single nucleotide polymorphisms in BW isolates. Future functional analyses are needed to understand the role of the identified effector candidates on BW or DW cultivars.

INTRODUCTION

Microbial plant pathogens cause devastating diseases, thereby threatening global food security (Fisher et al. 2012). However, the majority of microbes is harmless to plants that have evolved an immune system that is able to detect pathogen invasions and mount appropriate defence responses (Cook et al. 2015). To overcome plant immunity, pathogens secrete effector proteins to enable host colonization, often by deregulation of plant immune responses (Cook et al. 2015; Rovenich et al. 2014). In turn, plants evolved immune receptors that can recognize such effectors leading to the re-establishment of immunity (Cook et al. 2015). This poses a selection pressure driving the emergence of favourable mutations in pathogen populations to evade such recognition (Cook et al. 2015; Möller and Stukenbrock 2017) that, in a continuous co-evolution, leads to extensive genomic diversity in plant and pathogen populations.

To facilitate co-evolution with their hosts, many plant pathogens evolved a compartmentalized genome with gene-dense and gene-sparse regions (Dong et al. 2015; Raffaele and Kamoun 2012). Gene-sparse regions often contain effector genes and display signs of accelerated evolution (Dong et al. 2015; Raffaele and Kamoun 2012; Seidl and Thomma 2017). For example, effector genes of the potato late blight pathogen *Phytophthora infestans* are located in gene-sparse, repeat-rich regions that are characterized by extensive structural variation and increased effector gene diversity (Haas et al. 2009; Raffaele et al. 2010). In the vascular wilt pathogen *Verticillium dahliae*, *in planta* expressed effector genes are localized in repeat-rich, and thus relatively gene-sparse, lineage-specific (LS) regions that evolved by chromosomal rearrangements (de Jonge et al. 2013; Faino et al. 2016). Gene-sparse genomic regions are either located on the core chromosomes, or can reside on separate chromosomes that are often referred to as dispensable or accessory chromosomes (Croll et al. 2013; de Jonge et al. 2013; Faino et al. 2016; Raffaele and Kamoun 2012; Wittenberg et al. 2009).

Wheat is currently the third main human food crop after maize and rice (FAOSTAT 2015), and the production is dominated by bread wheat (BW) (*Triticum aestivum*) and durum wheat (DW) (*Triticum turgidum*) (Troccoli et al. 2000). However, the production of wheat is continuously theatened by various plant pathogens such as cereal rusts (Hubbard et al. 2015; Singh et al. 2011), and *Zymoseptoria tritici* that causes septoria tritici leaf blotch (STB), which is a major threat for wheat production worldwide (Kettles and Kanyuka 2016; O'Driscoll et al. 2014). The fungus shows clear host specificity, as individual isolates originating from DW are virulent on the majority of DW cultivars, yet generally avirulent on BW cultivars, and *vice versa* (Mirzadi Gohari et al. 2015). Additionally, *Z. tritici* isolates display cultivar specificity (Kema et al. 1996; Mirzadi Gohari et al. 2015), which is mediated by the presence of specific resistance genes in different cultivars. Thus far, 21 resistance genes, so called septoria tritici blotch (*Stb*) genes, have been genetically mapped (Brown et al. 2015), of which *Stb6* (Chartrain et

al. 2005a), as well as the corresponding avirulence effector gene *AvrStb6* were recently cloned (Chapter 6) (Kema et al. 2018; Saintenac et al. 2018; Zhong et al. 2017).

Sexual reproduction is an intricate part of the life cycle of *Z. tritici* (Kema et al. 1996; Wittenberg et al. 2009), which results in highly polymorphic populations (Hartmann et al. 2018; Hartmann et al. 2017; McDonald et al. 2016; Naouari et al. 2016; Stukenbrock et al. 2011; Zhan et al. 2003). The reference *Z. tritici* isolate IPO323 harbours 21 chromosomes, comprising 13 core chromosomes that are present in all studied isolates, and eight dispensable or accessory chromosomes that can be absent without obvious fitness effects (Wittenberg et al. 2009). Particularly the accessory chromosomes are highly dynamic, showing abundant genomic rearrangements and extensive length variations, as well as presence/absence polymorphisms (Croll et al. 2013; Habig et al. 2017; Stukenbrock et al. 2010; Wittenberg et al. 2009). These chromosomes are generally gene-poor and can undergo fusions. Taken together, these processes likely contribute to rapid adaptive evolution of *Z. tritici* (Croll et al. 2013; Fouché et al. 2018b). Interestingly, thus far, no effector genes were identified on *Z. tritici* accessory chromosomes (McDonald et al. 2015; Mirzadi Gohari et al. 2015). The previously characterized *Z. tritici* effectors Mg3LysM and Mg1LysM, which belong to a conserved group of LysM effectors that are found in many fungal species (de Jonge and Thomma 2009; Kombrink et al. 2016; Kombrink and Thomma 2013), are located on core chromosomes (Marshall et al. 2011). Both Mg1LysM and Mg3LysM can bind chitin and protect fungal hyphae against plant hydrolytic enzymes, while Mg3LysM also blocks the activation of chitin-induced host immunity (Marshall et al. 2011). Similarly, the other known effector genes *AvrStb6* (Kema et al. 2018; Zhong et al. 2017), *Zt_8_609* (Hartmann et al. 2017), *Avr3D1* (Meile et al. 2018b) and *Zt80707* (Poppe et al. 2015) are located on core chromosomes, usually in orphan genomic regions associated with transposable elements (TEs) or recombination hotspots (Croll et al. 2015; Meile et al. 2018b; Plissonneau et al. 2016). For example, genome-wide analyses revealed that *AvrStb6* resides in a recombination hotspot, which exhibits signatures of accelerated evolution including elevated SNP levels (Brunner and McDonald 2018; Kema et al. 2018; Zhong et al. 2017). Taken together, the natural diversification of *Z. tritici* effectors suggests a continuous co-evolutionary adaptation of pathogen populations to their wheat hosts.

Population genomics is a powerful method that can improve our understanding of diversity in natural pathogen populations and estimate their evolutionary potential (Grünwald et al. 2016). Initial studies showed genetic diversification in local and global populations, with high rates of gene flow and limited population structure (Linde et al. 2002; Zhan et al. 2003). These studies were based on a limited number of molecular markers, which likely resulted in a biased assessment of the overall genomic diversity of this species and could not reveal loci associated with *Z. tritici* cultivar specificity (McDonald et al. 2015). Recent population genomics studies further explored the genetic diversity of *Z. tritici* populations (Grandaubert et al. 2018; Hartmann et al.

2018; Hartmann et al. 2017; McDonald et al. 2016). Although these studies often assessed *Z. tritici* isolates that were not widely geographically separated and were often sampled from single wheat fields (Hartmann et al. 2018; Hartmann et al. 2017), they corroborated previous findings that *Z. tritici* is a highly dynamic pathogen that genetically differs between and within populations, even at small spatial scales (Croll et al. 2013; Hartmann et al. 2017; Linde et al. 2002; McDonald et al. 2016; Stukenbrock et al. 2011). Here, we aimed to explore the genomic diversity in 136 *Z. tritici* isolates that were sampled from bread and durum wheat accessions in 31 locations in the major global wheat producing regions. By comparing the chromosome and effector diversity of BW and DW isolates, along with the phenotyping of 118 isolates on four wheat cultivars, we provide a foundation for further genomic and functional studies of host specificity in *Z. tritici*.

MATERIALS AND METHODS

Fungal isolates and phenotyping assays

In total, 136 *Z. tritici* isolates were sampled from single wheat fields at 31 locations (Table S1). Upon sampling, collected isolates were prepared for long-term storage at -80°C (Kema et al. 1996).

Phenotyping assays were conducted on the bread wheat cvs. Taichung 29 (susceptible BW control; Kema et al. 1996), Bulgaria 88 (carries *Stb1 and Stb6*; (Adhikari et al. 2004; Chartrain et al. 2005a), TE9111 (carries *Stb11, Stb7*, and *Stb11*; (Chartrain et al. 2005b) and the DW cv. Volcani 447 (susceptible control; Kema et al. 1996; Mirzadi Gohari et al. 2015). Wheat seeds were planted in potting soil (Swedish sphagnum peat 5%, grinding clay granules 41%, garden peat 5%, beam structure 4%, steamed 140 compost 33%, PG-Mix-15-10-20- 12%) from the Unifarm greenhouse facility of Wageningen University and Research, Wageningen, The Netherlands, and after germination plants were grown until inoculation for 10 days in the greenhouse at 18°C, under 16 h light and 8 h dark cycles and a relative humidity (RH) of 75% (Kema et al. 2018). The isolates were taken from the -80°C stocks, cultured for 5-10 days on potato dextrose agar and from there transferred to yeast-glucose medium (yeast extract 10 g/L, glucose 30 g/L) and incubated in an orbital shaker (New Brunschwick (now Eppendorf) Innova 44r, orbit diameter 2.5 cm, The Netherlands) for 4-6 days at 16°C. Yeast-like conidiospores were collected by overnight sedimentation and inoculum was adjusted at 10^7 conidiospores/mL supplemented with two drops of Tween 20 (Merck, The Netherlands) surfactant and atomized over the primary leaves of the wheat seedlings. After inoculation, plants were kept in transparent plastic bags at 22°C at 100% relative humidity (RH) for 48 h, and were subsequently maintained in the same greenhouse at RH 95%. Disease development was monitored daily and after 10 days secondary and subsequent leaves were removed to ensure sufficient light on the inoculated leaves. After 21 days the inoculated leaves were collected and the percentage of leaf necrosis and the percentage of leaf area covered

with pycnidia were assessed on a 0-100 scale. We used a split plot statistical design and all assays were performed in two biological replicates with the *Z. tritici* isolates IPO-87016.MS1, IPO-90006.MS2, and a mock present in each tray as controls. The allocation of isolates over the trays was randomized, but both biological replicates were always in different blocks.

DNA extraction and whole-genome sequencing

Genomic DNA of all *Z. tritici* isolates was extracted from five-day-old liquid cultures grown in yeast extract-peptone-dextrose broth (YPD) using a standard phenol–chloroform procedure (Sambrook et al. 1989). Library preparation (~500 bp insert size) and whole-genome sequencing (150 bp paired-end reads) using Illumina HiSeq-2500 sequencing were performed at the Joint Genome Institute (JGI, Walnut Creek, CA, USA) and the generated raw sequencing data are deposited at the JGI under project CSP983.

Single nucleotide polymorphisms (SNPs) calling

We generated a repeat-masked genome of *Z. tritici* IPO323 (Goodwin et al. 2011) and used it as a reference for all subsequent analyses. To this end, *de novo* repeats were annotated using REPET (v2.2; default setting) (Flutre et al. 2010), which was further utilized to subsequently mask the genome sequence using RepeatMasker (v4.0.6) (Smit et al. 2016).

Paired-end reads of each isolate were mapped to the masked reference genome *Z. tritici* IPO323 (Goodwin et al. 2011) using BWA (MEM) (default options) (v0.7) (Li and Durbin 2010). PCR duplicates were marked using Picard tools (v1.1) (http://broadinstitute.github.io/picard). Single nucleotide polymorphisms (SNPs) were identified using the HaplotypeCaller of the Genome Analysis Toolkit (GATK) (v3.8) (McKenna et al. 2010). First, variations were detected for each isolate individually with the following settings: emitRefConfidence GVCF and ploidy 1. Joint genotyping for all isolates was performed using GenotypeGVCFs with the option maxAltAlleles 2. All non-SNPs variants were removed using SelectVariants (-selectType SNP) and to obtain high quality SNPs, they were further filtered using VariantFiltration with the following cut-offs: QUAL > 250, MQ > 30, QD > 20, FS < 0.1, and BaseQRankSum, ReadPosRankSum, MQRankSumPos between -2 and 2. Next, we excluded SNPs with missing genotype calls in 10% of the isolates and a minor allele frequency (MAF) <5% using vcftools (v0.1.5) (Danecek et al. 2011), similar to the procedure of Hartmann et al. (2017). Moreover, we excluded SNPs on the accessory chromosomes, as these are not conserved across the global panel (Goodwin et al. 2011). Principal component analysis (PCA), using the filtered SNP set, was performed with SNPRELATE (v1.6) (Zheng et al. 2012). SNPs were annotated with SnpEff (v3.2) (Cingolani et al. 2012), using the manually-refined gene annotation of *Z. tritici* isolate IPO323 (Kema et al. 2018). A phylogenetic tree was constructed using RealPhy (readLength 150) (Bertels et al. 2014) that used Bowtie2 (Langmead and Salzberg 2012) to map reads of individual *Z. tritici* isolates to the repeat-masked genome of *Z. tritici* IPO323 (Goodwin et al. 2011). The population structure was analysed using

Structure (v.2.3.2) (Falush et al. 2003; Pritchard et al. 2000) and the data were analysed with K ranging from one to eight with ten repetitions for each tested K value. We used 50,000 samples as a burn-in period and 100,000 samples per run for the Monte Carlo Markov Chain (MCMC) replicates.

Chromosomes and effector polymorphisms

Whole-genome sequencing data were used to determine chromosome polymorphisms across the entire panel. To this end, the read coverage of each isolate relative to the masked reference genome IPO323 was determined using BEDtools coverage (v2.25) (default parameters) (Quinlan and Hall, 2010). Subsequently, the individual read coverages for each isolate were normalized by averaging the read coverage for each chromosome over the whole-genome coverage and boxplots were produced using the R package ggplot2 (Wickham 2015). Next, we determined chromosome polymorphisms following the procedure of Fouché et al. (2018) by classifying a chromosome as (i) absent, if the normalized coverage ratio is close to zero (<0.3), (ii) present, if the normalized coverage ratio is close to one (≥0.7 and <1.3), (iii) duplicated, if the normalized coverage ratio is close to two (≥1.7), (iv) partially deleted, if the normalized coverage ratio is ≥0.3 and <0.7, or as (v) partially duplicated, if the normalized coverage ratio is ≥1.3 and <1.7. Heatmaps indicating the chromosome polymorphisms were generated using the R package pheatmap (Kolde and Kolde 2015). In addition, we carried out Pulsed field gel electrophoresis (PFGE), as described by (Mehrabi et al. 2007), for *Z. tritici* isolates IPO323 and ND95 to check whether the overall chromosome polymorphisms match with the whole-genome sequencing data.

To identify *Z. tritici* candidate effectors, N-terminal signal peptides were first predicted for all the manually-refined protein annotations *of Z. tritici* isolate IPO323 (Kema et al. 2018) using SignalP (v4.1) (Petersen et al. 2011). Subsequently, effector candidates were predicted from the set of secreted proteins using the machine-learning approach applied in EffectorP (v1.0) (default parameters) (Sperschneider et al. 2016). The effector polymorphisms across the entire global panel in relation to *Z. tritici* IPO323 were determined using BEDtools coverage (Quinlan and Hall 2010). To assess effector gene expression levels, we used an RNA-seq dataset (single-end) of *Z. tritici* isolate IPO323 collected over the entire pathogenesis on wheat cv. Riband (Rudd et al. 2015). RNA-seq data were mapped to the reference *Z. tritici* genome IPO323 using Tophat (v2.0) (min-intron-length 20, max-intron-length 2000, max-multihits 5) (Dobin et al. 2013). Subsequently, gene expression levels for each sample were estimated using Cuffdiff (v2.2) with default parameters (Trapnell et al. 2010) and reported as fragments per kilobase of exon per million fragments mapped (FPKM).

RESULTS

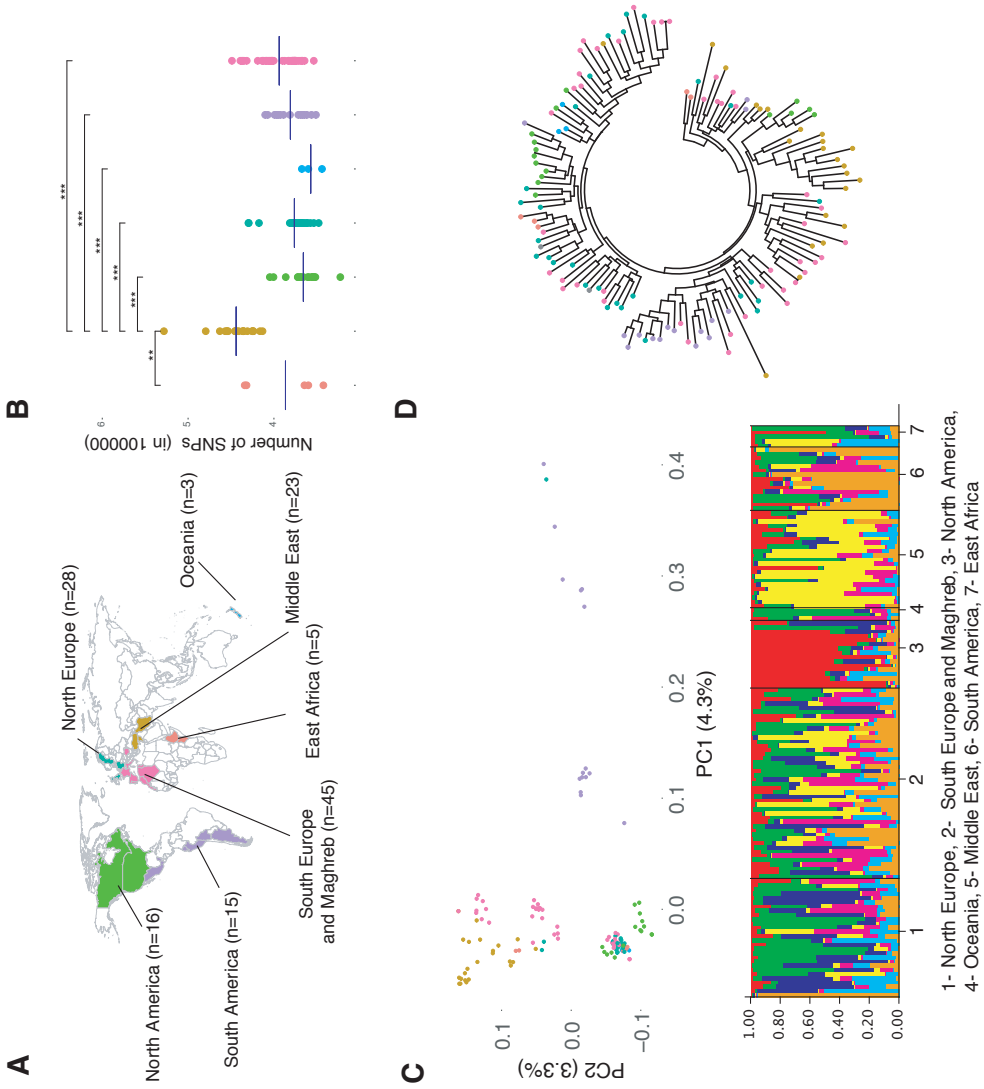
The global panel of *Zymoseptoria tritici* is highly polymorphic

Our collection comprised 136 *Z. tritici* isolates sampled from BW (95) or DW (14) and 27 isolates were sampled from unknown wheat accessions (Table S1). Most countries were represented by five (variable between 3-8) *Z. tritici* isolates, whereas some were represented by a single isolate, such as Bolivia, Peru, Italy and Romania, or more isolates (15-21) for France, Iran and the USA (Table S1). The isolates were grouped into seven distinct populations based on their geographical origin (North America, South America, South Europe and the North African countries Morocco, Tunisia and Algeria - collectively known as the Maghreb- East Africa, Middle East, North Europe, and Oceania) (Figure 1A; Table S1).

To estimate the genetic diversity across the collective panel, each isolate was genotyped and in total 877,279 bi-allelic SNPs were identified, with on average ~379,000 SNPs per *Z. tritici* isolate, indicating high levels of genetic diversity. The Oceania population displayed the least number of SNPs (average of 9.4 SNPs per kb) compared with the other populations (Figure 1B). The Middle East population showed the highest number of SNPs (average of 23.8 SNPs per kb), highlighting that genetic diversity in *Z. tritici* populations significantly varies between different geographical regions and that the maximum genetic diversity is observed in the Middle East population with an average number of 440,977 SNPs per isolate (ranging from 413,284 to 528,164 SNPs) (Figure 1B).

To analyse the population structure of the global panel, we performed a principal component analysis (PCA) using all identified SNPs, and reconstructed a maximum-likelihood phylogeny based on 291,000 biallelic SNPs as determined by RealPhy (Bertels et al. 2014). We observed that *Z. tritici* isolates are highly polymorphic and do not consistently group by geographical origin (Figure 1 C and D). We therefore performed an additional population genetic analysis and observed a geographical pattern for the populations from North Europe, North America, and the Middle-East (Figure 1C), suggesting a moderate genetic differentiation at a global scale.

Figure 1. Highly diverse global *Zymoseptoria tritici* isolates. (A) The number of isolates sampled in seven color-coded major wheat producing regions of the world. **(B)** The number of identified single nucleotide polymorphisms in all *Z. tritici* isolates in each population. The blue horizontal lines represent the average number of SNPs for each population. **(C)** Principal component analysis (PCA) based on all identified 877,279 SNPs in all *Z. tritici* populations (upper panel). The percentage of variance explained by each component is shown in parentheses. Lower panel shows the clustering assignments of each *Z. tritici* genotype inferred using the software Structure. The distribution of individual assignations estimated for K=7 clusters, from Bayesian inference cluster analysis performed with 9,874 *Z. tritici* SNPs. Each vertical line represents an individual isolate. **(D)** Phylogenetic tree of all *Z. tritici* isolates based on 291,000 SNPs. Colored nodes correspond to the geographic origins of the *Z. tritici* isolates.

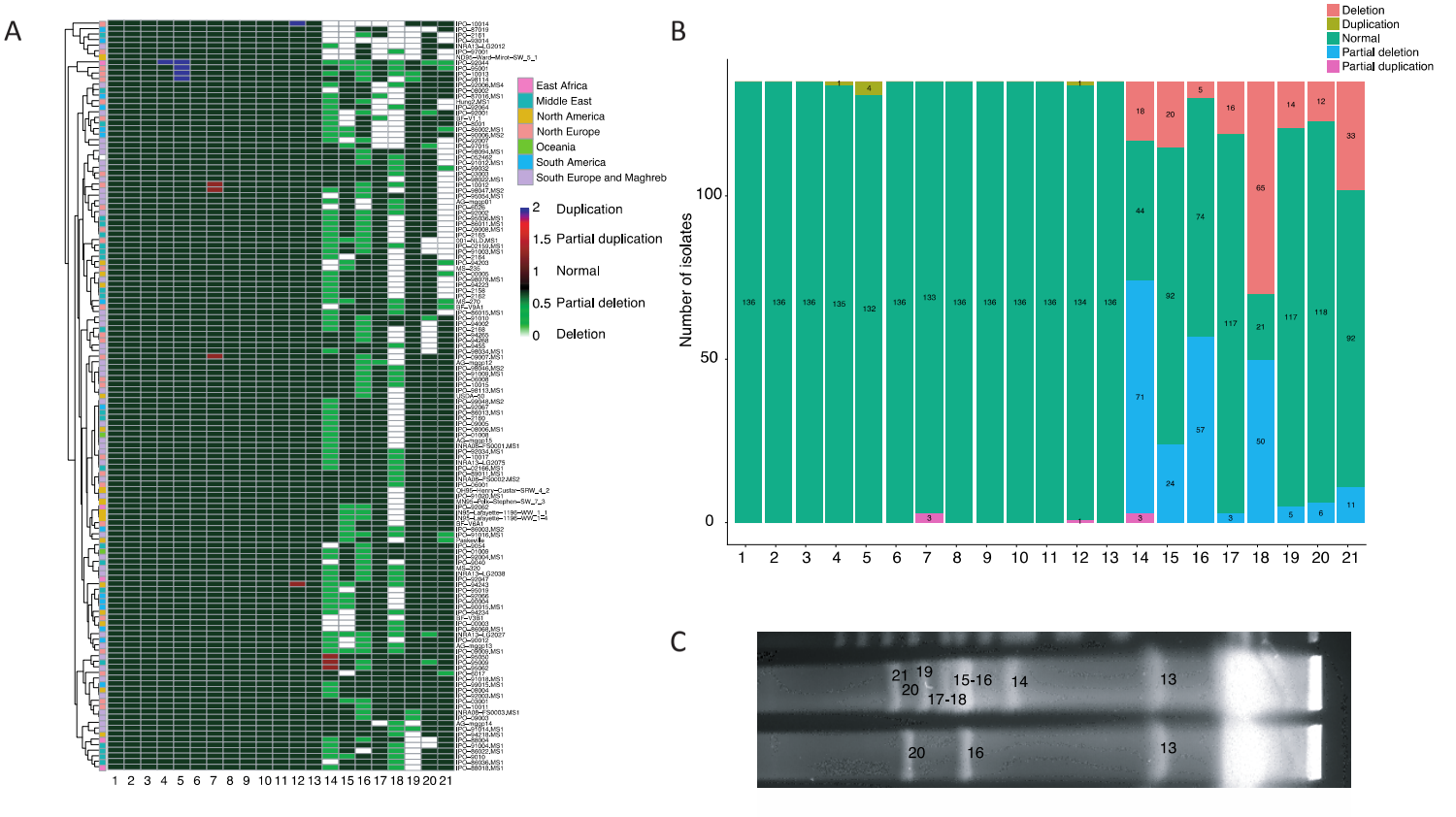


The highly polymorphic accessory chromosomes of *Zymoseptoria tritici*

The *Z. tritici* reference isolate IPO323 contains 21 chromosomes comprising 13 core chromosomes and eight accessory chromosomes (Goodwin et al. 2011; Habig et al. 2017; Wittenberg et al. 2009). We assessed chromosome size and number polymorphisms across the entire panel and observed that all core chromosomes are always present and only display size polymorphisms (Figure 2A and B). However, we observed a few cases of partial chromosomal duplications in isolates IPO-10014 (duplication of chromosome 12), IPO-92044 (duplication of chromosome 4), and IPO-98114, IPO-10013, IPO-95001, and IPO-92044 (duplication of chromosome 5) (Figure 2A). In addition, we observed partial duplication of chromosome 7 in isolates IPO-09007, IPO-98047, and IPO-10012 and also of chromosome 12 in isolate IPO-94243 (Figure 2A). The accessory chromosomes were extremely variable as multiple chromosomes were either partially or completely absent (Figure 2A and B). Chromosomes 18 and 21 showed the highest absence frequency, 48% and 24%, respectively, while partial deletions mainly affected chromosomes 14, 16 and 18 (52%, 41% and 36%, respectively). On the other hand, chromosomes 17, 19 and 20 showed the least polymorphisms and were present in 86% of the isolates. We did not encounter an isolate without any accessory chromosome, but isolate IPO-93014 only carried chromosome 20 and isolate ND95 carried only chromosomes 16 and 20, which we validated by pulsed field gel electrophoresis (Figure 2A and C). Taken together, core chromosomes of *Z. tritici* isolates are always present but individual chromosomes may be duplicated, while the number of accessory chromosomes is highly variable with frequent total and partial deletions, although they are never completely absent (Figure 2).

The SNP density of the reference isolate *Z. tritici* IPO323 is higher in accessory chromosomes than in core chromosomes (Goodwin et al. 2011; Stukenbrock et al. 2011). Here, we assessed the SNP density in the entire panel by averaging the SNP counts over 1-kb stretches of core and accessory chromosomes over the genome-wide SNP counts. The SNP densities in accessory chromosomes (average 1.29 SNPs per kb) are higher than in the core chromosomes (average 1.01 SNPs per kb). This was particularly evident for chromosome 16, which showed a higher SNP density than the other chromosomes, while the lowest SNP density was observed in chromosome 14 (Figure S2). Thus, *Z. tritici* accessory chromosomes are much more polymorphic than core chromosomes with frequent large-scale structural variations and higher SNP rates.

Figure 2. Conserved core chromosomes and highly polymorphic accessory chromosomes in a global *Zymoseptoria tritici* panel. (A) Heatmap indicating the presence or absence of chromosomes in 136 *Z. tritici* isolates based on whole-genome sequencing data. Colors indicate whether a chromosome is present (black), partially duplicated (red), duplicated (red), partially deleted (green), or deleted (white). **(B)** Summary of chromosomal polymorphisms for all *Z. tritici* isolates. The frequencies of presence, partial duplication, duplication, partial deletion, and deletion are shown for each of the 21 chromosomes depicted as vertical bars. **(C)** Pulsed-field gel electrophoresis (PFGE) of ND95 and reference *Z. tritici* isolate IPO323. Chromosomal bands corresponds to indicated accessory chromosomes.


A
East Africa
Middle East
North America
North Europe
Oceania
South America
South Europe and Maghreb
2 Duplication
1.5 Partial duplication
1 Normal
0.5 Partial deletion
0 Deletion
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
B
Deletion
Duplication
Normal
Partial deletion
Partial duplication
Number of isolates
100
50
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
C
21 19
20
15-16
14
13
17-18
20
16
13

Diversified effector catalogues among *Z. tritici* isolates

We predicted the effector repertoire of the reference isolate *Z. tritici* IPO323, and used it to assess effector diversity in the entire panel. We identified 1,034 genes that encode secreted proteins, of which 299 were predicted as effector candidates. The identified effector candidates, which also include the previously characterized effector AvrStb6 (Kema et al. 2018; Zhong et al. 2017), lacked conserved or functional domains (Table S3).

To identify presence/absence polymorphisms of effector genes in the entire panel, we mapped the genomic reads of all isolates to the reference strain IPO323 gene set and identified on average 12,250 genes per isolate, representing 93% of the 13,157 genes determined in IPO323. Among these genes, approximately 270 effector genes were predicted for each isolate, representing 90% of the 299 effectors predicted by SignalP. Comparison of these data revealed that 183 (60%) effectors are conserved in all *Z. tritici* isolates. Among the remaining presence/absence polymorphisms, we identified 88 effectors that were partially absent and 28 effectors that were only found in the reference isolate *Z. tritici* IPO323 (Figure S3). The majority (285, 95%) of these effectors was located on the core chromosomes and they were either conserved or showed presence/absence polymorphisms, whereas only 14 effectors were present on the accessory chromosomes and merely showed presence/absence polymorphisms (Figure S3). Of the known core effectors, we found that the previously characterized effector AvrStb6 was present in all isolates, except in isolate IPO-03003 (Figure 3).

To further investigate the differences between effector genes and other genes across the panel, we assessed gene length, distance to TEs, inter-genic length, and *in planta* expression levels. We observed that genes with presence/absence polymorphisms were significantly shorter than conserved genes (Wilcoxon rank-sum test, $P < 0.05$) (Figure 4A). Effector genes are generally shorter than other genes, but we did not find differences between conserved effectors and those showing presence/absence polymorphisms (Figure 4A). We also determined that genes and effector genes showing presence/absence polymorphisms were significantly closer to TEs than conserved ones (Wilcoxon rank-sum test, $P < 0.05$) (Figure 3B), but we did not observe an over-representation of effectors in gene-poor, TE-rich regions of *Z. tritici* genome (Figure 4A). Moreover, compared to conserved genes, the *in planta* expression of genes and effector genes with presence/absence polymorphisms was significantly lower (Wilcoxon rank-sum test, $P < 0.05$) (Figure 4A). Finally, we assessed the relative diversity of candidate effector genes in the global panel and determined that other genes displayed more synonymous SNPs than effector genes and that these carry more non-synonymous SNPs that lead to amino acid changes than other genes (Figure 4B). For instance, the previously characterized effector gene *AvrStb6* (Kema et al. 2018; Zhong et al. 2017) showed seven non-synonymous SNPs across the panel, leading to missense mutations and promoting virulence of *Z. tritici* isolates to the bread wheat cultivar Shafir (Kema et al. 2018).

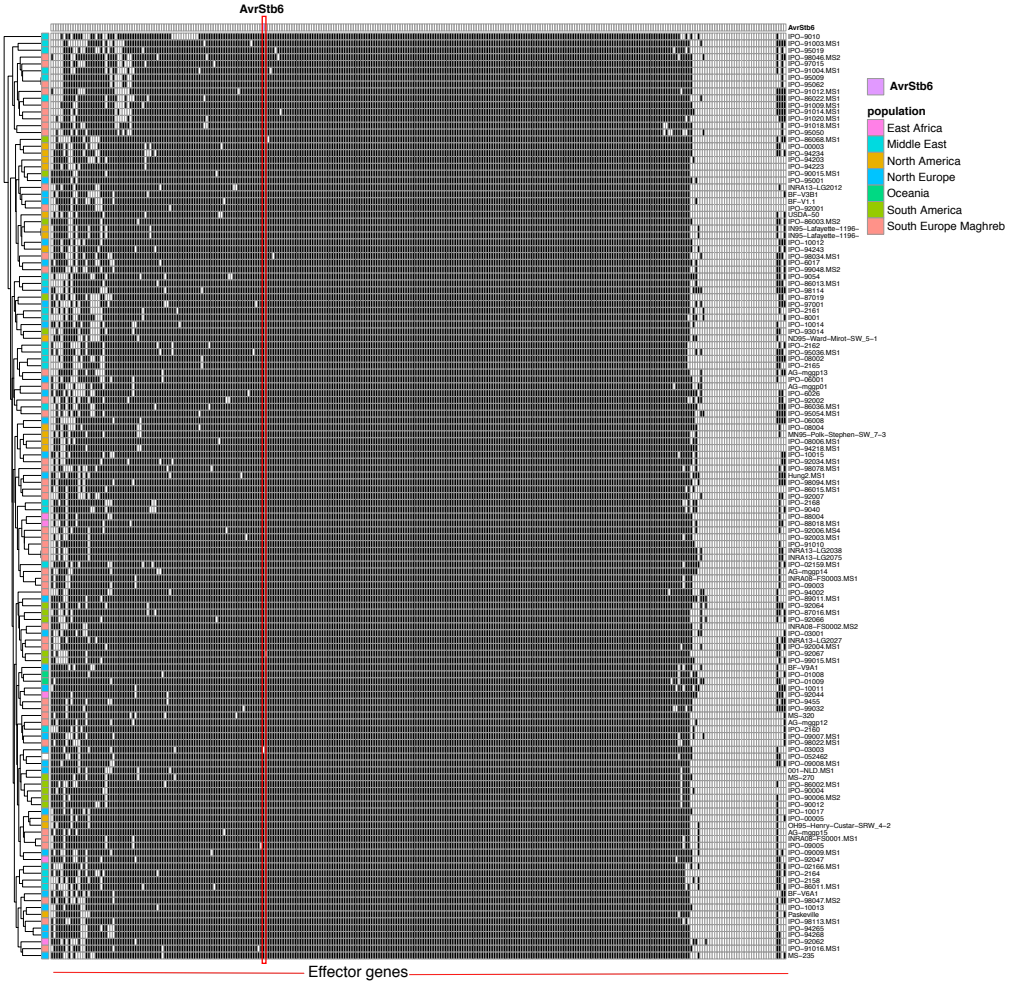


Figure 3. Heatmap indicating the presence or absence of effector genes in 136 *Z. tritici* isolates based on the gene annotation of reference *Z. tritici* isolate IPO323. The identified *Z. tritici* effector AvrStb6 is indicated and the red box indicates that the effector is absent in one isolates originating from Germany.

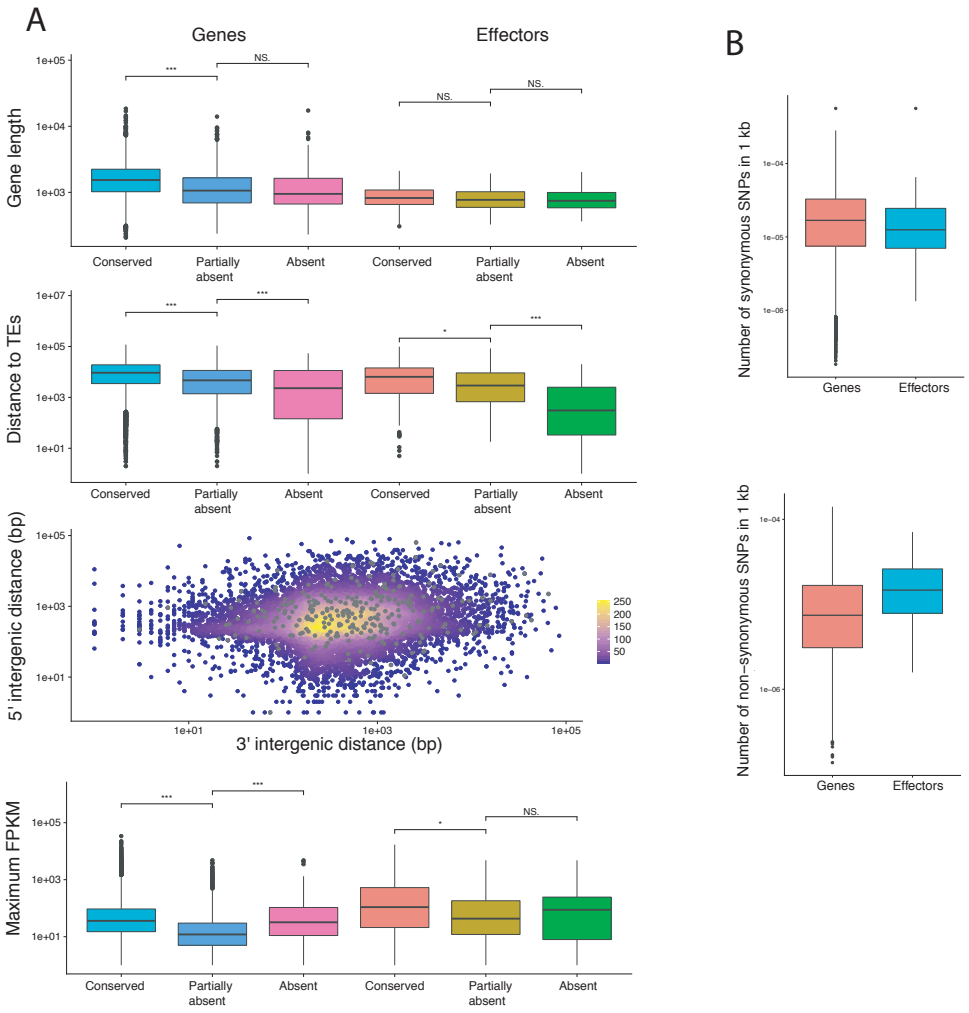


Figure 4. Effector diversification across a global *Zymoseptoria tritici* panel. (A) Features of conserved and polymorphic genes and effector genes in *Z. tritici,* including (first panel) gene length (kb), (second panel) distance to closest TEs, (third panel) inter-genic length of genes (blue) and effector genes (gray), (fourth panel) and *in planta* expression. **(B)** Boxplots showing the number of synonymous and non-synonymous SNPs for genes and effectors in one kb.

Divergence of *Z. tritici* pathogenicity to wheat species and cultivars

In total, 118 isolates from the global panel (Table S1) were phenotyped on a defined set of host genotypes and we observed clear host specialization. Among the 83 *Z. tritici* isolates originating from BW, 61 isolates were only virulent on at least one of the three BW cultivars, while four out of the ten *Z. tritici* isolates originating from DW were virulent on

the DW cv. Volcani 447 (Table S2). For example, DW isolate IPO95019 is non-pathogenic on all the BW cultivars (5% pycnidia), while it is highly virulent on DW cv. Volcani 447 (40% pycnidia) (Figure 5A). In contrast, BW isolate IPO02158 was non-pathogenic on DW cv. Volcani 447 (5% pycnidia), but highly virulent on all three BW cultivars (30-40% pycnidia) (Figure 5A).

We observed that Iranian isolates were the most aggressive on BW cvs. Taichung 29 (77% pycnidia), Bulgaria 88 (66% pycnidia), and TE9111 (57% pycnidia) (Table S2). In general, the most agressive isolates on BW cvs. originated from the Middle East population with an average pycnidia percentage of 30% (Figure 5B), while the most agressive isolates on cv. Volcani 447 originated from South Europe and the Maghreb (Figure 5B). Thus, our study shows that the most virulent isolates on BW or DW cvs. originate from geographic regions associated with their origin or primary production area. Moreover, we identified 12 isolates with virulence to both BW and DW cultivars (Table S2).

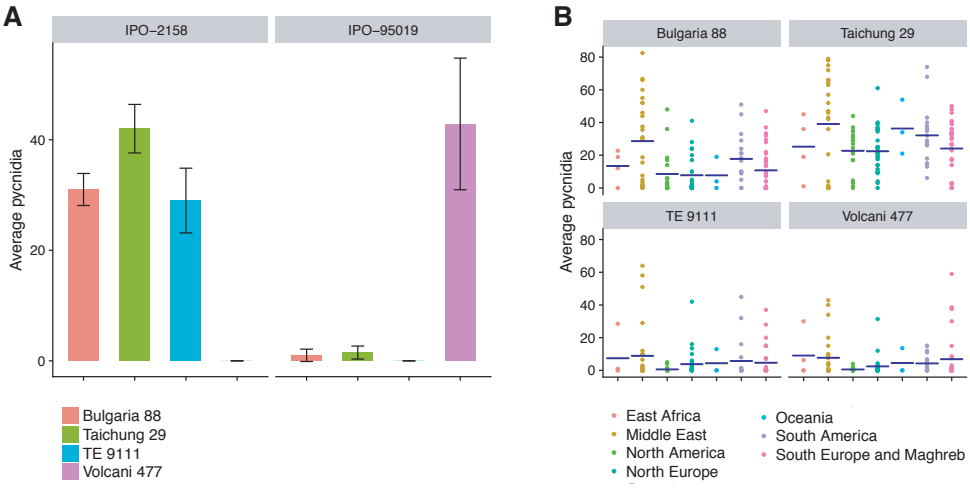


Figure 5. Divergence of wheat host and cultivar specificity in a global *Zymoseptoria tritici* panel. (A) Average pathogenicity of two exemplary *Z. tritici* isolates adapted to bread wheat (IPO-2158) or durum wheat (IPO-95019) on the bread wheat cvs. Bulgaria 88, Taichung 29, and TE9111, and the durum wheat cv. Volcani 477. Error bars represent standard error of two biological replicates. **(B)** Average pathogenicity of all *Z. tritici* isolates on the four phenotyped wheat cultivars. Blue horizontal lines represent average pycnidia levels for each population.

Genome-wide differences between bread and durum wheat adapted *Z. tritici* isolates
To further explore the observed differences and adaptation to BW or DW, we assessed the genetic make-up of isolates with specificity to either BW or DW with respect to the presence of core and accessory chromosomes and effectors. As expected, core chromosomes are maintained amongst all isolates, but accessory chromosomes are

highly variable (Figure S4A and B). We also determined that many candidate effectors are conserved between bread and durum isolates, including *AvrStb6* (Figure S4C). Hence, it seems that pathogenicity differences between BW and DW isolates are associated with other polymorphisms in effector genes such as SNPs. Therefore, we assessed SNPs in the protein coding regions of all genes of the BW and DW isolates and found that isolates with BW specificity harbour 349 genes with SNPs leading to an amino acid change or a stop codon, but these included only four effector genes (Figure 6A and B).

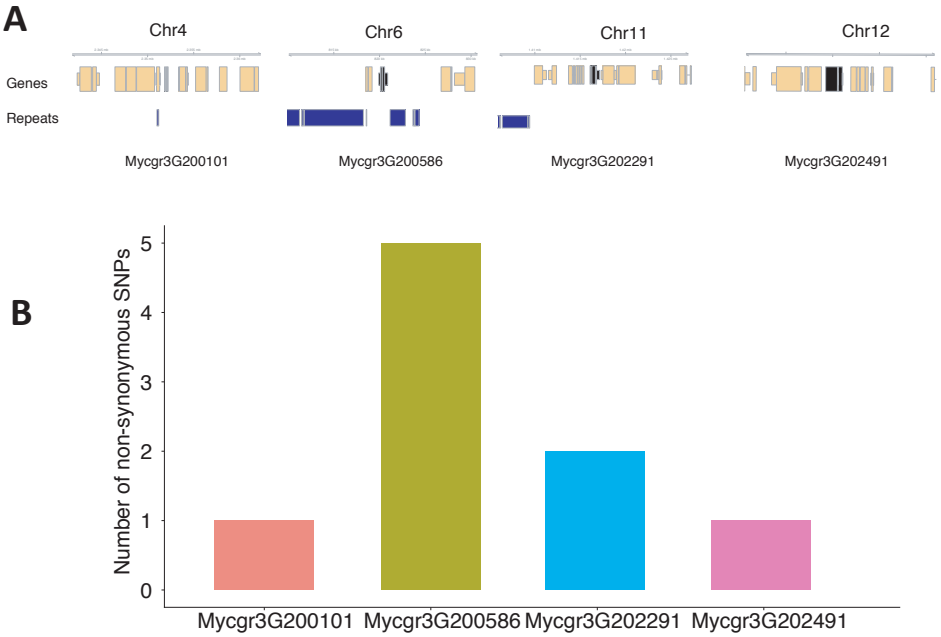


Figure 6. Genome-wide differences reveal candidate effector genes affected by SNPs in *Zymoseptoria tritici* isolates adapted to bread or durum wheat. (A) Genomic environment of identified effector candidates. Genes (yellow), and TEs (blue) are shown on chromosomes 4, 6, 11 and 12 in 10kb regions upstream and downstream of each effector gene (black). **(B)** The identified number of non-synonymous SNPs in each effector gene.

DISCUSSION

Global food security is one of the most important issues for humanity, which is threatened by plant pathogens that largely impact crop production (Fisher et al. 2018). To develop durable disease management strategies, knowledge on the evolutionary potential of pathogens in agricultural and natural environments is required (Plissonneau et al. 2017). Using population genomics, pathogen dynamics and genetic diversity can be assessed, thereby providing insights into pathogen adaptation to their hosts (Gibriel et al. 2016; Plissonneau et al. 2017; Stukenbrock and Bataillon 2012). For instance, population genomics of the ash dieback pathogen *Hymenoscyphus fraxineus* revealed the recent migration of this pathogen form East to West Europe and its extended genetic diversity in native Asian populations (McMullan et al. 2018). Similarly, population genomic analyses of *Z. tritici* revealed a rapidly evolving pathogen adapted to infect wheat, which emphasized the challenges imposed by this pathogen on future breeding programs (McDonald et al. 2016; Plissonneau et al. 2017; Plissonneau et al. 2018; Stukenbrock and Bataillon 2012; Stukenbrock et al. 2010). Recently, we provided new insights into the epidemiology of *Z. tritici* by showing its ability for sexual reproduction under adverse conditions that result in population sweeps for fungicide resistance and the slow down of the invasion of virulence alleles in natural populations (Chapter 6). Here, we performed comparative population genomics, using a global collection of *Z. tritici* isolates that were sampled from BW and DW accessions in the major global wheat producing regions. We demonstrated their high SNP rate, extensive structural variation on accessory chromosomes, and genome-wide differences between BW and DW isolates, which presumably contribute to the dynamic nature of *Z. tritici* populations (Croll et al. 2013; Hartmann et al. 2017; Linde et al. 2002; McDonald et al. 2016; Stukenbrock et al. 2011).

By phenotyping 118 isolates from the global panel on a defined set of BW and DW cultivars, we observed clear host specialization, which is consistent with previous studies showing that *Z. tritici* isolates are either virulent on BW or DW cultivars (Mirzadi Gohari et al. 2015). Thus, we aimed to assess the genomes of BW and DW isolates in order to identify causal presence/absence polymorphisms. We observed large scale structural variation in accessory chromosomes and conserved core chromosomes in BW and DW *Z. tritici* isolates, as expected for this fungal pathogen (Croll et al. 2013; Stuckenbrock et al. 2011). Next, we aimed to assess the differences in effector repertoires of BW and DW isolates, as modifications in effectors enable pathogens to adapt to their hosts (Seidl et al. 2014). One mechanism to diversify effectors is through the occurrence of SNPs. Once in protein coding regions of effector genes these might be shared among virulent isolates but should be absent in avirulent ones. Thus, we reasoned that conserved effectors of *Z. tritici*, albeit being present in virulent and avirulent isolates, facilitate adaptation to infect particular wheat cultivars. By assessing genome-wide differences between BW and DW isolates, we identified four effector genes with non-synonymous SNPs only leading to a different amino acid or a stop codon in bread wheat isolates.

The previously identified conserved effectors *AvrStb6* and *Avr3D1* carry SNPs shared by virulent isolates allowing them to circumvent wheat immunity (Kema et al. 2018; Meile et al. 2018a; Zhong et al. 2017). Future functional analyses are needed to investigate the role of the identified effectors and SNPs in *Z. tritici* host specificity. Additionally, since the *Z. tritici* reference strain IPO323 is a bread wheat isolate, mapping genomic reads of DW isolates to this reference genome will not disclose DW lineage-specific regions. Therefore, a reference DW isolate genome is required to further assess the difference between BW and DW isolates, as gene absence also has considerable impact on pathogen virulence (Fouché et al. 2018a). For example, aggressiveness of *Z. tritici* isolate 3D7 towards wheat cv. Tornonit is associated with deletion of the effector gene *Zt_6_809* (Hartmann et al. 2017). Similarly, in *V. dahliae*, deletion of the *Ave1* effector gene is associated with virulence on tomato plants that carry the Ve1 immune receptor (de Jonge et al. 2012).

A truly comprehensive global overview of genetic diversity of *Z. tritici* is unavailable, as many wheat-growing regions throughout are significantly under-represented due to lack of sampling. For example, little is known about the *Z. tritici* genetic diversity in Asia, Africa, as well as South America, despite their importance for global wheat production (Cordo et al. 2017). We demonstrated the genome-wide diversity by comparing chromosome and effector diversity of 136 *Z. tritici* isolates that were sampled from many and diverse wheat-growing regions. First, we identified a high number of SNPs in these *Z. tritici* populations, indicating a high level of genetic diversity, comparable with previous population studies (Hartmann et al. 2017; McDonald et al. 2016). Of all populations, the Middle East population displayed the highest number of SNPs (average number of 440,977 SNPs per isolate), highlighting the huge genetic diversity among *Z. tritici* isolates originating from the centre of origin (Stukenbrock et al. 2006; Stukenbrock et al. 2011), which likely explains why these isolates are pathogenic across the phenotyped BW and DW cultivars. Subsequent clustering of the isolates based on the identified SNPs showed highly polymorphic populations that did not align with their geographical origin or hosts. This is in accordance with previous studies that showed no or little population structure of global *Z. tritici* isolates at local, regional, and worldwide scales (Linde et al. 2002; Zhan et al. 2003), indicating extensive genetic diversity between *Z. tritici* populations. Similarly, we observed mixed *Z. tritici* populations when we performed the structure analysis, likely suggesting high rates of gene flow (Linde et al. 2002; Zhan et al. 2003). In contrast, recent population genomics studies displayed genetic relatedness between pathogen populations of the same origin (El Chartouni et al. 2011; Hartmann et al. 2017). When we assessed the diversity of *Z. tritici* chromosomes in our global collection, we observed, in line with previous work, highly conserved core chromosomes with extensive size polymorphisms and variable numbers and sizes of accessory chromosomes (Croll et al. 2013; Habig et al. 2017; Wittenberg et al. 2009). *Z. tritici* accessory chromosomes can also be lost during asexual propagation (Moeller et al. 2018), and their loss might be linked to increased aggressiveness on wheat as suggested for the cvs. Titlis and Runal (Habig et al. 2017; Stewart et al. 2018).

Lastly, we assessed the polymorphism of effector genes and observed high numbers of conserved effectors affected by non-synonymous SNPs, highlighting the importance of such effectors in *Z. tritici* virulence. This is in contrast to other fungal pathogens, such as *Fusarium oxysporum* and *Alternaria alternata*, where effector genes or host-specific toxin genes are mainly located on accessory chromosomes (Ma et al. 2010; Thomma 2003). Thus, our study highlights the significant genetic diversity in global populations of *Z. tritici* and signifies the importance of conserved effectors in the evolution of virulence on host plants.

In conclusion, our data demonstrate the extensive variability in a global population of *Z. tritici* isolates. We have demonstrated that the largest genetic diversity was observed in isolates originating from the centre of origin of the pathogen and the host, and observed four effector genes affected by non-synonymous SNPs that might contribute to host specificity. The extensive genetic diversity among *Z. tritici* isolates highlights its evolutionary potential to infect and overcome resistance in new wheat varieties. This challenges the overall question of durable management of this important wheat disease.

ACKNOWLEDGEMENTS

MFS acknowledges the receipt of a VENI grant of ALW-NWO.

REFERENCES

Adhikari, T. B., Yang, X., Cavaletto, J. R., Hu, X., Buechley, G., Ohm, H. W., Shaner, G., and Goodwin, S. B. 2004. Molecular mapping of *Stb1*, a potentially durable gene for resistance to septoria tritici blotch in wheat. Theoretical and Applied Genetics 109:944-953.

Bertels, F., Silander, O. K., Pachkov, M., Rainey, P. B., and van Nimwegen, E. 2014. Automated reconstruction of whole-genome phylogenies from short-sequence reads. Molecular Biology and Evolution 31:1077-1088.

Brown, J. K. M., Chartrain, L., Lasserre-Zuber, P., and Saintenac, C. 2015. Genetics of resistance to *Zymoseptoria tritici* and applications to wheat breeding. Fungal Genetics and Biology 79:33-41.

Brunner, P. C., and McDonald, B. A. 2018. Evolutionary analyses of the avirulence effector AvrStb6 in global populations of *Zymoseptoria tritici* identify candidate amino acids involved in recognition. Molecular Plant Pathology. https://doi.org/10.1111/mpp.12662E-pub ahead of print.

Chartrain, L., Brading, P. A., and Brown, J. K. M. 2005a. Presence of the *Stb6* gene for resistance to septoria tritici blotch (*Mycosphaerella graminicola*) in cultivars used in wheat-breeding programmes worldwide. Plant Pathology 54:134-143.

Chartrain, L., Joaquim, P., Berry, S. T., Arraiano, L. S., Azanza, F., and Brown, J. K. 2005b. Genetics of resistance to septoria tritici blotch in the Portuguese wheat breeding line TE 9111. Theoretical and Applied Genetics 110:1138-1144.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., and Ruden, D. M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly 6:80-92.

Cook, D. E., Mesarich, C. H., and Thomma, B. P. H. J. 2015. Understanding plant immunity as a surveillance system to detect invasion. Annual Review of Phytopathology 53:541-563.

Cordo, C. A., Mónaco, C. I., Altamirano, R., Perelló, A., Larrán, S., Kripelz, N. I., and Simón, M. 2017. Weather conditions associated with the release and dispersal of *Zymoseptoria tritici* spores in the Argentine Pampas region. International Journal of Agronomy 2017.

Croll, D., Lendenmann, M. H., Stewart, E., and McDonald, B. A. 2015. The impact of recombination hotspots on genome evolution of a fungal plant pathogen. Genetics 201: 1213–1228.

Croll, D., Zala, M., and McDonald, B. A. 2013. Breakage-fusion-bridge cycles and large insertions contribute to the rapid evolution of accessory chromosomes in a fungal pathogen. PLoS Genetics 9:e1003567.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G. 2011. The variant call format and VCFtools. Bioinformatics 27:2156-2158.

de Jonge, R., Bolton, M. D., Kombrink, A., van den Berg, G. C., Yadeta, K. A., and Thomma, B. P. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Research 23:1271-1282.

de Jonge, R., Peter van Esse, H., Maruthachalam, K., Bolton, M. D., Santhanam, P., Saber, M. K., Zhang, Z., Usami, T., Lievens, B., Subbarao, K. V., and Thomma, B. P. H. J. 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. Proceedings of the National Academy of Sciences 109:5110-5115.

de Jonge, R., and Thomma, B. P. H. J. 2009. Fungal LysM effectors: extinguishers of host immunity? Trends in Microbiology 17:151-157.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.

Dong, S., Raffaele, S., and Kamoun, S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Current Opinion in Genetics & Development 35:57-65.

El Chartouni, L., Tisserant, B., Siah, A., Duyme, F., Leducq, J.-B., Deweer, C., Fichter-Roisin, C., Sanssené, J., Durand, R., Halama, P., and Reignault, P. 2011. Genetic diversity and population structure in French populations of *Mycosphaerella graminicola*. Mycologia 103:764-774.

Faino, L., Seidl, M. F., Shi-Kunne, X., Pauper, M., van den Berg, G. C. M., Wittenberg, A. H. J., and Thomma, B. P. H. J. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome Research 26:1091-1100.

Falush, D., Stephens, M., and Pritchard, J. K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.

FAOSTAT, F. 2015. FAO Statistics Division; 2014.

Fisher, M. C., Hawkins, N. J., Sanglard, D., and Gurr, S. J. 2018. Worldwide emergence of resistance to antifungal drugs challenges human health and food security. Science 360:739-742.

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484:186-194.

Flutre, T., Inizan, O., Hoede, C., and Quesneville, H. 2010. REPET: pipelines for the identification and annotation of transposable elements in genomic sequences. Pages 9-13 in: Proceedings of the Plant & Animal Genome (PAG) XVIII Conference, San Diego, CA, USA.

Fouché, S., Plissonneau, C., and Croll, D. 2018a. The birth and death of effectors in rapidly evolving filamentous pathogen genomes. Current Opinion in Microbiology. 46:34-42.

Fouché, S., Plissonneau, C., McDonald, B. A., and Croll, D. 2018b. Meiosis leads to pervasive copy-number variation and distorted inheritance of accessory chromosomes of the wheat pathogen *Zymoseptoria tritici*. Genome Biology Evolution 10:1416-1429.

Gibriel, H. A. Y., Thomma, B. P. H. J., and Seidl, M. F. 2016. The age of effectors: genome-based discovery and applications. Phytopathology 106: 1206–1212.

Goodwin, S. B., M'Barek, S. B., Dhillon, B., Wittenberg, A. H., Crane, C. F., Hane, J. K., Foster, A. J., Lee, T. A., Grimwood, J., and Aerts, A. 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLoS Genetics 7:e1002070.

Grandaubert, J., Dutheil, J. Y., and Stukenbrock, E. H. 2018. The genomic determinants of adaptive evolution in a fungal pathogen. bioRxiv: 176727.

Grünwald, N. J., McDonald, B. A., and Milgroom, M. G. 2016. Population genomics of fungal and oomycete pathogens. Annual Review of Phytopathology 54:323-346.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M. V., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I. B., Bulone, V., Cai, G., Cakir, C., Carrington, J. C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M. A., Fugelstad, J., Gilroy, E. M., Gnerre, S., Green, P. J., Grenville-Briggs, L. J., Griffith, J., Grünwald, N. J., Horn, K., Horner, N. R., Hu, C.-H., Huitema, E., Jeong, D.-H., Jones, A. M. E., Jones, J. D. G., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J. G., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzón, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., Savidor, A., Schornack, S., Schwartz, D. C., Schumann, U. D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D. J., Sykes, S., Thines, M., van de Vondervoort, P. J. I., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R. J., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Habig, M., Quade, J., and Stukenbrock, E. H. 2017. Forward genetics approach reveals host genotype-dependent importance of accessory chromosomes in the fungal wheat pathogen *Zymoseptoria tritici*. mBio 8:e01919-17.

Hartmann, F. E., McDonald, B. A., and Croll, D. 2018. Genome-wide evidence for divergent selection between populations of a major agricultural pathogen. Molecular Ecology 9:656-617.

Hartmann, F. E., Sanchez-Vallet, A., McDonald, B. A., and Croll, D. 2017. A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. ISME Journal 11:1189-1204.

Hubbard, A., Lewis, C., Yoshida, K., Ramirez-Gonzalez, R., de Vallavieille-Pope, C., Thomas, J., Kamoun, S., Bayles, R., Uauy, C., and Saunders, D. 2015. Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. Genome Biology 16:23.

Kema, G., Annone, J., Sayoud, R., and Van Silfhout, C. 1996. Genetic variation for virulence and resistance in the wheat-*Mycosphaerella graminicola* pathosystem. I. Interactions between pathogen isolates and host cultivars. Phytopathology 86:200-212.

Kema, G. H. J., Mirzadi Gohari, A., Aouini, L., Gibriel, H. A. Y., Ware, S. B., van den Bosch, F., Manning-Smith, R., Alonso-Chavez, V., Helps, J., Ben M'Barek, S., Mehrabi, R., Diaz-Trujillo, C., Zamani, E., Schouten, H. J., van der Lee, T. A. J., Waalwijk, C., de Waard, M. A., De Wit, P. J. G. M., Verstappen, E. C. P., Thomma, B. P. H. J., Meijer, H. J. G., and Seidl, M. F. 2018. Stress and sexual reproduction affect the dynamics of the wheat pathogen effector AvrStb6 and strobilurin resistance. Nature Genetics 23:678.

Kettles, G. J., and Kanyuka, K. 2016. Dissecting the molecular interactions between wheat and the fungal pathogen *Zymoseptoria tritici*. Frontiers in Plant Science 7:508.

Kolde, R., and Kolde, M. R. 2015. Package 'pheatmap'.

Kombrink, A., Rovenich, H., Shi-Kunne, X., Rojas-Padilla, E., van den Berg, G. C. M., Domazakis, E., de Jonge, R., Valkenburg, D.-J., Sánchez-Vallet, A., Seidl, M. F., and Thomma, B. P. H. J. 2016. *Verticillium dahliae* LysM effectors differentially contribute to virulence on plant hosts. Molecular Plant Pathology 18: 596–608.

Kombrink, A., and Thomma, B. P. H. J. 2013. LysM Effectors: Secreted Proteins Supporting Fungal Life. PLoS Pathogens 9:e1003769.

Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357-359.

Li, H., and Durbin, R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589-595.

Linde, C. C., Zhan, J., and McDonald, B. A. 2002. Population structure of *Mycosphaerella graminicola*: from lesions to continents. Phytopathology 92:946-955.

Ma, L. J., Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M. J., Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., and Henrissat, B. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367-373.

Marshall, R., Kombrink, A., Motteram, J., Loza-Reyes, E., Lucas, J., Hammond-Kosack, K. E., Thomma, B. P. H. J., and Rudd, J. J. 2011. 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 Physiology 156:756-769.

McDonald, M. C., McDonald, B. A., and Solomon, P. S. 2015. Recent advances in the *Zymoseptoria tritici*-wheat interaction: insights from pathogenomics. Frontiers in Plant Science 6:102.

McDonald, M. C., McGinness, L., Hane, J. K., Williams, A. H., Milgate, A., and Solomon, P. S. 2016. Utilizing gene tree variation to identify candidate effector genes in *Zymoseptoria tritici*. G3 6:779-791.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20:1297-1303.

McMullan, M., Rafiqi, M., Kaithakottil, G., Clavijo, B. J., Bilham, L., Orton, E., Percival-Alwyn, L., Ward, B. J., Edwards, A., Saunders, D. G. O., Garcia Accinelli, G., Wright, J., Verweij, W., Koutsovoulos, G., Yoshida, K., Hosoya, T., Williamson, L., Jennings, P., Ioos, R., Husson, C., Hietala, A. M., Vivian-Smith, A., Solheim, H., MaClean, D., Fosker, C., Hall, N., Brown, J. K. M., Swarbreck, D., Blaxter, M., Downie, J. A., and Clark, M. D. 2018. The ash dieback invasion of Europe was founded by two genetically divergent individuals. Nature Ecology Evolution 381:235.

Mehrabi, R., Taga, M., and Kema, G. H. 2007. Electrophoretic and cytological karyotyping of the foliar wheat pathogen *Mycosphaerella graminicola* reveals many chromosomes with a large size range. Mycologia 99:868-876.

Meile, L., Croll, D., Brunner, P. C., Plissonneau, C., Hartmann, F. E., McDonald, B. A., and Sánchez-Vallet, A. 2018a. A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance to septoria tritici blotch. bioRxiv:264226.

Meile, L., Croll, D., Brunner, P. C., Plissonneau, C., Hartmann, F. E., McDonald, B. A., and Sánchez-Vallet, A. 2018b. A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance to septoria tritici blotch. New Phytologist 65:512.

Mirzadi Gohari, A., Ware, S. B., Wittenberg, A. H. J., Mehrabi, R., Ben M'Barek, S., Verstappen, E. C. P., van der Lee, T. A. J., Robert, O., Schouten, H. J., de Wit, P. P. J. G. M., and Kema, G. H. J. 2015. Effector discovery in the fungal wheat pathogen *Zymoseptoria tritici*. Molecular Plant Pathology 16:931-945.

Moeller, M., Habig, M., Freitag, M., and Holtgrewe Stukenbrock, E. 2018. Extraordinary genome instability and widespread chromosome rearrangements during vegetative growth. bioRxiv:304915.

Möller, M., and Stukenbrock, E. H. 2017. Evolution and genome architecture in fungal plant pathogens. Nature Publishing Group 484:1-16.

Naouari, M., Siah, A., Elgazzah, M., Reignault, P., and Halama, P. 2016. Mitochondrial DNA-based genetic diversity and population structure of *Zymoseptoria tritici* in Tunisia. European Journal of Plant Pathology 146:305-314.

O'Driscoll, A., Kildea, S., Doohan, F., Spink, J., and Mullins, E. 2014. The wheat-Septoria conflict: a new front opening up? Trends in Plant Science 19:602-610.

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods 8: 785-786.

Plissonneau, C., Benevenuto, J., Mohd-Assaad, N., Fouché, S., Hartmann, F. E., and Croll, D. 2017. Using population and comparative genomics to understand the genetic basis of effector-driven fungal pathogen evolution. Frontiers in Plant Sciences 8:119.

Plissonneau, C., Hartmann, F. E., and Croll, D. 2018. Pangenome analyses of the wheat pathogen *Zymoseptoria tritici* reveal the structural basis of a highly plastic eukaryotic genome. BMC Biology 16:673.

Plissonneau, C., Stürchler, A., and Croll, D. 2016. The evolution of orphan regions in genomes of a fungal pathogen of wheat. mBio 7: e01231-16.

Poppe, S., Dorsheimer, L., Happel, P., and Stukenbrock, E. H. 2015. Rapidly evolving genes are key players in host specialization and virulence of the fungal wheat pathogen *Zymoseptoria tritici* (*Mycosphaerella graminicola*). PLoS Pathogens 11:e1005055.

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Quinlan, A. R., and Hall, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26: 841-842.

Raffaele, S., Farrer, R. A., Cano, L. M., Studholme, D. J., MacLean, D., Thines, M., Jiang, R. H., Zody, M. C., Kunjeti, S. G., Donofrio, N. M., Meyers, B. C., Nusbaum, C., and Kamoun, S. 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science 330:1540-1543.

Raffaele, S., and Kamoun, S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nature Reviews Microbiology 10:417-430.

Rovenich, H., Boshoven, J. C., and Thomma, B. P. H. J. 2014. Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. Current Opinion in Plant Biology 20:96-103.

Rudd, J. J., Kanyuka, K., Hassani-Pak, K., Derbyshire, M., Andongabo, A., Devonshire, J., Lysenko, A., Saqi, M., Desai, N. M., Powers, S. J., Hooper, J., Ambroso, L., Bharti, A., Farmer, A., Hammond-Kosack, K. E., Dietrich, R. A., and Courbot, M. 2015. Transcriptome and metabolite profiling of the infection cycle of *Zymoseptoria tritici* on wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle definition. Plant Physiology 167:1158-1185.

Saintenac, C., Lee, W.-S., Cambon, F., Rudd, J. J., King, R. C., Marande, W., Powers, S. J., Bergès, H., Phillips, A. L., Uauy, C., Hammond-Kosack, K. E., Langin, T., and Kanyuka, K. 2018. Wheat receptor-kinase-like protein Stb6 controls gene-for-gene resistance to fungal pathogen *Zymoseptoria tritici*. Nature Genetics 51:291.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.

Seidl, M. F., and Thomma, B. 2017a. Transposable elements direct the coevolution between plants and microbes. Trends in Genetics 33:842-851.

Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., Herrera-Foessel, S., Singh, P. K., Singh, S., and Govindan, V. 2011. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annual Review of Phytopathology 49:465-481.

Smit, A., Hubley, R., and Green, P. 2016. RepeatMasker Open-4.0. Available from: http://www.repeatmasker.org

Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., Manners, J. M., and Taylor, J. M. 2015b. EffectorP: predicting fungal effector proteins from secretomes using machine learning. New Phytologist 210:743–761.

Stewart, E. L., Croll, D., Lendenmann, M. H., Sánchez-Vallet, A., Hartmann, F. E., Palma-Guerrero, J., Ma, X., and McDonald, B. A. 2018. Quantitative trait locus mapping reveals complex genetic architecture of quantitative virulence in the wheat pathogen *Zymoseptoria tritici*. Molecular Plant Pathology 19:201-216.

Stukenbrock, E. H., Banke, S., Javan-Nikkhah, M., and McDonald, B. A. 2006. Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. Molecular Biology and Evolution 24:398-411.

Stukenbrock, E. H., and Bataillon, T. 2012. A population genomics perspective on the emergence and adaptation of new plant pathogens in agro-ecosystems. PLoS Pathogens 8:e1002893.

Stukenbrock, E. H., Bataillon, T., Dutheil, J. Y., Hansen, T. T., Li, R., Zala, M., McDonald, B. A., Wang, J., and Schierup, M. H. 2011. The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. Genome Research 21:2157-2166.

Stukenbrock, E. H., Jørgensen, F. G., Zala, M., Hansen, T. T., McDonald, B. A., and Schierup, M. H. 2010. Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen *Mycosphaerella graminicola*. PLoS Genetics 6:e1001189.

Thomma, B. P. H. J. 2003. Alternaria spp.: from general saprophyte to specific parasite. Molecular Plant Pathology 4:225-236.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., and Pachter, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology 28: 511-515.

Troccoli, A., Borrelli, G. M., De Vita, P., Fares, C., and Di Fonzo, N. 2000. Mini Review: Durum wheat quality: A multidisciplinary concept. Journal of Cereal Science 32:99-113.

Wickham, H. 2015. ggplot2: elegant graphics for data analysis.[Internet]. 2009. New York, USA 907.

Wittenberg, A. H. J., van der Lee, T. A. J., Ben M'Barek, S., Ware, S. B., Goodwin, S. B., Kilian, A., Visser, R. G. F., Kema, G. H. J., and Schouten, H. J. 2009. Meiosis drives extraordinary genome plasticity in the haploid fungal plant pathogen *Mycosphaerella graminicola*. PLoS One 4:e5863.

Zhan, J., Pettway, R. E., and McDonald, B. A. 2003. The global genetic structure of the wheat pathogen *Mycosphaerella graminicola* is characterized by high nuclear diversity, low mitochondrial diversity, regular recombination, and gene flow. Fungal Genetics and Biology: 38:286-297.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28:3326-3328.

Zhong, Z., Marcel, T. C., Hartmann, F. E., Ma, X., Plissonneau, C., Zala, M., Ducasse, A., Confais, J., Compain, J., Lapalu, N., Amselem, J., McDonald, B. A., Croll, D., and Palma-Guerrero, J. 2017. A small secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat cultivars carrying the *Stb6* resistance gene. New Phytologist 31:166.

SUPPLEMENTAL INFORMATION

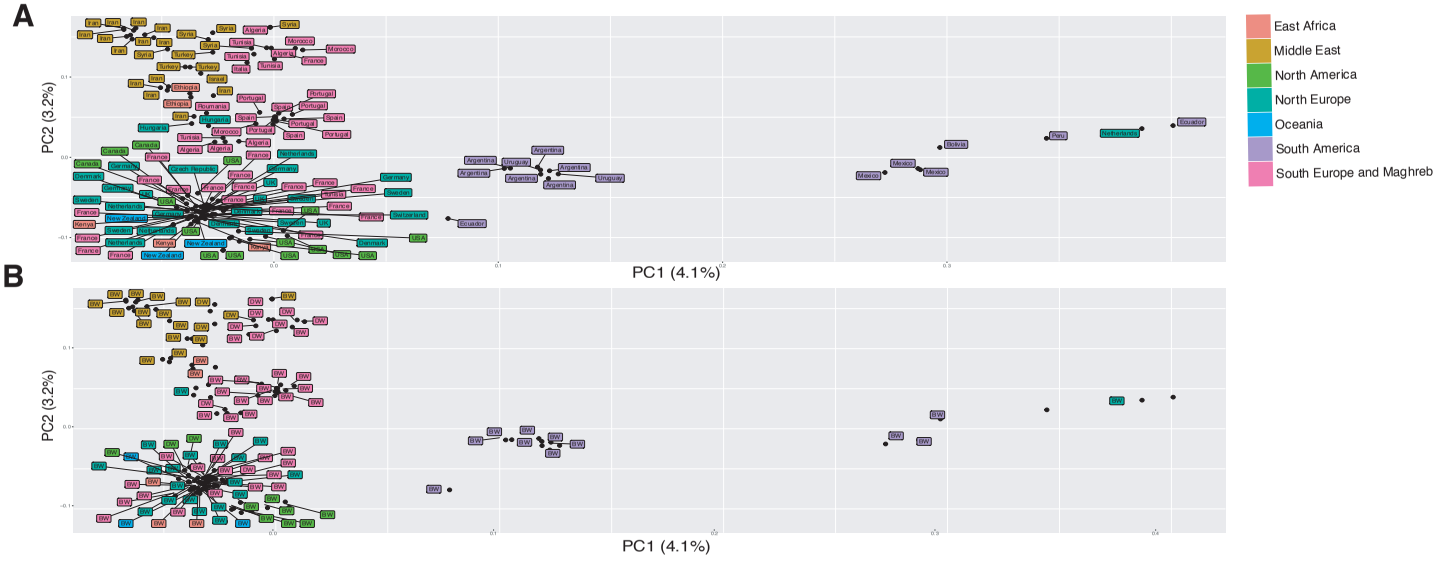


Figure S1. *Zymoseptoria tritici* isolates do not cluster by geographical or host origin. PCAs based on 877,279 SNPs across all *Z. tritici* isolates linked to **(A)** country of origin, and **(B)** host origin.

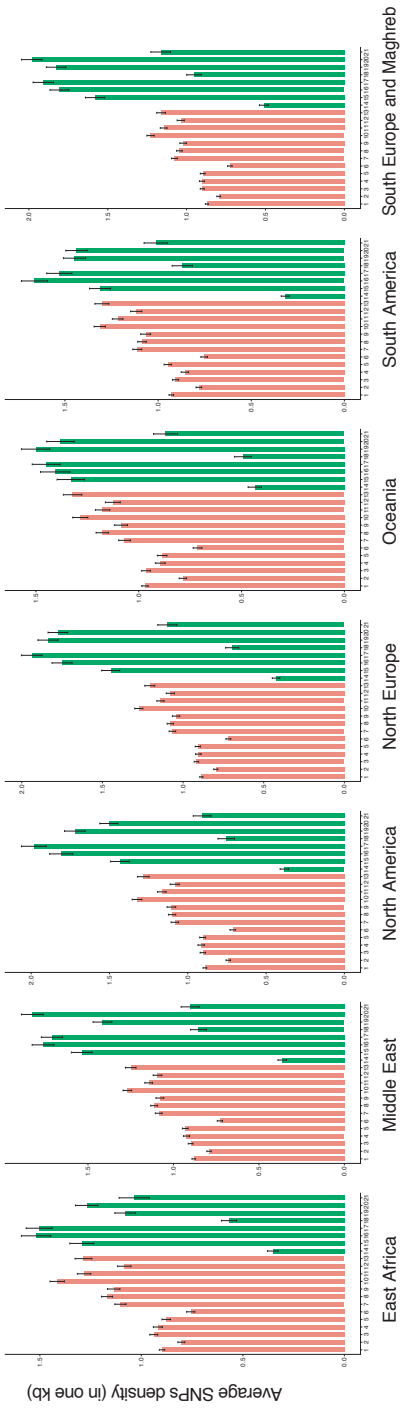


Figure S2. Comparative SNP densities in core and accessory chromosomes in all *Zymoseptoria tritici* isolates. SNPs densities are determined in one kb windows for core (red) and accessory chromosomes (green).

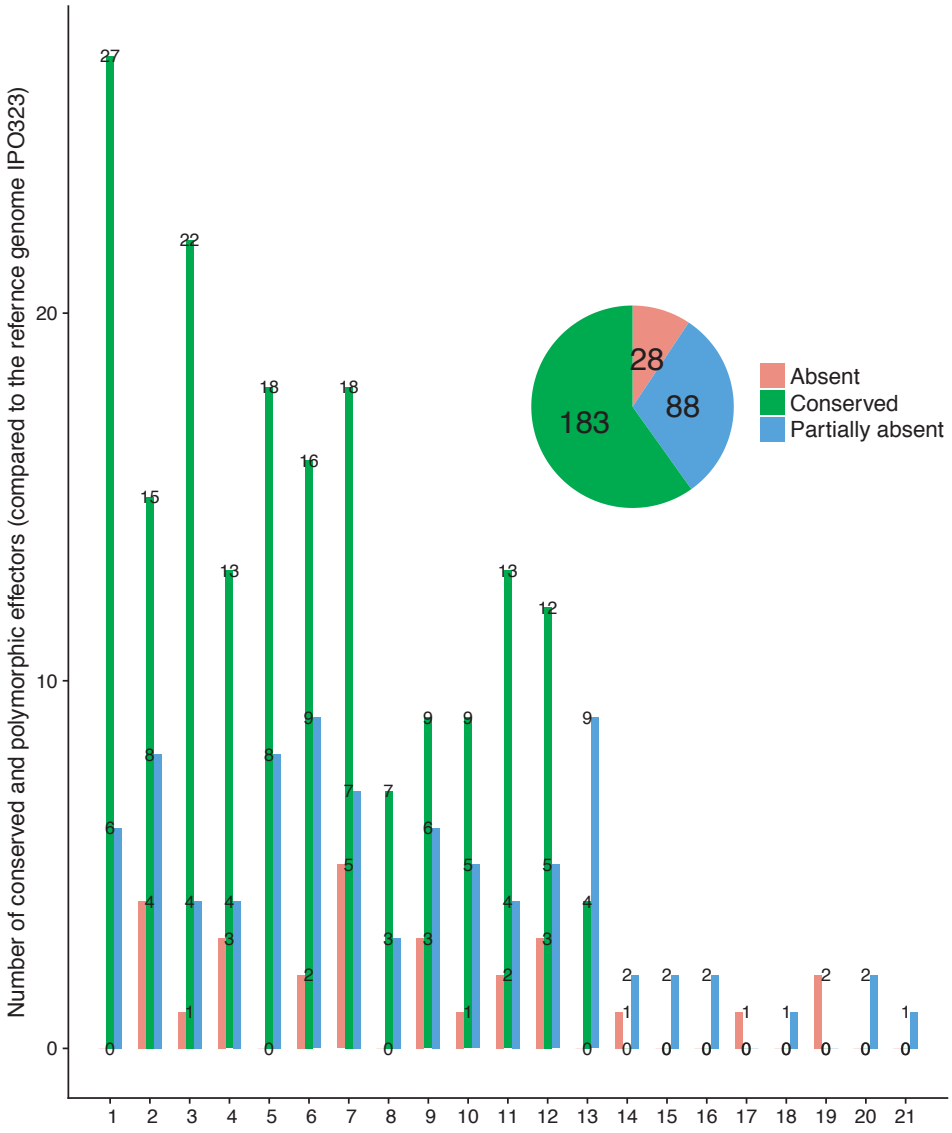


Figure S3. Abundance of conserved effectors in the global *Zymoseptoria tritici* panel. The pie chart shows the total number of conserved, partially absent, and absent effectors. Bar plots indicate the number of effectors per class on each chromosome.

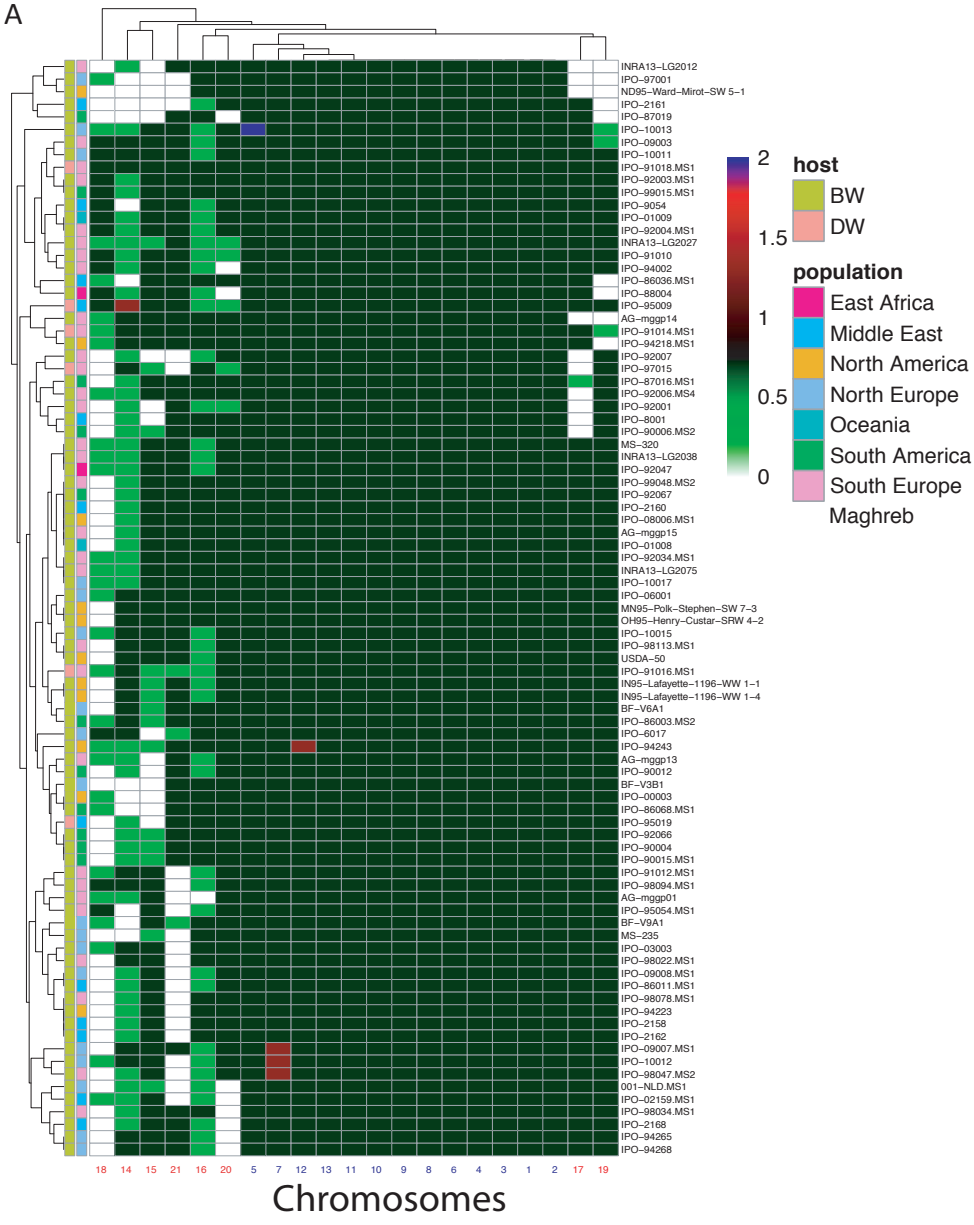


Figure S4. Presence/absence chromosomal polymorphisms in *Zymoseptoria tritici* isolates adapted to bread or durum wheat. (A) Heatmap indicating presence/absence polymorphisms in each chromosome in bread and durum wheat isolates. Colors indicate presence (black), partially duplicated (red), duplicated (red), partially deleted (green) or absent (white).

B

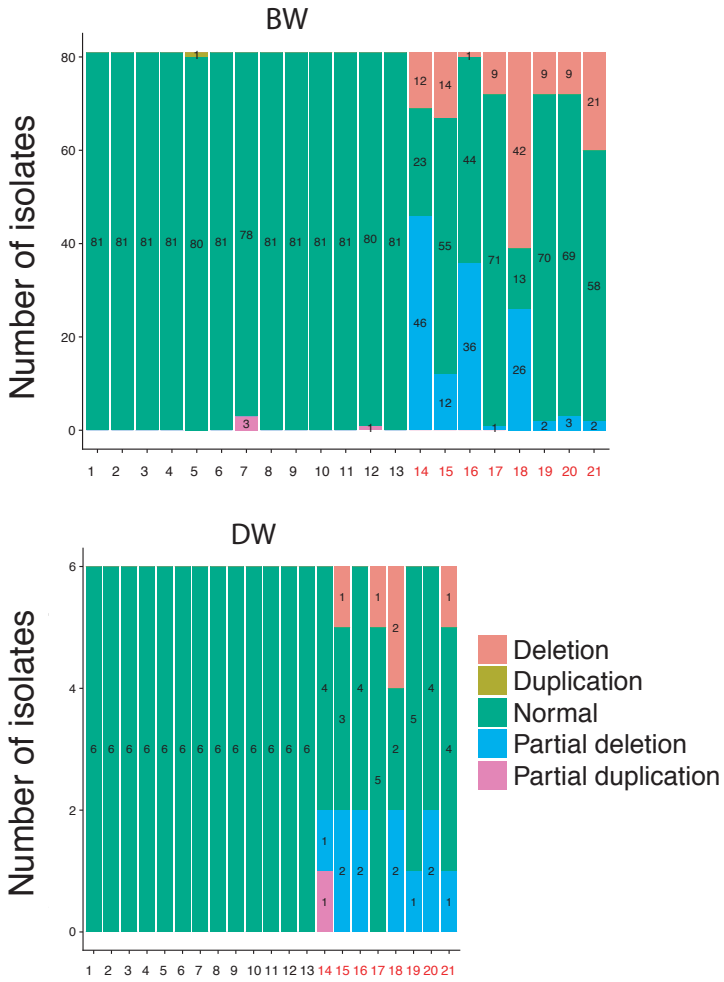


Figure S4 *continued*. **(B)** The frequencies of presence, partial duplication, duplication, partial deletion, and deletion are shown for each of the 21 chromosomes depicted as vertical bars. **(C)** Heatmap indicating the presence (black) and absence (white) of effectors in bread and durum wheat *Z. tritici* isolates based on IPO323 gene annotation. The identified *Z. tritici* effector AvrStb6 is indicated (green color).

C

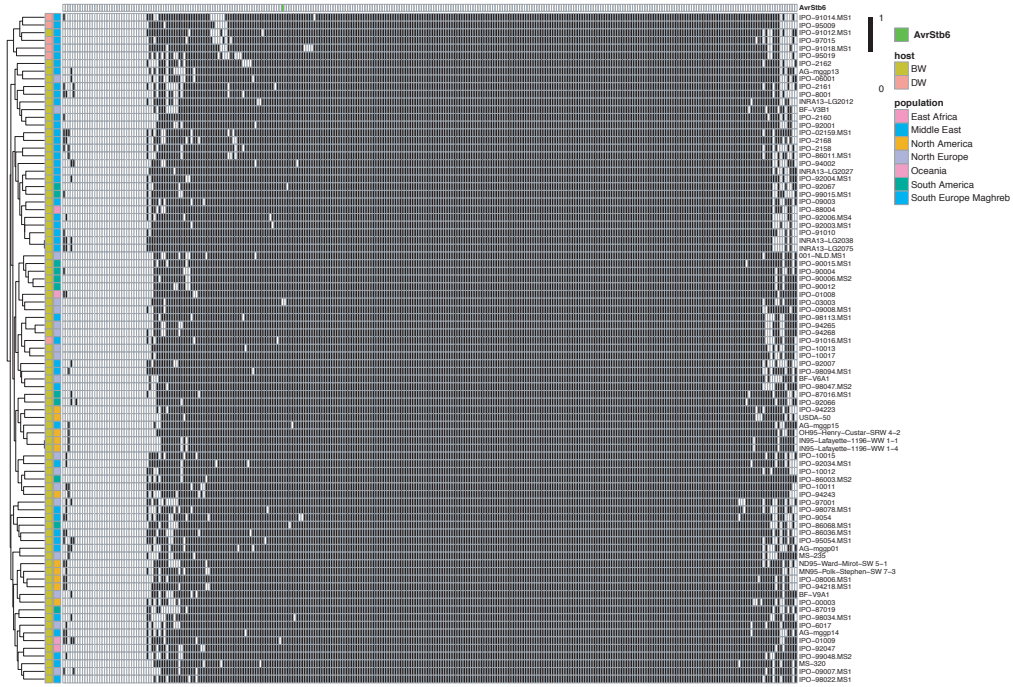

AvrStb6
IPO−91014.MS1
IPO−95009
IPO−91012.MS1
IPO−97015
IPO−91018.MS1
IPO−95019
IPO−2162
AG−mggp13
IPO−08001
IPO−2161
IPO−8001
INRA13−LG2012
BF−V3B1
IPO−2160
IPO−92001
IPO−02159.MS1
IPO−2168
IPO−2158
IPO−86011.MS1
IPO−94002
INRA13−LG2027
IPO−92004.MS1
IPO−92067
IPO−99015.MS1
IPO−09003
IPO−88004
IPO−92006.MS4
IPO−92003.MS1
IPO−91010
INRA13−LG2038
INRA13−LG2075
001−NLD.MS1
IPO−90015.MS1
IPO−90004
IPO−90006.MS2
IPO−90012
IPO−01008
IPO−03003
IPO−09008.MS1
IPO−98113.MS1
IPO−94265
IPO−94268
IPO−91016.MS1
IPO−10013
IPO−10017
IPO−92007
IPO−98094.MS1
BF−V6A1
IPO−98047.MS2
IPO−87016.MS1
IPO−92066
IPO−94223
USDA−50
AG−mggp15
CH95−Henry−Custar−SRW 4−2
IN95−Lafayette−1196−WW 1−1
IN95−Lafayette−1196−WW 1−4
IPO−10015
IPO−92034.MS1
IPO−10012
IPO−86003.MS2
IPO−10011
IPO−94243
IPO−97001
IPO−98078.MS1
IPO−9054
IPO−86068.MS1
IPO−86036.MS1
IPO−95054.MS1
AG−mggp01
MS−235
ND95−Ward−Mirot−SW 5−1
MN95−Polk−Stephen−SW 7−3
IPO−08006.MS1
IPO−94218.MS1
BF−V9A1
IPO−00003
IPO−87019
IPO−98034.MS1
IPO−6017
AG−mggp14
IPO−01009
IPO−92047
IPO−99048.MS2
MS−320
IPO−09007.MS1
IPO−98022.MS1
1
0
AvrStb6
host
BW
DW
population
East Africa
Middle East
North America
North Europe
Oceania
South America
South Europe Maghreb

Table S1. *Z. tritici* isolates used in this study.

Isolate	Country	Population	Host origin
001-NLD.MS1	Netherlands	North Europe	BW
AG-mggp01	France	South Europe and Maghreb	BW
AG-mggp12	France	South Europe and Maghreb	BW
AG-mggp13	France	South Europe and Maghreb	BW
AG-mggp14	France	South Europe and Maghreb	BW
AG-mggp15	France	South Europe and Maghreb	BW
BF-V1.1	UK	North Europe	BW
BF-V3B1	UK	North Europe	BW
BF-V6A1	UK	North Europe	BW
BF-V9A1	UK	North Europe	BW
Hung2.MS1	Hungaria	North Europe	NA
IN95-Lafayette-1196-WW 1-1	USA	North America	NA
IN95-Lafayette-1196-WW 1-4	USA	North America	NA
INRA08-FS0001.MS1	France	South Europe and Maghreb	BW
INRA08-FS0002.MS2	France	South Europe and Maghreb	BW
INRA08-FS0003.MS1	France	South Europe and Maghreb	BW
INRA13-LG2012	Spain	South Europe and Maghreb	BW
INRA13-LG2027	Spain	South Europe and Maghreb	BW
INRA13-LG2038	Spain	South Europe and Maghreb	BW
INRA13-LG2075	Spain	South Europe and Maghreb	BW
IPO-00003	USA	North America	BW
IPO-00005	USA	North America	BW
IPO-01008	New Zealand	Oceania	BW
IPO-01009	New Zealand	Oceania	BW
IPO-02159.MS1	Iran	Middle East	BW
IPO-02166.MS1	Iran	Middle East	BW
IPO-03001	Germany	North Europe	BW
IPO-03003	Germany	North Europe	BW
IPO-052462	NA	NA	NA
IPO-06001	Denmark	North Europe	BW
IPO-06008	Denmark	North Europe	BW
IPO-08002	Iran	Middle East	BW
IPO-08004	Canada	North America	DW

Table S1. ***continued***

Isolate	Country	Population	Host origin
IPO-08006.MS1	Canada	North America	DW
IPO-09003	France	South Europe and Maghreb	BW
IPO-09005	France	South Europe and Maghreb	BW
IPO-09007.MS1	Germany	North Europe	BW
IPO-09008.MS1	Germany	North Europe	BW
IPO-09009.MS1	Germany	North Europe	BW
IPO-10011	Sweden	North Europe	NA
IPO-10012	Sweden	North Europe	NA
IPO-10013	Sweden	North Europe	NA
IPO-10014	Sweden	North Europe	NA
IPO-10015	Sweden	North Europe	NA
IPO-10017	Sweden	North Europe	NA
IPO-1004	New Zealand	Oceania	BW
IPO-2158	Iran	Middle East	BW
IPO-2160	Iran	Middle East	BW
IPO-2161	Iran	Middle East	BW
IPO-2162	Iran	Middle East	BW
IPO-2164	Iran	Middle East	BW
IPO-2165	Iran	Middle East	BW
IPO-2168	Iran	Middle East	BW
IPO-6017	Denmark	North Europe	BW
IPO-6026	Denmark	North Europe	BW
IPO-8001	Iran	Middle East	BW
IPO-86002.MS1	Bolivia	South America	BW
IPO-86003.MS2	Ecuador	South America	BW
IPO-86011.MS1	Turkey	Middle East	BW
IPO-86013.MS1	Turkey	Middle East	BW
IPO-86015.MS1	Morocco	South Europe and Maghreb	BW
IPO-86022.MS1	Turkey	Middle East	DW
IPO-86036.MS1	Israel	Middle East	NA
IPO-86068.MS1	Argentina	South America	BW
IPO-87016.MS1	Uruguay	South America	BW
IPO-87019	Uruguay	South America	BW
IPO-88004	Ethiopia	East Africa	BW

Table S1. ***continued***

Isolate	Country	Population	Host origin
IPO-88018.MS1	Ethiopia	East Africa	BW
IPO-89011.MS1	Netherlands	North Europe	BW
IPO-90004	Mexico	South America	BW
IPO-90006.MS2	Mexico	South America	NA
IPO-90012	Mexico	South America	BW
IPO-90015.MS1	Peru	South America	NA
IPO-9010	Iran	Middle East	NA
IPO-9040	Iran	Middle East	NA
IPO-9054	Iran	Middle East	NA
IPO-91003.MS1	Syria	Middle East	NA
IPO-91004.MS1	Syria	Middle East	DW
IPO-91009.MS1	Tunisia	South Europe and Maghreb	DW
IPO-91010	Tunisia	South Europe and Maghreb	DW
IPO-91012.MS1	Tunisia	South Europe and Maghreb	DW
IPO-91014.MS1	Tunisia	South Europe and Maghreb	DW
IPO-91016.MS1	Tunisia	South Europe and Maghreb	DW
IPO-91018.MS1	Morocco	South Europe and Maghreb	DW
IPO-91020.MS1	Morocco	South Europe and Maghreb	DW
IPO-92001	Portugal	South Europe and Maghreb	BW
IPO-92002	Portugal	South Europe and Maghreb	BW
IPO-92003.MS1	Portugal	South Europe and Maghreb	BW
IPO-92004.MS1	Portugal	South Europe and Maghreb	BW
IPO-92006.MS4	Portugal	South Europe and Maghreb	BW
IPO-92007	Portugal	South Europe and Maghreb	BW
IPO-92034.MS1	Algeria	South Europe and Maghreb	BW
IPO-92044	Kenya	East Africa	BW
IPO-92047	Kenya	East Africa	BW
IPO-92062	Kenya	East Africa	BW
IPO-92064	Argentina	South America	BW
IPO-92066	Argentina	South America	BW
IPO-92067	Argentina	South America	NA
IPO-93014	Argentina	South America	BW
IPO-94002	Algeria	South Europe and Maghreb	BW
IPO-94203	USA	North America	BW

Table S1. ***continued***

Isolate	Country	Population	Host origin
IPO-94218.MS1	Canada	North America	BW
IPO-94223	USA	North America	BW
IPO-94234	USA	North America	BW
IPO-94243	USA	North America	BW
IPO-94265	Netherlands	North Europe	BW
IPO-94268	Netherlands	North Europe	BW
IPO-9455	France	South Europe and Maghreb	BW
IPO-95001	Switzerland	North Europe	NA
IPO-95009	Syria	Middle East	BW
IPO-95019	Syria	Middle East	DW
IPO-95036.MS1	Syria	Middle East	BW
IPO-95050	Algeria	South Europe and Maghreb	DW
IPO-95054.MS1	Algeria	South Europe and Maghreb	BW
IPO-95062	Algeria	South Europe and Maghreb	DW
IPO-97001	Czech Republic	North Europe	BW
IPO-97015	Italia	South Europe and Maghreb	BW
IPO-98022.MS1	France	South Europe and Maghreb	BW
IPO-98034.MS1	France	South Europe and Maghreb	BW
IPO-98046.MS2	France	South Europe and Maghreb	BW
IPO-98047.MS2	France	South Europe and Maghreb	BW
IPO-98078.MS1	France	South Europe and Maghreb	BW
IPO-98094.MS1	France	South Europe and Maghreb	BW
IPO-98113.MS1	France	South Europe and Maghreb	BW
IPO-98114	Hungaria	North Europe	BW
IPO-99015.MS1	Argentina	South America	BW
IPO-99032	France	South Europe and Maghreb	NA
IPO-99048.MS2	France	South Europe and Maghreb	NA
MN95-Polk-Stephen-SW 7-3	USA	North America	NA
MS-235	Netherlands	North Europe	BW
MS-270	Ecuador	South America	NA
MS-320	Roumania	South Europe and Maghreb	NA
ND95-Ward-Mirot-SW 5-1	USA	North America	NA
OH95-Henry-Custar-SRW 4-2	USA	North America	NA
Paskeville	USA	North America	NA
USDA-50	USA	North America	BW

Table S2. Phenotyped *Z. tritici* isolates in this study

Isolate	Country	Population	Host origin	Average pycnidia Taichung	Average pycnidia Bulgaria	Average pycnidia TE	Average pycnidia Volcani
001-NLD.MS1	Netherlands	North Europe	BW	21	0	5	1
AG-mggp01	France	South Europe and Maghreb	BW	27	0.5	0.5	0
AG-mggp13	France	South Europe and Maghreb	BW	46	32	2	0
AG-mggp14	France	South Europe and Maghreb	BW	32	0	0	0
AG-mggp15	France	South Europe and Maghreb	BW	30.5	21	15	0.5
BF-V3B1	UK	North Europe	BW	25	2.222222222	0	0
BF-V6A1	UK	North Europe	BW	39	0.5	0.5	1.666666667
BF-V9A1	UK	North Europe	BW	14	0	1.6	0
IN95-Lafayette-1196-WW 1-1	USA	North America	NA	35	0	0	0
IN95-Lafayette-1196-WW 1-4	USA	North America	NA	21	0	0	0
INRA13-LG2012	Spain	South Europe and Maghreb	BW	33	37	0	8.571428571
INRA13-LG2027	Spain	South Europe and Maghreb	BW	14	17	0	0.5
INRA13-LG2038	Spain	South Europe and Maghreb	BW	50	4	37	0.5
INRA13-LG2075	Spain	South Europe and Maghreb	BW	34	14	7	1.5
IPO-00003	USA	North America	BW	22	36	4	0
IPO-01008	New Zealand	Oceania	BW	21	19	0	0
IPO-01009	New Zealand	Oceania	BW	34	0	0	0
IPO-02159.MS1	Iran	Middle East	BW	43	52	12	7.5
IPO-03003	Germany	North Europe	BW	25	28	6	1.666666667

Table S2. *continued*

Isolate	Country	Population	Host origin	Average pycnidia Taichung	Average pycnidia Bulgaria	Average pycnidia TE	Average pycnidia Volcani
IPO-06001	Denmark	North Europe	BW	35.55555556	24	13.5	2
IPO-08006.MS1	Canada	North America	DW	34.5	18.5	0	1.428571429
IPO-09003	France	South Europe and Maghreb	BW	24	0	15	0
IPO-09007.MS1	Germany	North Europe	BW	22	41.11111111	1	1.666666667
IPO-09008.MS1	Germany	North Europe	BW	11.5	24	10	0.714285714
IPO-10011	Sweden	North Europe	NA	34.5	0	1	0
IPO-10012	Sweden	North Europe	NA	40	10	1.5	0.714285714
IPO-10013	Sweden	North Europe	NA	17.5	1	0.5	0.555555556
IPO-10015	Sweden	North Europe	NA	22	0	0	2.857142857
IPO-10017	Sweden	North Europe	NA	20.5	0.5	0	2.5
IPO-2158	Iran	Middle East	BW	42	31	29	0
IPO-2160	Iran	Middle East	BW	66	46	51	0.833333333
IPO-2161	Iran	Middle East	BW	72	66	2	4.444444444
IPO-2162	Iran	Middle East	BW	79	37.5	0	10
IPO-2168	Iran	Middle East	BW	65.55555556	35.55555556	64	0
IPO-6017	Denmark	North Europe	BW	36	1	3	2.857142857
IPO-8001	Iran	Middle East	BW	64	66.66666667	0.5	9
IPO-86003.MS2	Ecuador	South America	BW	26.5	10	1	0
IPO-86011.MS1	Turkey	Middle East	BW	63	15.5	10	0
IPO-86036.MS1	Israel	Middle East	NA	36	3.5	3	4

Table S2. *continued*

Isolate	Country	Population	Host origin	Average pycnidia Taichung	Average pycnidia Bulgaria	Average pycnidia TE	Average pycnidia Volcani
IPO-86068.MS1	Argentina	South America	BW	40	16.66666667	0	6
IPO-87016.MS1	Uruguay	South America	BW	74	51	1.5	0
IPO-87019	Uruguay	South America	BW	18	29	0	4.285714286
IPO-88004	Ethiopia	East Africa	BW	45	18.88888889	0	0
IPO-90004	Mexico	South America	BW	13	0	0	0
IPO-90006.MS2	Mexico	South America	NA	68	45	45	0
IPO-90012	Mexico	South America	BW	15	0	0	0
IPO-90015.MS1	Peru	South America	NA	32	33	32	3.75
IPO-9054	Iran	Middle East	NA	20.55555556	1.5	10	0
IPO-91010	Tunisia	South Europe and Maghreb	DW	39	7	0	1.25
IPO-91012.MS1	Tunisia	South Europe and Maghreb	DW	36	8	0	0
IPO-92001	Portugal	South Europe and Maghreb	BW	40	47	0	0
IPO-92003.MS1	Portugal	South Europe and Maghreb	BW	14	0	0	0
IPO-92004.MS1	Portugal	South Europe and Maghreb	BW	26.66666667	1	0	1.5
IPO-92006.MS4	Portugal	South Europe and Maghreb	BW	50	29.5	1	2
IPO-92007	Portugal	South Europe and Maghreb	BW	17	22	28	0
IPO-92034.MS1	Algeria	South Europe and Maghreb	BW	33	18	7.5	7.5
IPO-92047	Kenya	East Africa	BW	19	12	1	6.25
IPO-92066	Argentina	South America	BW	28	9.5	0	2
IPO-92067	Argentina	South America	NA	36.5	8.888888889	0	0
IPO-94002	Algeria	South Europe and Maghreb	BW	26	0	0	0.5

Table S2. *continued*

Isolate	Country	Population	Host origin	Average pycnidia Taichung	Average pycnidia Bulgaria	Average pycnidia TE	Average pycnidia Volcani
IPO-94218.MS1	Canada	North America	BW	29	0	1.5	0
IPO-94223	USA	North America	BW	31	48	0	4
IPO-94243	USA	North America	BW	32.22222222	17	0	1
IPO-94265	Netherlands	North Europe	BW	14	5	0	0
IPO-94268	Netherlands	North Europe	BW	25	20	0	0
IPO-95054.MS1	Algeria	South Europe and Maghreb	BW	46	30	0.5	0
IPO-97001	Czech Republic	North Europe	BW	37	0	0	1.666666667
IPO-98022.MS1	France	South Europe and Maghreb	BW	2.5	13.33333333	0	0
IPO-98034.MS1	France	South Europe and Maghreb	BW	41	28	15	0
IPO-98047.MS2	France	South Europe and Maghreb	BW	17.77777778	14	0.5	0
IPO-98078.MS1	France	South Europe and Maghreb	BW	34	0	0	0
IPO-98094.MS1	France	South Europe and Maghreb	BW	23	16	0	0
IPO-98113.MS1	France	South Europe and Maghreb	BW	33	4	15	3
IPO-99015.MS1	Argentina	South America	BW	15	5	0	0
IPO-99048.MS2	France	South Europe and Maghreb	NA	36	0	0	0
MN95-Polk-Stephen-SW 7-3	USA	North America	NA	17	6	0	0
MS-235	Netherlands	North Europe	BW	24	24	16	0
MS-320	Roumania	South Europe and Maghreb	NA	48	9	0	0
ND95-Ward-Mirot-SW 5-1	USA	North America	NA	17	0	0	0

Table S2. *continued*

Isolate	Country	Population	Host origin	Average pycnidia Taichung	Average pycnidia Bulgaria	Average pycnidia TE	Average pycnidia Volcani
OH95-Henry-Custar-SRW 4-2	USA	North America	NA	37	2	0	0
USDA-50	USA	North America	BW	24.5	14	5	1.428571429
IPO-91014.MS1	Tunisia	South Europe and Maghreb	DW	0	0	0	37.5
IPO-91016.MS1	Tunisia	South Europe and Maghreb	DW	0	0	0	59
IPO-91018.MS1	Morocco	South Europe and Maghreb	DW	0	0	0	38
IPO-95009	Syria	Middle East	BW	0	0	0	34
IPO-95019	Syria	Middle East	DW	1.5	1	0	42.85714286
IPO-97015	Italia	South Europe and Maghreb	BW	0	0	0	38.57142857
IPO-02166.MS1	Iran	Middle East	BW	66	55	0	15
IPO-03001	Germany	North Europe	BW	19	1	1	12
IPO-1004	New Zealand	Oceania	BW	54	4	13	13.57142857
IPO-2164	Iran	Middle East	BW	78	52	58	20
IPO-2165	Iran	Middle East	BW	75	82.5	6.5	40
IPO-6026	Denmark	North Europe	BW	61	17	42	31.42857143
IPO-88018.MS1	Ethiopia	East Africa	BW	36	22.77777778	28.5	30
IPO-92002	Portugal	South Europe and Maghreb	BW	30	33.33333333	20	30
IPO-92064	Argentina	South America	BW	35	21.25	0	15
IPO-93014	Argentina	South America	BW	35	24	8	15
IPO-94203	USA	North America	BW	38	18	16	11
MS-270	Ecuador	South America	NA	29	18	0	12

Table S2. *continued*

Isolate	Country	Population	Host origin	Average pycnidia Taichung	Average pycnidia Bulgaria	Average pycnidia TE	Average pycnidia Volcani
BF-V1.1	UK	North Europe	BW	9	1.666666667	0	4.285714286
INRA08-FS0003.MS1	France	South Europe and Maghreb	BW	NA	10	NA	0
IPO-06008	Denmark	North Europe	BW	3	0	0	0
IPO-08004	Canada	North America	DW	0	0	0	0
IPO-09005	France	South Europe and Maghreb	BW	3	0.5	0	0
IPO-10014	Sweden	North Europe	NA	9.5	1.111111111	0.5	0
IPO-86002.MS1	Bolivia	South America	BW	6	10	0	6.875
IPO-86013.MS1	Turkey	Middle East	BW	0	0	0	0
IPO-86022.MS1	Turkey	Middle East	DW	1	0.5	2	0.714285714
IPO-89011.MS1	Netherlands	North Europe	BW	4	4	0	0
IPO-9010	Iran	Middle East	NA	0	2	0.7	0
IPO-9040	Iran	Middle East	NA	4	7.5	0	0
IPO-91003.MS1	Syria	Middle East	NA	0	1	0	0
IPO-92044	Kenya	East Africa	BW	1	0	0	0
IPO-94234	USA	North America	BW	3	0	0	0
IPO-95001	Switzerland	North Europe	NA	0	0	2	0
IPO-95062	Algeria	South Europe and Maghreb	DW	0	5	0	3.333333333
IPO-99032	France	South Europe and Maghreb	NA	2.5	0	0	0
Paskeville	USA	North America	NA	1	3.5	0	0

Table S3. Predicted functions for *Z. tritici* effector genes

Gene ID	Position	Polymorphism	PFAM id	Function
Mycgr3G100167	5:857034-857796	conserved	PF00254	FKBP-type peptidyl-prolyl cis-trans isomerase
Mycgr3G103792	3:2156483-2157362	conserved	PF00024	PAN domain
Mycgr3G107904	2:794600-795810	conserved	PF14856	Pathogen effector; putative necrosis-inducing factor
Mycgr3G109582	5:1696538-1698653	conserved	PF01328	Peroxidase, family 2
Mycgr3G109710	6:184812-185788	conserved	PF00188	Cysteine-rich secretory protein family
Mycgr3G201240	10:1543107-1543600	conserved	PF02015	Glycosyl hydrolase family 45
Mycgr3G202488	4:891874-892579	conserved	PF04560	RNA polymerase Rpb2, domain 7
Mycgr3G202660	8:1256459-1257283	conserved	PF01476	LysM domain
Mycgr3G28809	1:422439-423429	conserved	PF03928	Haem-degrading
Mycgr3G29060	2:1412565-1413332	conserved	PF11578	Protein of unknown function (DUF3237)
Mycgr3G31813	1:2087562-2087976	conserved	PF07876	Stress responsive A/B Barrel Domain
Mycgr3G39947	4:189199-189986	conserved	PF07249	Cerato-platanin
Mycgr3G40724	4:2107426-2108487	conserved	PF06766	Fungal hydrophobin
Mycgr3G44978	7:922940-924117	conserved	PF00160	Cyclophilin type peptidyl-prolyl cis-trans isomerase/CLD
Mycgr3G48129	9:166802-167681	conserved	PF06766	Fungal hydrophobin
Mycgr3G58567	5:94206-95180	conserved	PF00106	short chain dehydrogenase
Mycgr3G69392	3:670144-671342	conserved	PF00071	Ras family
Mycgr3G72138	5:1287999-1288716	conserved	PF02221	ML domain
Mycgr3G74892	8:1367543-1368605	conserved	PF01105	emp24/gp25L/p24 family/GOLD
Mycgr3G77055	11:1276921-1277777	conserved	PF12296	Hydrophobic surface binding protein A
Mycgr3G77196	12:168340-170199	conserved	PF12708	Pectate lyase superfamily protein
Mycgr3G77282	12:378015-378694	conserved	PF01083	Cutinase
Mycgr3G78786	1:3793465-3794370	conserved	PF00025	ADP-ribosylation factor family

Table S3. *continued*

Gene ID	Position	Polymorphism	PFAM id	Function
Mycgr3G96467	11:149923-150859	conserved	PF07335	Fungal chitosanase of glycosyl hydrolase group 75
Mycgr3G98257	1:1349889-1351018	conserved	PF01105	emp24/gp25L/p24 family/GOLD
Mycgr3G99331	2:3479067-3479856	conserved	PF01083	Cutinase
Mycgr3G111027	10:703511-704143	partially absent	PF16541	Alternaria alternata allergen 1
Mycgr3G23786	9:1159349-1160328	partially absent	PF00067	Cytochrome P450
Mycgr3G36978	2:219579-220555	partially absent	PF02469	Fasciclin domain
Mycgr3G38394	3:1634812-1635789	partially absent	PF00106	short chain dehydrogenase
Mycgr3G51082	13:471718-472201	partially absent	PF14087	Domain of unknown function (DUF4267)
Mycgr3G96437	10:1661128-1662066	partially absent	PF06951	Group XII secretory phospholipase A2 precursor (PLA2G12)
Mycgr3G97214	12:1323781-1324352	partially absent	PF00024	PAN domain

Chapter 6

Stress and sexual reproduction affect the dynamics of the wheat pathogen effector AvrStb6 and strobilurin resistance

Gerrit H.J. Kema[1,2§], Amir Mirzadi Gohari[1§], Lamia Aouini[1§], Hesham Gibriel[2§], Sarah Ware[1*], Frank Van den Bosch[3], Robbie Manning-Smith[3], Vasthi Alonso-Chavez[3], Joe Helps[3], Sarrah Ben M'Barek[4], Rahim Mehrabi[5], Caucasella Diaz-Trujillo[1,2], Elham Zamani[6], Henk Schouten[1], Theo van der Lee[1], Cees Waalwijk[1], Maarten de Waard[2], Pierre de Wit[2], Els Verstappen[1], Bart P.H.J. Thomma[2], Harold J.G. Meijer[1], and Michael F. Seidl[2]

1 Wageningen University and Research, Wageningen Plant Research, PO Box 16, 6700AA Wageningen, the Netherlands.
2 Wageningen University and Research, Laboratory of Phytopathology, PO box 16, 6700AA Wageningen, the Netherlands.
3 Rothamsted Research, West Common, Harpenden, Hertfordshire, AL5 2JQ, United Kingdom.
4 Biotechnology Center of Borj Cedria (CBBC), Laboratory of Molecular Plant Physiology, P.O. Box 901, 2050 Hammam-Lif, Tunisia.
5 Department of Agricultural Biotechnology, College of Agriculture, Isfahan University of Technology, Isfahan, Iran.
6 Department of Plant Pathology, Tarbiat Modares University, Tehran, Iran.

§ These authors contributed equally

Published in Nature Genetics 50:375-380 (2018)

ABSTRACT

Host resistance and fungicide treatments are cornerstones of plant-disease control. Here, we show that these treatments allow sex and modulate parenthood in the fungal wheat pathogen *Zymoseptoria tritici*. We demonstrate that the *Z. tritici*–wheat interaction complies with the gene-for-gene model by identifying the effector AvrStb6, which is recognized by the wheat resistance protein Stb6. Recognition triggers host resistance, thus implying removal of avirulent strains from pathogen populations. However, *Z. tritici* crosses on wheat show that sex occurs even with an avirulent parent, and avirulence alleles are thereby retained in subsequent populations. Crossing fungicide-sensitive and fungicide-resistant isolates under fungicide pressure results in a rapid increase in resistance-allele frequency. Isolates under selection always act as male donors, and thus disease control modulates parenthood. Modeling these observations for agricultural and natural environments reveals extended durability of host resistance and rapid emergence of fungicide resistance. Therefore, fungal sex has major implications for disease control.

INTRODUCTION

Sexual reproduction is common in nearly all branches of the eukaryotic tree of life, including microbial organisms like fungi (Baret et al. 2008; Peterson et al. 2014), and has been considered an important driver for rapid adaption to novel or changing environments (Seidl et al. 2014). Dothideomycete fungi represent the largest and most ecologically diverse group of ascomycetes with approximately 20,000 species (Goodwin et al. 2011), and most of them reproduce sexually and asexually. One of them is the plant pathogen *Zymoseptoria tritici* that causes septoria tritici blotch in wheat. At the onset of the wheat growing season, *Z. tritic* produces air-borne sexual ascospores, thereby releasing genetically diverse founding populations in commercial wheat fields (Hunter et al. 1999; Kema et al. 1996a; Shaw et al. 1989), and splash-dispersed asexual conidia that drive epidemics during the growing season (Shaw et al. 1993). Fungicides and host resistance are paramount for disease control. Until now, 21 resistance genes to septoria tritici blotch (*Stb* genes) have been identified (Table S1) and mapped, and *Stb6,* which is ubiquitous in European wheat cultivars (Chartrain et al. 2005), is the first resistance gene that was recently cloned (Saintenac et al. 2018). However, the molecular processes underlying the *Z. tritici*-wheat interaction are still relatively poorly understood (Linde et al. 2002; Mehrabi et al. 2009; Thrall et al. 2015).

MATERIALS AND METHODS

Primer development and PCR conditions

We developed a mismatch amplification mutation assay (MAMA) (Cha et al. 1992) on part of the cytochrome b (*cytb*) gene to determine azoxystrobin sensitivity or resistance among generated *Zymoseptoria tritici* ascospore progenies. Primers were designed with a mismatch on the penultimate nucleotide and the ultimate nucleotide was at position 143 of *cytb*. The primer set to specifically amplify a DNA fragment in sensitive isolates used a sense primer StrobSNP2fwd [5'-3' (404-428)] with a mismatch of T instead of G at nucleotide 427 of *cytb* and an antisense primer StrobSNP1rvs [5'-3' (1024-1043)]. The primer set to specifically amplify a DNA fragment in resistant isolates used an antisense primer StrobSNPrcF7[5'-3'(428-453)] with a mismatch of T instead of G at nucleotide 429 and a sense primer StrobSNPrcR1 [5'-3' (152-173)]. One and 0.5 µl of DNA were used for the MAMA and mating-type PCR assays, respectively.

Mating type PCR primers and thermal cycling conditions were as previously described (Waalwijk et al. 2002). Amplicons were analysed on 1.2% agarose gels using 25 µl aliquots of the PCR products. PCRs to amplify simple sequence repeats (SSRs) were in a 20 µl volume containing 20 ng DNA, 2 µl 10X PCR buffer with $MgCl^{2+}$, 2 µl each forward and reverse primers (2 µM), 0.8 µl dNTPs (5 mM), 0.2 µl Taq DNA polymerase (5U/ µl), and x µl sdd water. Thermal cycling was as follows: cycle 1; 94°C for 2 mins., cycle 2 (repeated 12x); 94°C for 30 secs. then 66°C for 30 secs. minus 1°C per cycle, then 72°C

for 30 secs., cycle 3 (repeated 27x); 94°C for 30 secs., then 53°C for 30 secs., then 72°C for 30 secs. and cycle 4; 72°C for 7 mins., followed by a cooling-off step to 10°C. Fragments were separated on a Mega-Gel Dual High-Throughput Vertical Electrophoresis Unit (CBS Scientific, Del Mar, CA, USA) with 6% non-denaturing acrylamide gels stained with ethidium bromide during the run.

To monitor biomass of isolates in crossing and infection assays, we designed specific TaqMan® probe/primer combinations for quantitative PCRs (qPCR) based on the *mat1-1* and *mat1-2* idiomorph sequences of the two reference *Z. tritici* isolates IPO323 and IPO94269, respectively (Waalwijk et al. 2002). Primers that specifically amplify DNA fragments in *mat1-1* isolates were Mmat1F3/Mmat1R3, with a FAM-fluorescent probe IP3, and primers to specifically amplify DNA fragments in *mat1-2* isolates were Mmat2F7/Mmat2R7, with a YY-fluorescent probe 2P4. Both quantitative real-time amplifications were performed in a single PCR on an Applied Biosystems 7500 Real-time PCR System (Foster City, CA, USA). Total reaction volumes were 25 µl, including 3 µl DNA, 12.5 µl Premix Ex Taq™ (2X) (TaKaRa, Shiga, Japan), 1 µl each forward and reverse primers (6 µM), 0.67 µl for each probe (5 mM), 0.5 µl ROX Reference Dye II (50x), and 8.33 µl ultraPURE™ nuclease-free water (Gibco, Paisley, Scotland). Thermal cycling was as follows: cycle 1; 50°C for 2 mins., cycle 2; 95°C for 10 mins., cycle 3 (repeated 39x); 95°C for 15 secs., then 60°C for 20 secs. Results were analysed using Sequence Detection Software version 1.2.3 (Applied Biosystems, Foster City, CA, USA). Standard curves from serial dilutions of known concentrations of pure fungal DNA of the six parental isolates plus the DNA from the reference isolates (Table S5) gave highly similar results in CT values. Therefore, serial dilutions of DNA from isolates IPO323 and IPO94269 were included in each TaqMan® PCR run to calculate the unknown concentrations of fungal DNA in inoculated wheat seedlings. The standard curves had very high R2 values (0.990-0.996) for all data points from 3 pg to 30 ng and, therefore, CT values within this range were reliable (data not shown). See Table S6 for all used probes and primers.

Generation and analyses of segregating *Zymoseptoria tritici* populations

Crossing assays

We used an *in planta* crossing protocol for all mating assays (Kema et al. 1996a). For mapping, we extended the existing *Z. tritici* mapping population IPO323/IPO94269 to 400 progeny isolates and the IPO323/IPO95052 population to 165 progeny isolates by manually collecting individual ascospores. For the EPP-biotic stress validation, we independently performed six crosses between avirulent and virulent isolates (IPO323, IPO94269, IPO95052) on five wheat cultivars (Obelisk, Shafir, Taichung 29, Inbar or Volcani 447) in multiple (>=2) biological replications. In addition, we used eight isolates in 19 crosses on nine wheat varieties (seven bread wheat and two durum wheat) and one barley accession (Tables S2 and 3) to test the occurrence of sex despite one of the parents is avirulent. For the EPP-abiotic stress validation we conducted 42 crosses between three sets of fungicide resistant and sensitive isolates on cv. Taichung 29 (Figure 3 and Figures S8 and 10; Table S5). Single sequence repeat (SSR) genotyping was

routinely used to either confirm that segregating populations resulted from the applied parental isolates (Table 1, Figures S7 and 11) or to determine the genotype of asexual fructifications that appeared in crossing assays (Figures S12 and 13). Populations were maintained at -80°C (Kema et al. 1996b) for further detailed analyses, including DArTSeq as well as MAMA, diagnostic PCRs for mating type determinations (Waalwijk et al. 2002) and the maternal/paternal contributions to sexual development, sequencing/phenotyping to determine (a)virulence in progeny and wild type strains (Tables 1 and 2; Tables S4 and 5; Figure S8) and qPCR (Figure 3, Figures S8 and 10).

Phenotyping
We prepared inoculum following published procedures (Mirzadi Gohari et al. 2015) and performed seedling assays at growth stage (GS) 11-12 (Zadoks et al. 1974) either by painting a spore suspension using a soft brush (mapping populations) or by atomizing a spore suspension onto the potted seedlings that were placed at the perimeter of a circular rotary table in an inoculation cabinet, adjusted at 15 rpm, which is equipped with interchangeable atomizers and a water cleaning device to avoid cross-contamination between isolates (all other assays). Infected plants were incubated in transparent plastic bags for 48h at 100% RH in the aforementioned greenhouse. Disease severity was assessed at 21 days post-inoculation using necrosis and pycnidial development estimated as percentage of the total primary leaf area of individual seedlings. Following these procedures, we screened 190 IPO323/IPO94269 offspring isolates, partly in three independent replicates (81 isolates) or singular tests (Figure S14) with the parental strains as controls, on cv. Shafir, carrying *Stb6*, and the susceptible control cv. Taichung 29.

Genetic mapping
Fungal genomic DNA was isolated using a standard CTAB-chloroform protocol. The parents and off-spring (N=282) of the *Z. tritici* mapping population (IPO323/IPO94269) (Goodwin et al. 2011; Wittenberg et al. 2009), were assayed of which 171 isolates showed distinct avirulence/virulence phenotypes on cv. Shafir. We used DArTSeq™, a genotyping-by-sequencing (GBS) method that combines diversity-arrays-technology (DArT) and next generation sequencing platforms (Courtois et al. 2013). In total 5,392 polymorphic DArTSeq markers *Z. tritici* isolates were obtained. Marker sequences (max 69 nt) were placed on the *Z. tritici* reference genome (Figure 1), using NCBI BLASTn (megablast) (Morgulis et al. 2008) and visualized using the GViz package (Hahne et al. 2013) (Figure 1). Multi-mapping markers were only placed on the genome at the best position if there was a considerable difference in bit-scores (difference ≥5).

For fine mapping, the 5,392 generated DArTSeq markers were sorted according to their discrimination power for avirulent/virulent isolates by calculating the squared differences of genotype frequencies, and 60 DArTSeq markers linked with avirulence were selected. These markers were sorted into a genetic linkage map, using JoinMap® 4 software with settings LOD (Log of Odds) ≥3 for grouping, and the maximum likelihood

mapping option for linkage group generation (Van Ooijen et al. 2006). Since the segregation of avirulence fitted the model of single gene inheritance (Figure S14) (Kema et al. 2000), phenotypic data were converted to an appropriate marker (*AvrStb6*) using scoring codes that are required for JoinMap, and this was integrated in the mapping procedure.

Offspring isolates with more than 10% missing genotypic values were removed from the analysis. Moreover, isolates without recombination near the (a)virulence locus, and eight showing discrepancies between the genotyping and phenotyping were not considered for analysis. To delimit the physical region harboring *AvrStb6*, we deployed a graphical mapping approach using the recombinant offspring isolates and clustered the markers that co-segregated with *AvrStb6* into bins with the marker order as estimated by JoinMap as a reference (Figure S15). The generated genetic linkage map was compared to the IPO323 reference genome sequence by aligning the DArTSeq data to determine the physical position of *AvrStb6*.

Gene annotation

Gene annotation was performed on the *Z. tritici* reference genome isolate IPO323 (Goodwin et al. 2011) using the Maker2 pipeline (Holt et al. 2011), combining *ab initio* protein-coding gene evidence from SNAP (Johnson et al. 2008), Augustus (Stanke et al. 2004), and GeneMark-HMM (Lukashin et al. 1998). Additionally, Maker2 was provided with protein alignments to 35 predicted fungal proteomes, *Z. tritici* reference gene models annotated by the Joint Genome Institute (JGI) (Goodwin et al. 2011), and transcriptome data (assembled transcripts and splice-junctions) derived from two previously published RNA-seq datasets (Rudd et al. 2015; Kellner et al. 2014). For gene annotation, RNA-seq data (single-end) were mapped to the *Z. tritici* reference genome with TopHat (version 2.0.13) (--min-intron-length 20 --max-intron-length 2000 --max-multihits 5) (Trapnell et al. 2009). *Z. tritici* transcripts were assembled using Cufflinks (Trapnell et al. 2010). Gene models predicted with Maker2 were manually evaluated and refined (Lee et al. 2013), for example by excluding protein-encoding genes <60 aa or lacking a starting methionine.

Identification of effector candidates

Gene expression, expressed as fragments per kilobase of exon per million fragments mapped (FPKM), during wheat colonization for newly predicted protein-coding genes was inferred using Cuffdiff (version 2.2.1) (Trapnell et al. 2010). Similar to previous observations (Rudd et al. 2015; Palma-Guerrero et al. 2016), the third replicate of the RNA-seq experiment of Rudd et al. (2015) behaved differently and was therefore excluded from all further analyses. Pair-wise log2-fold expression changes as well as multiple-testing corrected p-values ($P < 0.05$) were inferred for *in planta* RNA-seq samples compared to CDB (Rudd et al. 2015). N-terminal secretion signals were predicted in all proteins using SignalP (version 4.1) (Emanuelsson et al. 2007). Protein domains were predicted with InterProScan (Zdobnov et al. 2001).

Functional analyses of *AvrStb6*

Strains, media and growth conditions

Z. tritici strains IPO323 and IPO94269, which are avirulent and virulent on cv. Shafir, were used as wild type strains (WTs) and recipient strains for gene deletion and ectopic expression (Figure S16). The WTs and all deletion strains were kept at -80°C and were re-cultured on potato dextrose agar (PDA) (Sigma-Aldrich Chemie, Steinheim, Germany) at 15°C once desired for experimentation. Yeast-like spores were produced in yeast glucose broth (YGB) medium (yeast extract 10 $g.L^{-1}$, glucose 30 $g.L^{-1}$) after placement in an orbital shaker (Innova 4430, New Brunswick Scientific, Nijmegen, The Netherlands) at 15°C. For *in vitro* expression analyses in *Z. tritici* blastospores we used YGB and MM (Barratt et al. 1965) under similar conditions, whereas we adjusted the conditions in YGB to 25°C for expression in mycelium. *Escherichia coli* DH5α was used for general plasmid transformation and *Agrobacterium tumefaciens* strain AGL-1 was used for all fungal transformations.

Fungal transformation

All transformations were performed using *A. tumefaciens* mediated transformation (ATMT) as described previously (Zwiers et al. 2001; Mehrabi et al. 2006). Genomic DNA of stable transformants was extracted according to standard protocols (Sambrook et al. 2006). For ectopic complementation, the same procedure was utilized with minor modifications, including the use of 250 µg $m.L^{-1}$ geneticin for the selection of mutants.

RNA isolation and qRT-PCR

In vitro and *in planta* expression profiling of *AvrStb6* was performed using quantitative real-time PCR (qRT-PCR). For *in planta* analyses, wheat cv. Shafir was inoculated, in triplicate, with the WT isolates as described (Mehrabi et al. 2006), and leaf samples were collected at seven hours post-inoculation, and subsequently at 1, 2, 4, 8, 12, 16 and 20 dpi, followed by flash freezing and grinding in liquid nitrogen using a mortar and pestle. Total RNA was extracted either from ground leaves or fungal biomass produced in YGB using the RNeasy plant mini kit (Qiagen, MA, USA). DNA contamination was removed with the DNAfree kit (Ambion, Cambridgeshire, U.K.). First-strand cDNA was synthesized from approximately 2 µg of total RNA primed with oligo(dT) using the SuperScript III following manufacturers' instructions. One µl of the resulting cDNA was used in a 25 µl PCR reaction using a QuantiTect SYBR Green PCR Kit and run and analysed using an ABI 7500 Real-Time PCR System. The relative expression of each gene was initially normalized with the constitutively expressed *Z. tritici* beta-tubulin gene (Keon et al. 2007) and then calculated based on the comparative C (t) method described previously (Figure S2) (Schmittgen et al. 2008).

Pathogenicity assays and quantitative fungal biomass analyses

All assays were conducted as described above using wheat cvs. Shafir and Taichung 29 (Figures S5 and 17). Disease development was monitored and recorded every three days and leaves of cv. Taichung 29 were harvested at 2, 4, 8, 12, 16 and 20 dpi for

qRT-PCR expression analyses and for qPCR fugal biomass determination of all WTs and transformed *Z. tritici* strains (Waalwijk et al. 2002; Ware 2006) (Figure S17). Genomic DNA was extracted from approximately 100 mg of infected leaves using a standard phenol/chloroform DNA extraction (Sambrook et al. 2006).

Generation of gene deletion and ectopic integration constructs
To generate the *AvrStb6* deletion construct, pKOZtAvrStb6, the multisite Gateway® three-fragment vector construction kit was used, enabling the cloning of three fragments into the destination vector, which was compatible with the ATMT procedure. A 2 kb upstream and downstream sequence of *AvrStb6* was cloned in pDONR™P4-P1R and pDONR™P2R-P3. The generated constructs along with pRM250 (Mehrabi et al. 2009) containing the hygromycin phosphotransferase (*Hph*) gene as a selection marker were cloned into the destination vector, pPm43GW, via the LR reaction. In order to make the *AvrStb6* ectopic integration construct (pZtAvrStb6.com), the full ORF of *AvrStb6*, including a 1,020 bp upstream stretch as its promoter and 552 bp stretch downstream as terminator, were cloned into pDONR™P221 (Invitrogen, CA, USA) resulting in the generation of p221-ZtAvrStb6.com. The p221-ZtAvrStb6.com as well as two entry vectors pRM245 and pRM234 (Mehrabi et al. 2009), were used to clone these three fragments into the destination vector, pPm43GW, through the LR reaction.

Determining exclusive paternal parenthood
EPP-biotic stress
To determine parenthood in the conducted crosses, we analysed four crosses (Table 1; Figures S6 and 7) using four markers (*AvrStb6*, *mat*, ag-0006 and mt-SSR) and monitored fungal biomass development by qPCR (Figure 3, Figure S8).

EPP-abiotic stress
Strobilurin sensitivity was assayed in six strains (Supplementary Table 6) on potato dextrose agar (PDA) plates that were amended with kresoxim-methyl (BASF, Ludwigshafen, Germany) and trifloxystrobin (Bayer CropScience, Monheim am Rhein, Germany) and determined minimal inhibitory concentrations (MICs) of two different technical samples of the fungicides by spotting isolates on strobilurin amended PDA plates. The concentrations for kresoxim-methyl were 0.0025, 0.005, 0.01, 0.025, 0.05, 0.1, 0.25, 0.5, and 1.0 ppm, and the concentrations for trifloxystrobin were 0.00025, 0.0005, 0.001, 0.0025, 0.005, 0.01, 0.025, 0.05, 0.1 and 0.25 ppm. All isolates were spotted in triplicate in a volume of 5 µl per spot at a concentration of 4 x 10^5 spores ml^{-1}. As a positive control for growth isolates were also plated on PDA amended with the strobilurin solvent (1% methanol). Plates were placed at 18°C in the dark for 10 days, after which MIC values were assessed. A test progeny was generated by crossing *Z. tritici* isolates IPO03001 and IPO03003 and analysed it on amended PDA plates and with MAMA assays to conclude that both methods are congruous.

MIC values for the six parental isolates (Table S5) for the commercially available fungicide Amistar™ (Syngenta, Roosendaal, Netherlands) were determined, containing the active ingredient azoxystrobin, at 0.1, 1.0, and 10 ppm and then determined which concentrations of azoxystrobin to use for infection and crossing assays using an *in planta* dose response curves for the sensitive *Z. tritici* isolates using different preventive applications of azoxystrobin (250 $g.L^{-1}$ a.i. of azoxystrobin; 50% E.C.) on 10 day-old seedlings of cv. Taichung 29 that were preventatively treated (48h) using a track sprayer that was calibrated to deliver the recommended application of 1 $L.ha^{-1}$ sprayed at a rate of 250 $L.ha^{-1}$, with the following percentages of the full recommended dose: 0, 3.125, 6.25, 12.5, 25,50, 100 and 200% (which correspond with fungicide solutions of 0, 0.03125, 0.0625, 0.125, 0.25, 0.5, 1, and 2 g $azoxystrobin.L^{-1}$, respectively). We then inoculated with *Z. tritici* and percentages of leaf area covered by pycnidia were recorded at 20, 23, 26, and 29 dpi for dose response curve experiments, at 20 dpi for infection assays, and at 0, 5, 10, 15, 20, 25 and 30 dpi for qPCR biomass monitoring over time (Figure 3, Figures S9 and 10). Finally, three sets of *Z. tritici* field isolates IPO03001/IPO03003, IPO03002/IPO03005 and IPO04001/IPO04011, with equal pathogenicity, opposite mating types and contrasting sensitivity to azoxystrobin (Table 2, Figure 3, Figure S9) were used for the generation of 42 *in planta* ascospore progenies and fungal biomass development of each isolate in each crossing assay (Figure 3, Figure S10) individually and in pairwise mixtures on untreated and preventatively treated (48h, 100% azoxystrobin) seedlings of the wheat cv. Taichung 29 was monitored. Leaf samples were collected at 0, 5, 10, 15, 20, 25 and 30 dpi and were immediately frozen in liquid nitrogen before storage at -80°C until lyophilization, subsequent DNA extraction, and qPCR analyses. Two extractions were made from each sample (technical repeats), and the mean results were expressed in ng of fungal $DNA.mg^{-1}$ dry weight leaf material. A first set of 18 crosses was performed in seedlings of cv. Taichung 29 that were preventively treated (48h) with Amistar™ at 0 (control), 3.125, 6.25, 12.5, 25 and 50% of the full rate. In a second set of 24 crosses we repeated these conditions, but added two concentrations; full rate (100%), and the double rate (200%) (Table 2, Table S5). From six through 12 weeks after inoculation, material was harvested for ascospore discharge and collection (Kema et al. 1996a). Ascospores were isolated as much as possible from diverse locations within a plate or within several plates from each cross to obtain random ascospore progenies. Baseline germination frequencies on unamended WA plates for all 42 progeny sets (N=15,975) and randomly selected ascospores were determined. Germination frequencies of the 24 ascospore progenies for the second series of crosses were also determined on WA amended with 1 ppm active ingredient azoxystrobin (N=9,025), and these frequencies were expressed as percentages relative to the mean of the control germination frequencies on unamended water agar. We evaluated the percentage of resistant offspring by 2,100 independent MAMA PCRs (Table 2).

Developing the new population genetics model
The model equations (Table S7) were numerically solved in C++. Output was plotted using the graphics package Sigmaplot. The calculation showing that, independent of the parameter values (see also Figure S18), the frequency of virulence increases slower when the avirulent strain takes part in the sexual reproduction was done by hand and checked using the package Maple. We modelled the population genetic consequences of this new observation using an allele frequency model as introduced by Leonard (Leonard et al. 1969).

Data availability
All data are available and deposited in NCBI Genbank under accession number ACPE00000000 (Goodwin et al. 2011), in Gene Expression Omnibus under the accession number GSE54874 (Kellner et al. 2014), and as a BioProject with the accession number PRJEB8798 (Rudd et al. 2015).

URL section DArTSeq http://genome.jgi.doe.gov/Mycgr3/Mycgr3.home.html

RESULTS AND DISCUSSION

Gene-for-gene (GFG) interaction models have been suggested for a plethora of plant-pathogen interactions (Thrall et al. 2015), but genetic proof was only provided for a limited number of pathosystems (Brading et al. 2002; Brown et al. 2011). After more than a decade of genetic studies (Goodwin et al. 2011; Kema et al. 1996a; Wittenberg et al. 2009; Kema et al. 2002; Kema et al. 2000; Mirzadi Gohari et al. 2015; Waalwijk et al. 2002) we report the map-based cloning of the first *Z. tritici* avirulence effector AvrStb6, which triggers Stb6-mediated immunity (Saintenac et al. 2018) that underlies GFG in the *Z. tritici*-wheat interaction. We previously developed a mapping population between *Z. tritici* isolates IPO323 and IPO94269 (Goodwin et al. 2011; Wittenberg et al. 2009; Kema et al. 2002; Mirzadi Gohari et al. 2015) that we saturated here with Diversity Array Technology (DarTseq) markers (Tables S2, 3 and 4, Figure S1) and mapped a putative avirulence effector gene on the tip of chromosome 5 (Table 1, Figure 1). Public RNAseq data (Rudd et al. 2015) were used to predict a single gene candidate (four exons; Figure S2, 3 and 4), which was highly expressed *in planta,* encoding a small secreted protein (82 amino acids [aa], 12 cysteines, mature size 63 aa; Figure 1). Deletion in the avirulent strain IPO323 resulted in compatibility on cv. Shafir that carries *Stb6*, identifying the candidate as *AvrStb6*. Introducing *AvrStb6* into the compatible strain IPO94269 resulted in incompatibility on cv. Shafir, thereby demonstrating that AvrStb6 is recognized by Stb6 (Figure 1, Figure S5). Recently, *AvrStb6* was also identified in a genome-wide association study and subsequent ectopic integration in a virulent *Z. tritici* strain (Zhong et al. 2017). Analyses of the IPO323/IPO94269 mapping population and a

panel of *Z. tritici* isolates suggests that pathogenicity on cultivars carrying *Stb6* is consistent with two amino acid changes in the AvrStb6 protein (Table S4, Figure S6).

Table 1. Summary of segregation ratios for the nuclear markers for avirulence *(AvrStb6),* mating type *(mat),* and the microsatellite ag-0006, as well as the mitochondrial microsatellite marker mtSSR in four populations derived from crosses between *Zymoseptoria tritici* isolate IPO323 (adapted to bread wheat) with isolates IPO94269 (adapted to bread wheat) or IPO95052 (adapted to durum wheat) on three wheat cultivars with different modes of resistance (cv. Obelisk, generally susceptible to *Z. tritici* isolates derived from bread wheat; cv. Shafir carries *Stb6* (Chartrain et al. 2005); cv. Inbar is a tetraploid durum wheat cultivar that is resistant to virtually all *Z. tritici* from bread wheat, including IPO323 and IPO94269, but susceptible to the majority of *Z. tritici* isolates from durum wheat, including IPO95052) (Goodwin et al. 2011). Segregation ratios that differ significantly (Chi-square test, P<0.0001) from a 1:1 ratio are indicated by an asterisk.

Marker	IPO323	IPO94269	Progeny on bread wheat cv. Obelisk n = 74	χ^2, *P*<0,05	Progeny on bread wheat cv. Shafir n = 87	χ^2, *P*<0,05	IPO323	IPO95052	Progeny on bread wheat cv. Obelisk n = 99	χ^2, *P*<0,05	Progeny on durum wheat cv. Inbar n = 58	χ^2, *P*<0,05
AvrStb6	+	-	42:32	1,35	48:39	0,93	+	-	53:46	0,50	28:30	0,69
mat	*mat1-1*	*mat1-2*	41:33	1,11	47:40	0,56	*mat1-1*	*mat1-2*	55:44	1,22	31:27	0,28
ag-0006	B[a]	A[a]	36:38	0,05	52:35	3,32	B	A	43:56	1,71	28:30	0,69
mtSSR	B[b]	A[b]	19:55	17,51*[c]	0:87	87,00*	B	C[b]	99:0	99,00*	43:15	13,52*[d]

[a]A and B represent the different alleles of ag-0006 in the parental isolates; [b]A, B and C represent the different alleles of mtSSR in the parental isolates. [c]Both parents have expected equal chance on paternal and maternal parenthood, but the skewed segregation of mtSSR is likely caused by competition of both isolates during pathogenesis on the susceptible cv. Obelisk, similar with Figure 3a - see also Figure 2 - where both isolates were inoculated on cv. Taichung 29. [d]The expected segregation ratio is 0:58, but durum wheat cv. Inbar allows - atypically – pycnidia production by some bread wheat isolates, such as IPO323 (Kema et al. 1996c).

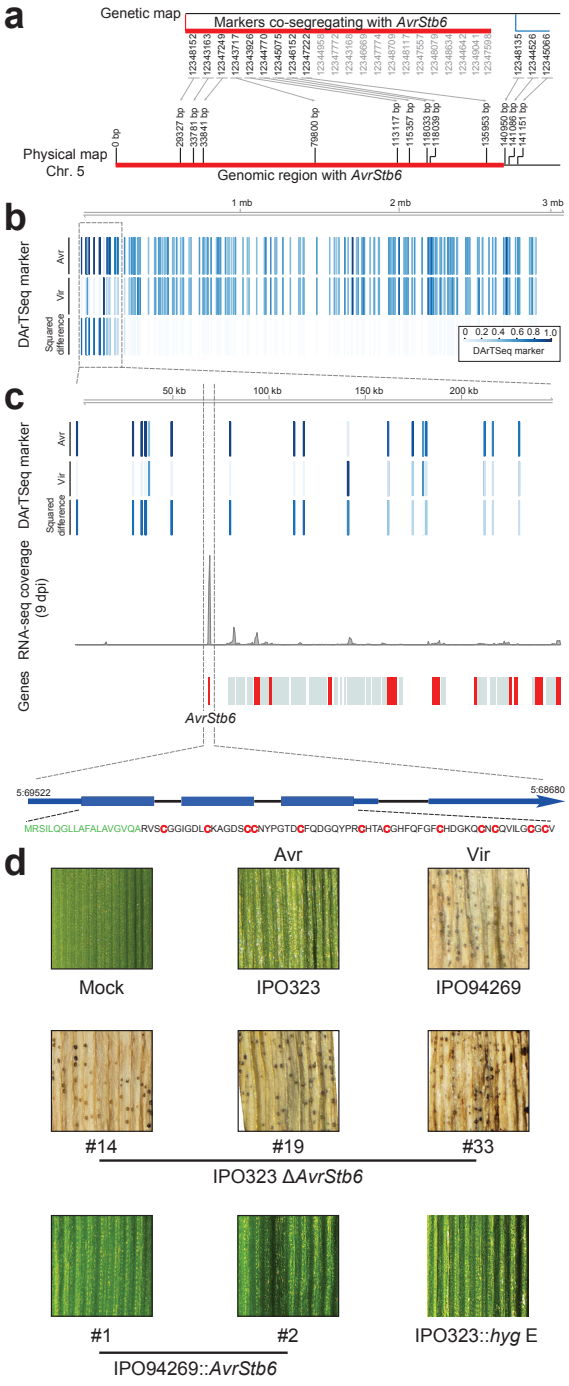


Figure 1. Cloning of *AvrStb6* in *Zymoseptoria tritici* isolate IPO323. (**a**) Genetic and physical maps of a tip of chromosome 5: cluster of 22 DArTSeq markers fully co-segregating with *AvrStb6* are highlighted in red (markers with non-unique mapping in grey) and flanking markers in blue. Marker locations (bp) are indicated by arrows and the genomic region harboring *AvrStb6* in red. (**b**) Genomic locations of DArTSeq markers on chromosome 5 are indicated by colored lines. The dashed rectangle highlights the only polymorphic region that is characterized by high squared differences of the fractions of DArTSeq markers in avirulent or virulent progeny (color-coded [scale 0-1]). (**c**) Magnification of the first 250 kb on chromosome 5 with the genomic location of the DArTSeq markers. The positions of predicted genes and genes encoding secreted proteins are indicated by grey and red bars, respectively. RNAseq reads (Rudd et al. 2015) indicate a single, highly expressed gene, designated *AvrStb6*, encoding a secreted, cysteine-rich effector protein. The blue line indicates exon-intron structure with coding regions shown with extended line width. Amino acids in green indicate the predicted signal peptide, whereas cysteines are shown in red. (**d**) Phenotyping of *Z. tritici* on cv. Sharif (carrying *Stb6*). Top: mock and wt *Z. tritici* isolates IPO323 and IPO94269. Middle: independent knock-outs of *AvrStb6* in IPO323 are virulent. Bottom: two independent Introductions of *AvrStb6* in virulent strain IPO94269 (#1 and #2) are avirulent. A strain (IPO323::*hyg* E) with an ectopic integration of the deletion construct in IPO323 is similar to the wt.

Z. tritici isolates are crossed *in planta,* which is similar to sex in nature, to generate segregating populations (Goodwin et al. 2011; Hunter et al. 1999; Kema et al. 1996a; Shaw et al. 1989; Wittenberg et al. 2009; Brading et al. 2002; Kema et al. 2002; Kema et al. 2000; Mirzadi Gohari et al. 2015). Apart from demonstrating a GFG interaction between wheat and *Z. tritici*, we observed unexpected sexual reproduction between IPO323 and IPO94269 on cv. Shafir, despite the presence of *AvrStb6* in the avirulent parent IPO323 (Table 1). Sexual reproduction was further confirmed by crossing IPO323 and IPO95052 on the cvs. Obelisk or Inbar, which are susceptible to IPO323 and resistant to IPO95052 or *vice versa,* respectively (Table 1). We analysed the four progenies with three nuclear markers (the avirulence gene *AvrStb6;* the mating type alleles *mat1-1* or *mat1*-2; a random nuclear SSR marker) and a mitochondrial SSR marker (mt-SSR) and conclude that IPO323, despite its avirulence, undergoes sexual reproduction with isolates IPO94269 or IPO95052 (Table 1, Figure S7). Thus, although IPO323 cannot infect cvs. Shafir and Inbar, it completes a sexual cycle, thereby maintaining *AvrStb6* in subsequent populations (Figure 2 and Figure S8). Moreover, crosses between sexually compatible *Z. tritici* strains never fail unless both parents are avirulent (Table S3). Notably, IPO323 is the exclusive paternal donor in the cross with the virulent isolate IPO94269 on cv. Shafir, but swaps to the exclusive maternal - and virulent - donor in crosses with the avirulent isolate IPO95052 on cv. Obelisk, as shown by the mt-SSR marker that is only maternally inherited (Rudd et al. 2015) (Table 1), as well as on cv. Taichung 29 (Figure 3a). Thus, isolate IPO323 circumvents unfavorable host conditions (i.e. resistance) via sexual reproduction as male partner; a mechanism that we here call exclusive paternal parenthood (EPP, Table 1, Figure 2). Hence, we conclude that host resistance is a biotic stress factor that modulates parenthood in fungal sex. Therefore, our data challenge the common belief (Thrall et al. 2015; Brown et al. 2011) that avirulent individuals disappear from natural populations since they can neither infect nor reproduce on resistant hosts.

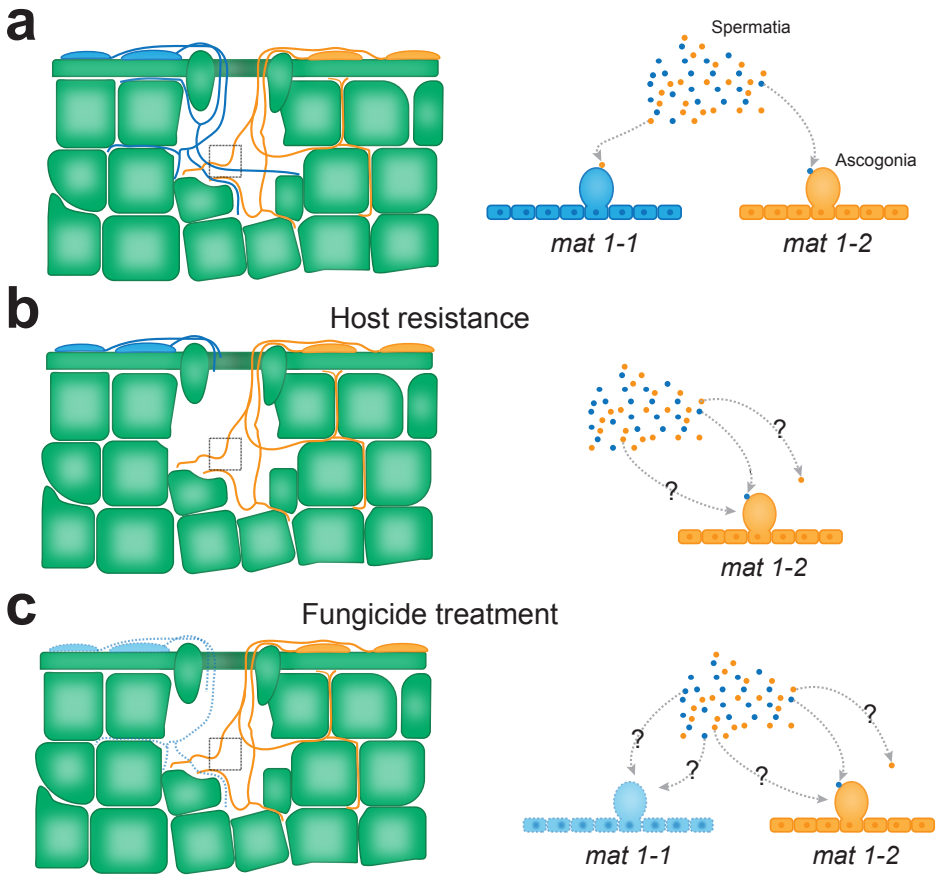


Figure 2. Sex in *Zymoseptoria tritici.* The fungus has a heterothallic bipolar mating system. Each strain has a unique mating type, either *mat1-1* (blue) or *mat1-2* (orange). When both strains infect the same host, they produce female (ascogonia) and male (microconidia or spermatia) (Annone 1984; Alexopoulos 1962) reproductive organs. Both strains have equal chances for maternal or paternal parenthood. Heterothallism defines that *mat1-1* ascogonia are exclusively fertilized by *mat1-2* spermatia and *vice versa*. (**a**) Optimal conditions for two pathogenic strains. (**b**) An avirulent strain (*mat1-1,* blue) encounters biotic stress on resistant wheat, despite penetration (Kema et al. 1996b). The virulent strain (*mat1-2,* orange) colonizes the mesophyll. Biotic stress reduces biomass of the avirulent strain, but allows the production of spermatia. Exclusive paternal parenthood (EPP) determines that ascogonia of the virulent strain are exclusively fertilized by the avirulent strain. Consequently, avirulence genes are transmitted to the progeny and distributed by airborne ascospores. (**c**) The sensitive isolate (*mat1-1,* blue dotted line) is under abiotic stress, while the resistant strain (*mat1-2,* orange solid line) colonizes the host after strobilurin application. Sensitive strains are shown during colonization or just after penetration for strobilurin applications under

field conditions are preventive/curative. Abiotic stress reduces biomass of the sensitive strain, but allows the production of spermatia. EPP determines that mating is exclusively accomplished by fertilizing the ascogonia of the resistant strain. Consequently, the entire progeny caries the *cytb* gene with the G143A mutation (fungicide resistance), which is maternally transferred and further disseminated by airborne ascospores.

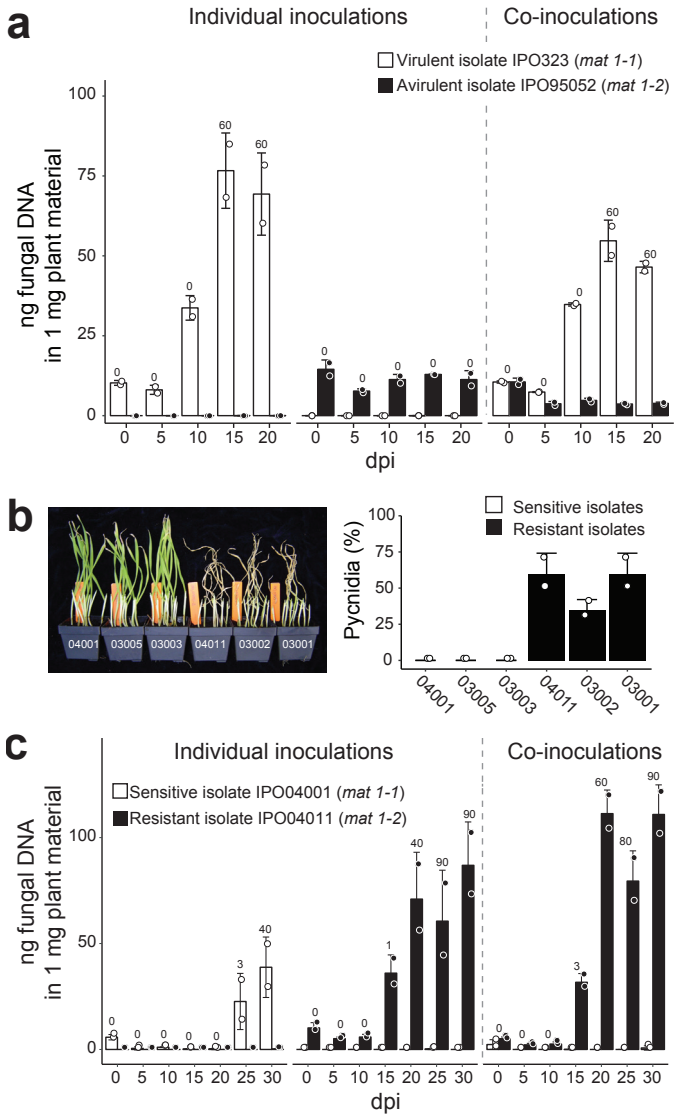


Figure 3. Inoculation and mating/competition assays with *Zymoseptoria tritici*. (a) Quantitative biomass detection of isolates IPO323 (virulent) and IPO95052 (avirulent)

and their co-inoculations on the bread wheat cv. Taichung 29 at 0, 5, 10, 15, and 20 dpi (bars: average of two independent experiments; whiskers show standard deviations). Percent leaf area covered by pycnidia at each time point shown as numbers over bars. (**b**) Fungicide sensitivity screen at 20 dpi. Plants of wheat cv. Taichung 29 treated (48h prior to inoculation) with the full recommended rate of the strobilurin Amistar® (active ingredient [ai] azoxystrobin) and inoculated with the sensitive isolates 04001, 03005 or 03003 and with the resistant isolates 04011, 03002 or 03001 (right panel; percent pycnidia based on visual observations, average of two independent experiments, whiskers show standard deviations). This resulted in significantly different disease severities between the sensitive and resistant strains (both panels). (**c**) Quantitative biomass detection of *Z. tritici* on cv. Taichung 29 after preventative treatment (48 h prior to inoculation) with the full recommended field rate of Amistar® at 0, 5, 10, 15, 20, 25, and 30 dpi (three independent crossing experiments for the phenomenon; bars are averages of two technical replicates, whiskers show standard deviations). Plants were inoculated with the sensitive isolate IPO04001, the resistant isolate IPO04011, and both (co-inoculations). Percent leaf area covered by pycnidia at each time point is shown as numbers over each bar.

To generalize these observations, we considered fungicides as abiotic stress factors for *Z. tritici* and hypothesized that they result in EPP of sensitive strains. We used the strobilurin fungicide Amistar® and resistance as the maternally inherited marker (Figure 3, Figure S9). Six *Z. tritici* field isolates originating from Germany and The Netherlands with equal pathogenicity, opposite mating types and contrasting fungicide resistance were crossed in three sets (Table 2, Table S5, Figure 3, Figure S10). We produced 42 progenies under various concentrations of Amistar® (Table 2) and the percentage of resistant ascospores was determined through either visual observation (9,025 ascospores) or by PCR assays on 2,100 progeny isolates (50 per cross) (Table 2). Despite the use of fungicides, we confirmed sexual reproduction for all crosses (Table 2, Figure S11). Sensitive strains were outcompeted in each crossing assay (Figure 3, Figure S10). Under normal and double azoxystrobin concentrations all progenies were entirel fixed for resistance in one generation (Table 2). Thus, Amistar® applications direct resistant and sensitive isolates into maternal and paternal parenthood, respectively, leading to a rapidly increasing frequency of resistance alleles in the generated progenies. In conclusion, we observed that biotic and abiotic stresses may hamper or restrict host colonization, but cannot preclude sexual reproduction as male gametes (spermatia) presumably survive (a)biotic stresses (Figure 2).

Table 2. Non-Mendelian inheritance of resistance to azoxystrobin in ascospore progeny populations of *Zymoseptoria tritici.* a, The percentage of strobilurin-resistant progeny was determined by monitoring the germination of 9,025 ascospores, originating from 21 in planta crosses between strobilurin resistant and sensitive *Z. tritici* isolates on seedlings of wheat cv. Taichung 29 that were preventively treated with six doses of Amistar® and then discharged onto water agar amended with 1 ppm (MIC value) Amistar® □normalized to germination frequencies of 15,975 ascospores from the

same crosses that were discharged to unamended water agar. **b,** The percentage of strobilurin resistant progeny determined by a strobilurin sensitivity PCR screen in 42 *Z. tritici* progenies.

a	IPO03001 [R] x IPO03003 [S]		IPO03002 [R] x IPO03005 [S]		IPO04011 [R] x IPO04001 [S]	
Rate	Total	% germination	Total	% germination	Total	% germination
0	630	93	856	15	2108	49
1/32	85	89	390	48	709	64
1/16	188	92	183	78	556	100
1/8	191	76	135	88	607	100
1/4	237	100	166	88	336	99
1/2	496	99	105	93	349	100
full dose	186	96	NA	NA	512	100
Total	2013		1835		5177	

b						
	% resistant in PCR					
Rate	1st round	2nd round	1st round	2nd round	1st round	2nd round
0	100	100	0	33	100	38
1/32	100	100	98	49	96	71
1/16	100	100	100	100	65	98
1/8	100	100	100	98	100	100
1/4	100	100	96	98	100	100
1/2	100	100	92	92	100	100
full dose	*	100	*	100	*	100
2x full dose	*	100	*	100	*	100

We developed a population genetic model by incorporating EPP into Leonard's seminal model of GFG coevolution of a plant-pathogen system (Leonard et al. 1969). In this model, a plant has one locus with alleles for resistance and susceptibility, and the pathogen has a corresponding locus with alleles for avirulence and virulence. The proportion of each allele in a well-mixed population is modelled over time. In real-life cases alleles often co-exist in stable or cyclic polymorphisms, however in Leonard's model the frequency of resistance and virulence alleles in the respective population only results in fixation of one of the genes – coexistence is not possible (Figure 4). Hence, Leonard's model forms the theoretical framework to identify traits whose inclusion can result in stable or cyclic polymorphisms (such as having multiple pathogen cycles per plant cycle, including a seed bank, or incorporating spatial structure) (Goodwin et al. 2011). We explore the consequences of incorporating the EPP reproduction mechanism into the Leonard model under two scenarios: firstly, when the frequency of the plant alleles is constant, as can be assumed in an agricultural system, and secondly, when the frequency of the plant is free to vary, as in a natural ecosystem. In the agricultural scenario, the frequency of virulence in the pathogen population increases slower when avirulent strains partake in sexual reproduction on resistant hosts (Figure 4). Additionally, polymorphisms (where the two alleles can coexist indefinitely) are possible, although unlikely, which is not the case without the EPP mechanism (Tellier et al. 2007). This implies that resistance in crop cultivars will erode slower, which can have significant consequences for the sustainability of disease control in crop production systems. In the natural scenario, our model (Figure 4) shows stable or cyclic polymorphisms occurring across a wide range of parameter values. We therefore showed that the presence of sex under biotic stress allows the occurrence of stable polymorphisms simply as a result of the pathogens' genetic system. Moreover, the model confirms that when fungicide sensitive strains partake in sex, the mitochondrially inherited *cytb* resistance allelle invades faster than any nuclear inherited fungicide resistance allelle (not shown).

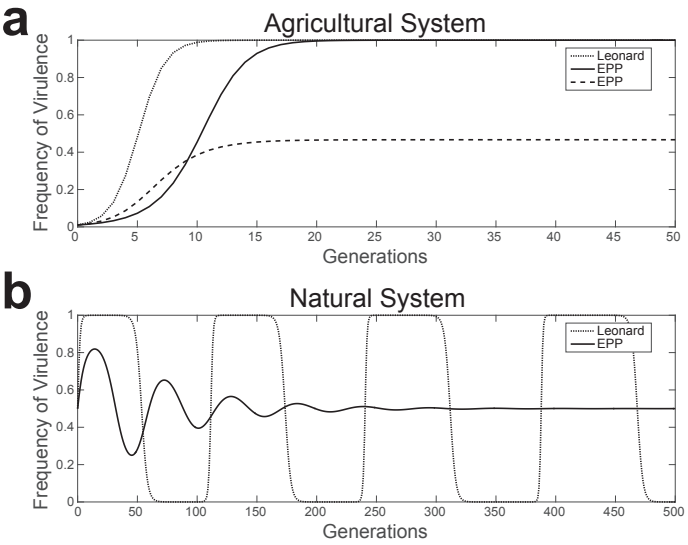


Figure 4. Comparison of the EPP model and Leonard's model in an agricultural and natural scenario. (**a**) In the agricultural scenario, the frequency of the resistance allele in the plant population is constant, as it is under the control of growers and not affected by selection pressures imposed by the pathogen. Comparing the EPP model and Leonard's model demonstrates the reduced rate of virulence build-up. The dashed EPP line represents a different set of parameters and demonstrates a polymorphism in the pathogen population. (**b**) In the natural scenario, the resistance allele frequency is dynamic and controlled by selection pressures in the system. Comparing Leonard's model with the EPP model demonstrates the possible stability of the internal equilibrium point in the EPP model, and the instability of the corresponding internal equilibrium point in Leonard's model (see Supplementary Note for different parameter sets).

The experimental data and our theoretical model provide explanations for practical observations. Slow decline of host resistance is commonly observed in the wheat-*Z. tritici* pathosystem matching the unanticipated, but ubiquitous presence of *Stb6* in many old and contemporary wheat cultivars around the world (Table S1) (Chartrain et al. 2005; Kema et al. 1997). Compared to the typical boom-and-bust cycle in the yellow rust pathogen *Puccinia striiformis*, resistance to septoria tritici blotch declined significantly slower over a period of 10 years in the United Kingdom (Brown et al. 2015). Strobilurin fungicides were commercially introduced in 1996 and showed initially excellent control of a wide range of plant pathogens including *Z. tritici*. However, in 1998 resistance appeared for powdery mildew in wheat, caused by *Blumeria graminis* (Sierotzki et al. 2000), in 2002 for *Z. tritici,* which then occurred throughout Europe one year later and presently strobilurin resistance is fixed in the vast majority of *Z. tritici* populations (Fraaije et al. 2005; Torriani et al. 2009). A similar trend for strobilurin

resistance dynamics was observed in *Pseudocercospora fijiensis*, the banana Black Sigatoka fungus (Amil et al. 2007; Arango Isaza et al. 2016).

Plant disease management mostly relies on host resistance or fungicide applications (Price et al. 2015; Diaz-Trujillo et al. 2017). Therefore, our observations on fungal sex have a broad relevance for developing resistant host varieties and shaping disease control strategies. This not only applies to plant pathogens, but also to human fungal pathogens such as *Aspergillus fumigatus,* where sex probably also contributes to the development of new life-threatening resistance mechanisms (O'Gorman et al. 2009; Verweij et al. 2016). We conclude that fungal sex is an underestimated aspect in disease control that requires much more attention.

ACKNOWLEDGEMENTS

We acknowledge financial support from the Sixth EU Framework Program Food Quality and Safety priority, the Dutch Ministry of Agriculture, Nature and Food Quality and Bayer CropScience, the Monsanto's Beachell-Borlaug International Scholars Program, the L'Oréal-UNESCO-For Women in Science Fellowship, the Consejo Nacional de Ciencia y Tecnología, Mexico, the Netherlands Organization for Scientific Research and the Dioraphte Foundation. We thank Odette Mendes, Ineke de Vries and Casper van Schaik for general support and Bruce A McDonald for sharing unpublished results.

REFERENCES

Alexopoulos, C. *Introductory Mycology*, 613 (John Wiley & Sons Inc., 1962).

Amil, A.F., Heaney, S.P., Stanger, C. & Shaw, M.W. Dynamics of QoI sensitivity in *Mycosphaerella fijiensis* in Costa Rica during 2000 to 2003. *Phytopathology* 97, 1451-1457 (2007).

Annone, J. Washington State University (1984).

Arango Isaza, R.E. *et al.* Combating a global threat to a clonal crop: banana black sigatoka pathogen *Pseudocercospora fijiensis* (synonym *Mycosphaerella fijiensis*) genomes reveal clues for disease control. *PLoS Genetics* 12, e1005876 (2016).

Barratt, R.W., Johnson, G.B. & Ogata, W.N. Wild-type and mutant stocks of *Aspergillus nidulans*. *Genetics* 52, 233-246 (1965).

Barrett, L.G., Thrall, P.H., Burdon, J.J. & Linde, C.C. Life history determines genetic structure and evolutionary potential of host–parasite interactions. *Trends in Ecology & Evolution* 23, 678-685 (2008).

Brading, P.A., Verstappen, E.C.P., Kema, G.H.J. & Brown, J.K.M. A gene-for-gene relationship between wheat and *Mycosphaerella graminicola*, the septoria tritici blotch pathogen. *Phytopathology* 92, 439-445 (2002).

Brown, J.K.M. & Tellier, A. Plant-parasite coevolution: bridging the gap between genetics and ecology. *Annual Review of Phytopathology* 49, 345-367 (2011).

Brown, J.K.M. Durable resistance of crops to disease: a darwinian perspective. *Annual Review of Phytopathology* 53, 513-539 (2015).

Cha, R.S., Zarbl, H., Keohavong, P. & Thilly, W.G. Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. *Genome Research* 2, 14-20 (1992).

Chartrain, L., Brading, P.A. & Brown, J.K.M. Presence of the *Stb6* gene for resistance to septoria tritici blotch (*Mycosphaerella graminicola*) in cultivars used in wheat-breeding programmes worldwide. *Plant Pathology* 54, 134-143 (2005b).

Courtois, B. *et al.* Genome-wde association mapping of root traits in a japonica rice panel. *PLoS One* 8, e78037 (2013).

Diaz-Trujillo, C. *et al.* A new mechanism for reduced sensitivity to demethylation-inhibitor fungicides in the fungal banana black Sigatoka pathogen *Pseudocercospora fijiensis*. *Molecular Plant Pathology*. https://doi.org/10.1111/ mpp.12637 (2017).

Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocols* 2, 953-971 (2007).

Fraaije, B.A. *et al.* Role of ascospores in further spread of QoI-resistant cytochrome *b* alleles (G143A) in field populations of *Mycosphaerella graminicola*. *Phytopathology* 95, 933-941 (2005).

Goodwin, S.B. *et al.* Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genetics* 7, e1002070 (2011).

Hahne, F. *et al.* Gviz: Plotting data and annotation information along genomic coordinates. *R package version* 1(2013).

Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12, 1-14 (2011).

Hunter, T., Coker, R.R. & Royle, D.J. The teleomorph stage, *Mycosphaerella graminicola*, in epidemics of septoria tritici blotch on winter wheat in the UK. *Plant Pathology* 48, 51-57 (1999).

Johnson, A.D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 24, 2938-2939 (2008).

Kellner, R. *et al.* Expression profiling of the wheat pathogen *Zymoseptoria tritici* reveals genomic patterns of transcription and host-specific regulatory programs. *Genome Biology and Evolution* 6, 1353-1365 (2014).

Kema, G.H.J. & van Silfhout, C.H. Genetic variation for virulence and resistance in the wheat-*Mycosphaerella graminicola* pathosystem III. Comparative seedling and adult plant experiments. *Phytopathology* 87, 266-272 (1997).

Kema, G.H.J. *et al.* A combined amplified fragment length polymorphism and randomly amplified polymorphism DNA genetic linkage map of *Mycosphaerella graminicola*, the septoria tritici leaf blotch pathogen of wheat. *Genetics* 161, 1497-1505 (2002).

Kema, G.H.J. *et al.* Genetic variation for virulence and resistance in the wheat-*Mycosphaerella graminicola* pathosystem. I. Interactions between pathogen isolates and host cultivars. *Phytopathology* 86, 200-212 (1996b).

Kema, G.H.J., Sayoud, R., Annone, J.G. & Van Silfhout, C.H. Genetic variation for virulence and resistance in the wheat-*Mycosphaerella graminicola* pathosystem. II. Analysis of interactions between pathogen isolates and host cultivars. *Phytopathology* 86, 213-220 (1996c).

Kema, G.H.J., Verstappen, E.C.P. & Waalwijk, C. Avirulence in the wheat septoria tritici leaf blotch fungus *Mycosphaerella graminicola* is controlled by a single locus. *Molecular Plant-Microbe Interactions* 13, 1375-1379 (2000).

Kema, G.H.J., Verstappen, E.C.P., Todorova, M. & Waalwijk, C. Successful crosses and molecular tetrad and progeny analyses demonstrate heterothallism in *Mycosphaerella graminicola*. *Current Genetics* 30, 251-258 (1996a).

Kema, G.H.J., Yu, D.Z., Rijkenberg, F.H.J., Shaw, M.W. & Baayen, R.P. Histology of the pathogenesis of *Mycosphaerella graminicola* in wheat. *Phytopathology* 86, 777-786 (1996).

Keon, J. *et al.* Transcriptional adaptation of *Mycosphaerella graminicola* to programmed cell death (PCD) of its susceptible wheat host. *Molecular Plant-Microbe Interactions* 20, 178-193 (2007).

Lee, E. *et al.* Web Apollo: a web-based genomic annotation editing platform. *Genome Biology* 14, R93-R93 (2013).

Leonard, K.J. Genetic equilibria in host-pathogen systems. *Phytopathology* 59, 1858–1863 (1969).

Linde, C.C., Zhan, J. & McDonald, B.A. Population structure of *Mycosphaerella graminicola*: from lesions to continents. *Phytopathology* 92, 946-955 (2002).

Lukashin, A.V. & Borodovsky, M. GeneMark.hmm: New solutions for gene finding. *Nucleic Acids Research* 26, 1107-1115 (1998).

Mehrabi, R. *et al.* Ga and Gb proteins regulate the cyclic AMP pathway that is required for development and pathogenicity of the phytopathogen *Mycosphaerella graminicola*. *Eukaryotic Cell* 8, 1001-1013 (2009).

Mehrabi, R., Zwiers, L.H., De Waard, M.A. & Kema, G.H.J. *MgHog1 r*egulates dimorphism and pathogenicity in the fungal wheat pathogen *Mycosphaerella graminicola. Molecular Plant-Microbe Interactions* 19, 1262-1269 (2006).

Mirzadi Gohari, A. *et al.* Effector discovery in the fungal wheat pathogen *Zymoseptoria tritici*. *Molecular Plant Pathology* 16, 931-945 (2015).

Morgulis, A. *et al.* Database indexing for production MegaBLAST searches. *Bioinformatics* 24, 1757-1764 (2008).

O'Gorman, C.M., Fuller, H.T. & Dyer, P.S. Discovery of a sexual cycle in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Nature* 457, 471-474 (2009).

Palma-Guerrero, J. *et al.* Comparative transcriptomic analyses of *Zymoseptoria tritici* strains show complex lifestyle transitions and intraspecific variability in transcription profiles. *Molecular Plant Pathology* 17, 845-859 (2016).

Peterson, P.D., Leonard, K.J., Roelfs, A.P. & Sutton, T.B. Effect of barberry eradication on changes in populations of *Puccinia graminis* in Minnesota. *Plant Disease* 89, 935-940 (2005).

Price, C.L., Parker, J.E., Warrilow, A.G.S., Kelly, D.E. & Kelly, S.L. Azole fungicides – understanding resistance mechanisms in agricultural fungal pathogens. *Pest Management Science* 71, 1054-1058 (2015).

Rudd, J.J. *et al.* Transcriptome and metabolite profiling of the infection cycle of *Zymoseptoria tritici* on wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle definition. *Plant Physiology* 167, 1158-1185 (2015).

Saintenac, C. *et al.* An evolutionary conserved pattern-recognition receptor like protein controls gene-for-gene resistance to a fungal pathogen in wheat. *Nature Genetics* 50, 368–374.

Sambrook, J. & Russell, D.W. Purification of nucleic acids by extraction with phenol:chloroform. *Cold Spring Harbor Protocols* 2006, pdb.prot4455 (2006).

Schmittgen, T.D. & Livak, K.J. Analyzing real-time PCR data by the comparative CT method. 3, 1101-1108 (2008).

Seidl, M.F. & Thomma, B.P.H.J. Sex or no sex: Evolutionary adaptation occurs regardless. *Bioessays* 36, 335-345 (2014).

Shaw, M.W. & Royle, D.J. Airborne inoculum as a major source of *Septoria-Tritici* (*Mycosphaerella graminicola*) infections in winter-wheat crops in the UK. *Plant Pathology* 38, 35-43 (1989).

Shaw, M.W. & Royle, D.J. Factors determining the severity of epidemics of *Mycosphaerella graminicola* (*Septoria tritici*) on winter-wheat in the UK. *Plant Pathology* 42, 882-899 (1993).

Sierotzki, H., Wullschleger, J. & Gisi, U. Point mutation in cytochrome *b* gene conferring resistance to strobilurin fungicides in *Erysiphe graminis* f. sp *tritici* field isolates. *Pesticide Biochemistry and Physiology* 68, 107-112 (2000).

Stanke, M., Steinkamp, R., Waack, S. & Morgenstern, B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research* 32, W309-W312 (2004).

Tellier, A. & Brown, J.K.M. Polymorphism in multilocus host–parasite coevolutionary interactions. *Genetics* 177, 1777-1790 (2007).

Thrall, P.H., Barrett, L.G., Dodds, P.N. & Burdon, J.J. Epidemiological and evolutionary outcomes in gene-for-gene and matching allele models. *Frontiers in Plant Science* 6, 1084 (2015).

Torriani, S.F.F., Brunner, P.C., McDonald, B.A. & Sierotzki, H. QoI resistance emerged independently at least 4 times in European populations of *Mycosphaerella graminicola*. *Pest Management Science* 65, 155-162 (2009).

Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* 28, 511-515 (2010).

Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105-1111 (2009).

Van Ooijen. JoinMap®4, Software for the calculation of genetic linkage maps in experimental populations. *Kyazma B.V., Wageningen, Netherlands* (2006).

Verweij, P.E., Chowdhary, A., Melchers, W.J.G. & Meis, J.F. Azole resistance in *Aspergillus fumigatus*: can we retain the clinical use of mold-active antifungal azoles? *Clinical Infectious Diseases* 62, 362-368 (2016).

Waalwijk, C., Mendes, O., Verstappen, E.C.P., de Waard, M.A. & Kema, G.H.J. Isolation and characterization of the mating-type idiomorphs from the wheat septoria leaf blotch fungus *Mycosphaerella graminicola*. *Fungal Genetics and Biology* 35, 277-286 (2002).

Ware, S.B. PhD thesis, Wageningen University, The Netherlands (2006).

Wittenberg, A.H.J. *et al.* Meiosis drives extraordinary genome plasticity in the haploid fungal plant pathogen *Mycosphaerella graminicola*. *PLoS One* 4, e5863 (2009).

Zadoks, J.C., Chang, T.T. & Konzak, C.F. A decimal code for the growth stages of cereals. *Weed Research* 14, 415-421 (1974).

Zdobnov, E.M. & Apweiler, R. InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17, 847-848 (2001).

Zhong, Z. *et al.* A small secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat cultivars carrying the Stb6 resistance gene. *New Phytologist* 214, 619-631 (2017).

Zwiers, L.H. & De Waard, M.A. Efficient *Agrobacterium tumefaciens*-mediated gene disruption in the phytopathogen *Mycosphaerella graminicola*. *Current Genetics* 39, 388-393 (2001).

SUPPLEMENTARY INFORMATION

Figure S1. Genome positions of the mapped DArT markers.
Genome positions for individual DArT markers that are present (blue bars) or absent in the avirulent (**A**) or virulent (**B**) progeny (282 *Z. tritici* isolates) is displayed using the core chromosomes (Chr. 1-13; left) and the dispensable chromosomes (Chr.14-21; right) of *Z. tritici* IPO323 as a reference. The values of presence (1) or absence (0) of individual DArT markers was summarized for all avirulent (**C**) or virulent (**D**) progeny. This was calculated by taking the fraction of DArT marker present in the avirulent or virulent progeny (range between 0 and 1). The squared difference between the fractions of DArT markers present in the avirulent or virulent progeny is calculated, identifying a single polymorphic region in the genome of *Z. tritici* located on the tip of chromosome 5.

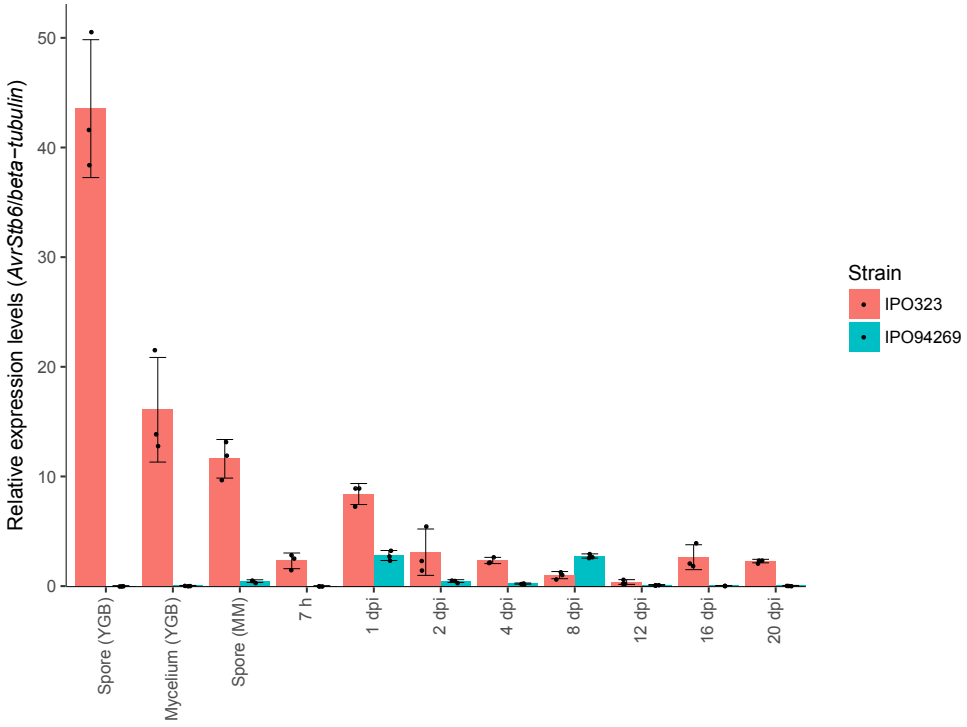


Figure S2. *In vitro* and *in planta* expression of *AvrStb6*. For *in vitro* conditions expression was profiled in blastospores in YGB and MM and in mycelium. For *in planta* expression, leaves of cv. Shafir were inoculated with the WT strains (IPO323 and IPO94269) and harvested 7 hour, 1, 2, 4, 8, 12, 16 and 20 days post-inoculation (dpi). Data were normalized with the constitutively expressed *Z. tritici* β-tubulin gene.

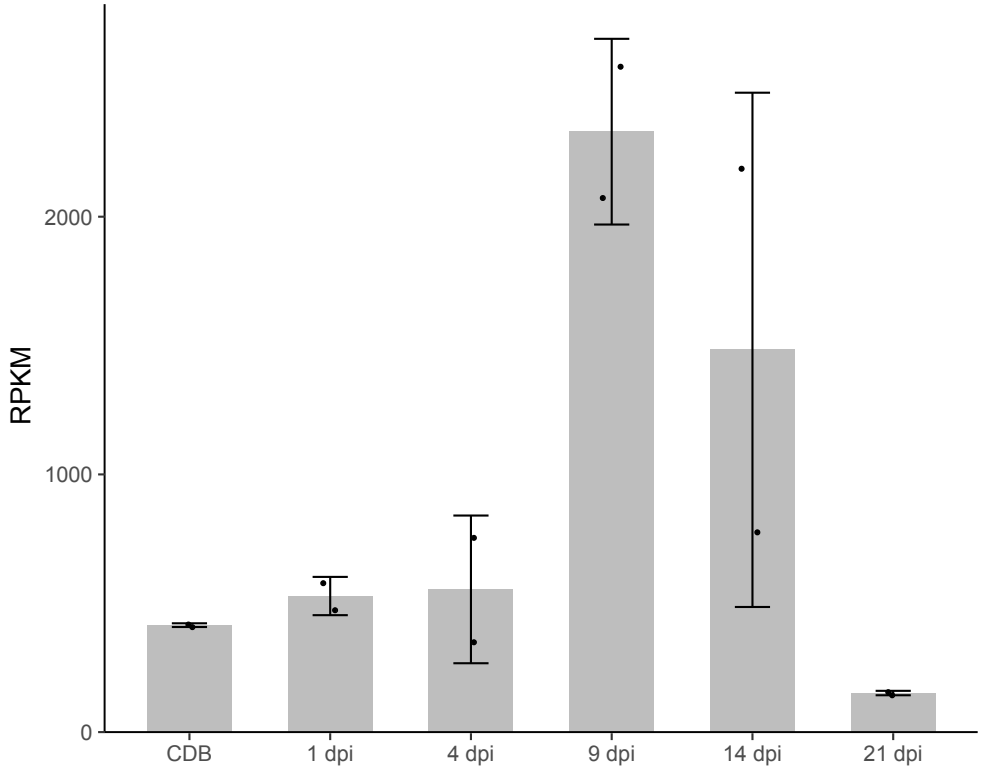


Figure S3. Expression of *AvrStb6* as measured in an RNAseq experiment using various *in vitro* and *in planta* conditions (Rudd et al. 2015).

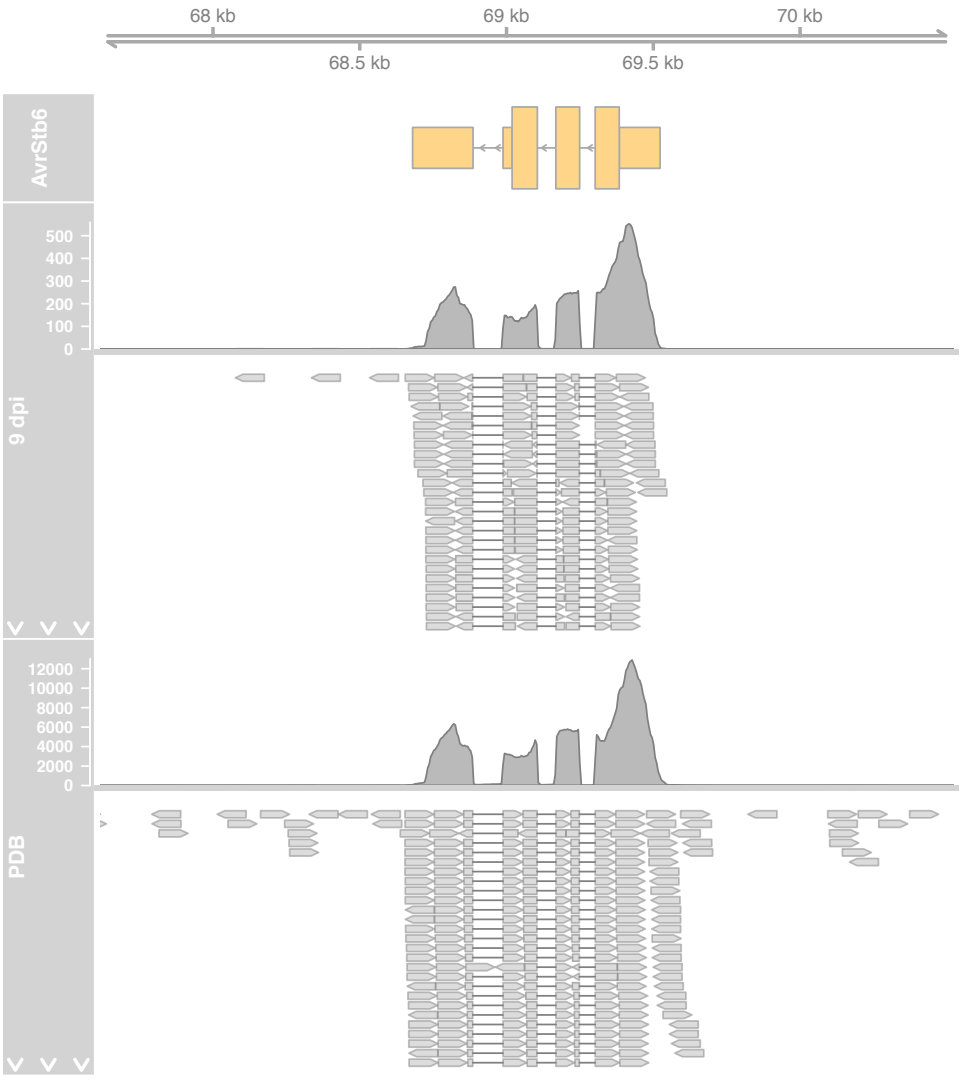


Figure S4. Genomic location of *AvrStb6* on chromosome 5. The gene model (exon-intron structure) of *AvrStb6* is displayed. Mapping of RNA sequencing reads to the reference genome of *Z. tritici* IPO323 is shown as a coverage and as a read alignment track. The transcriptomic data was derived from *in planta* (9 dpi) and *in vitro* (PDB) condition, respectively.

Figure S5. The effect of *Zymoseptoria tritici AvrStb6* deletion on disease development in the wheat cv. Shafir (*Stb6*). Primary leaves were inoculated with *Z. tritici* IPO323 and IPO94269 (WTs), all disruptants (IPO323Δ*AvrStb6*#14-19-33 and IPO94269Δ*AvrStb6*#1-2) along with the ectopic strain IPO323E*AvrStb6*#2. Experiments were triplicated and photographs were taken at 20 days post-inoculation.

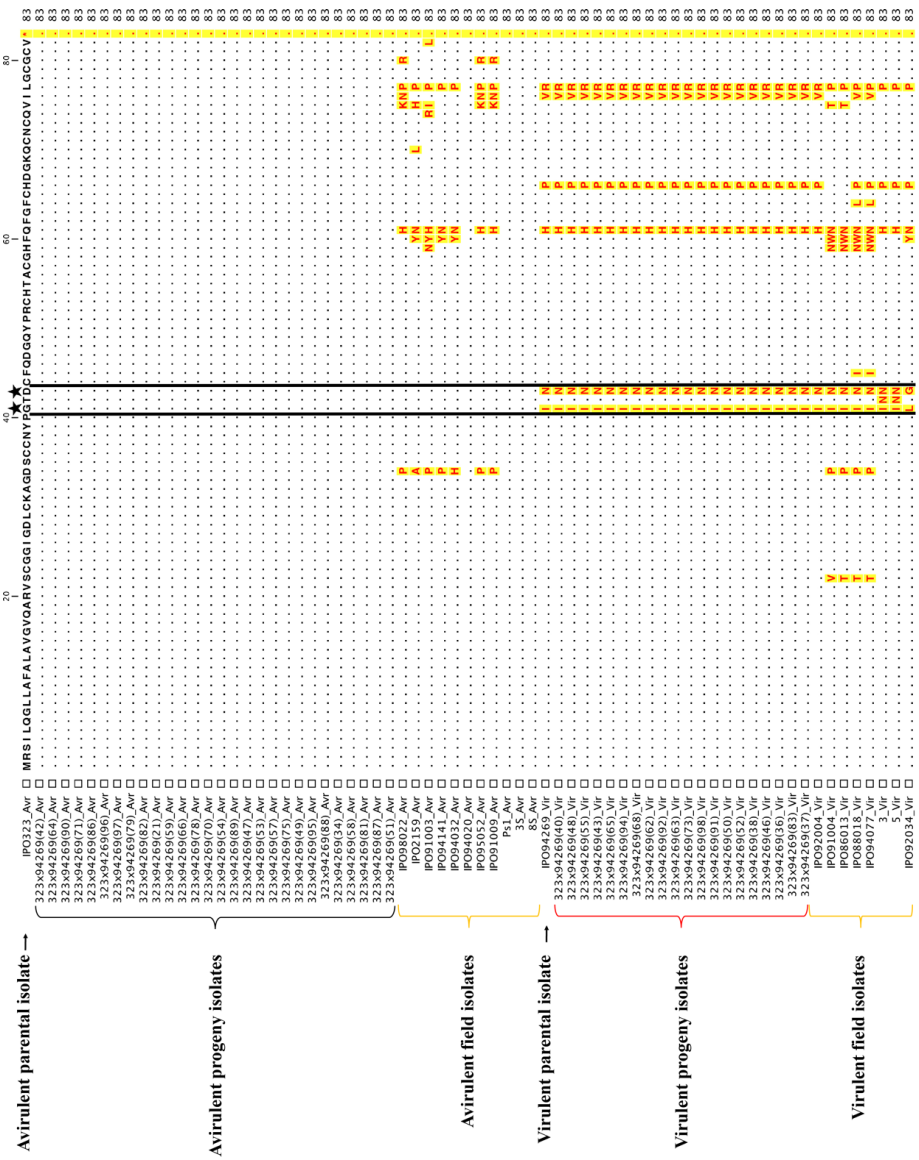


Figure S6. Protein alignment highlighting amino acid differences between virulent and avirulent *Zymoseptoria tritici* wild type isolates and IPO323/IPO94269 progeny isolates. In the alignment amino acid substitutions are indicated, identical amino acids are denoted with "-".

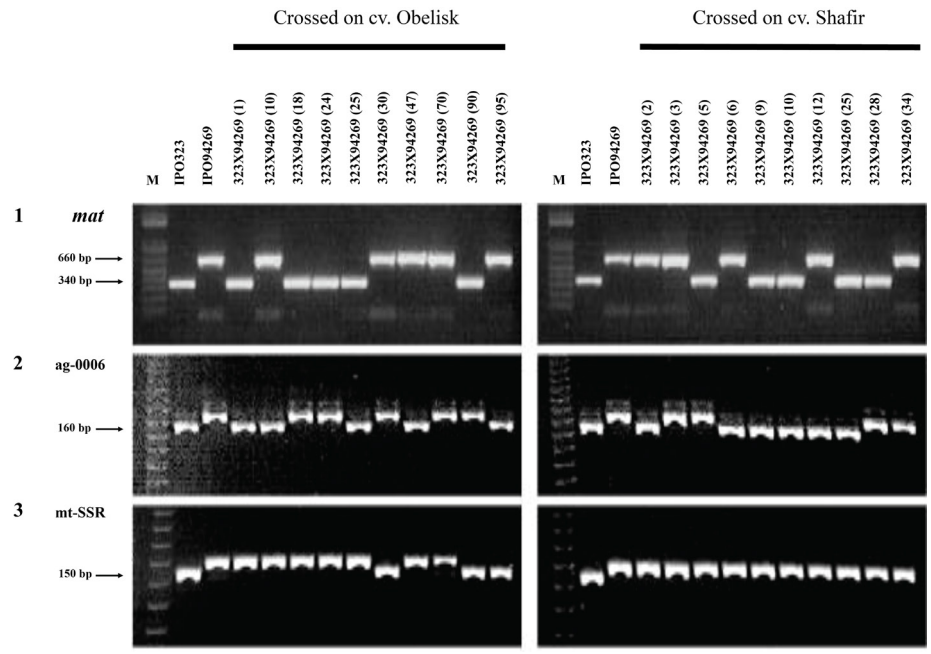


Figure S7. Examples of segregating markers in populations that were derived from *in planta Zymoseptoria tritici* crosses. *Z. tritici* isolates IPO323 and IPO94269 crossed on the bread wheat cvs. Obelisk (left) and Shafir (right). 1, *mat*. Upper band = *mat 1-1*, lower band = *mat 1-2*. 2, ag-0006. 3, mt-SSR (see also Table 1).

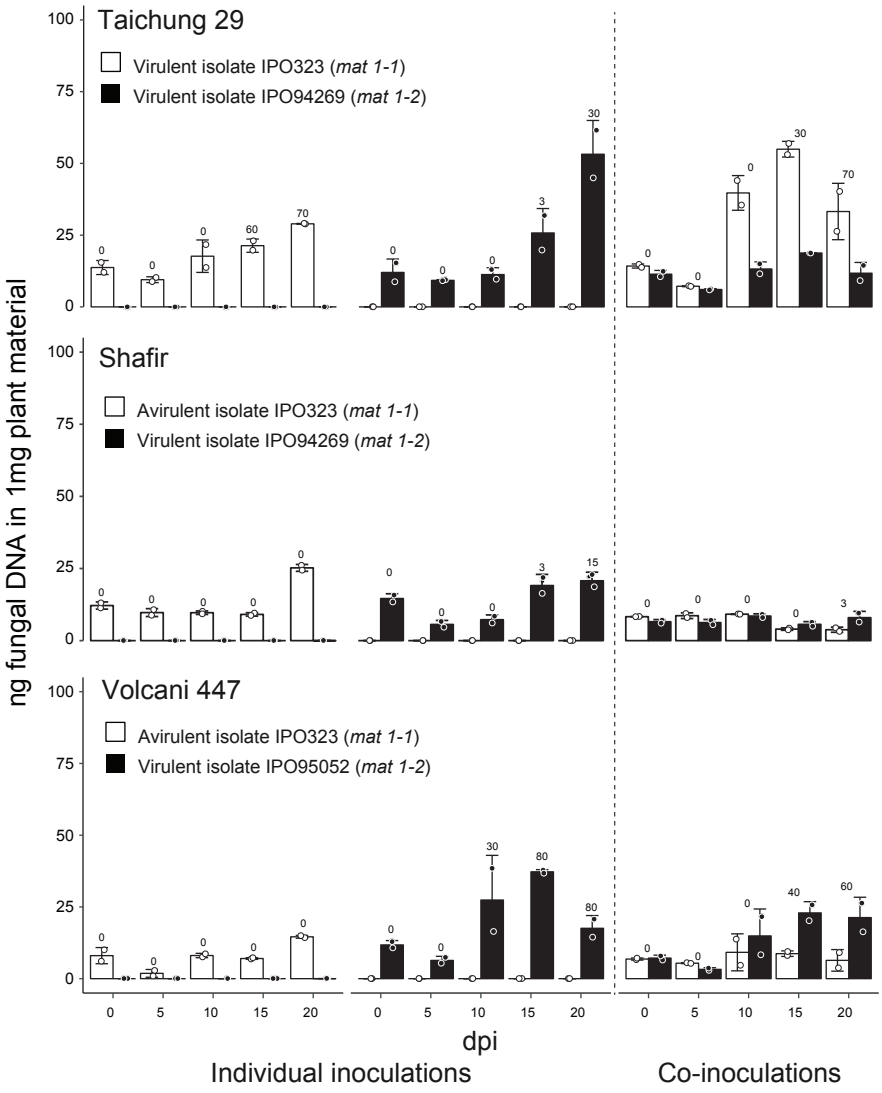


Figure S8. Quantitative fungal biomass detection of *Zymoseptoria tritici* isolates IPO3233 and IPO94269 on bread wheat cvs. Taichung29 and Shafir, and isolates IPO323 and IPO95052 on durum wheat cv. Volcani 447 at 0, 5, 10, 15, and 20 dpi (bars; average of two independent experiments; whiskers indicate standard deviations) and percent leaf area covered by pycnidia at each time point (numbers above each bar).

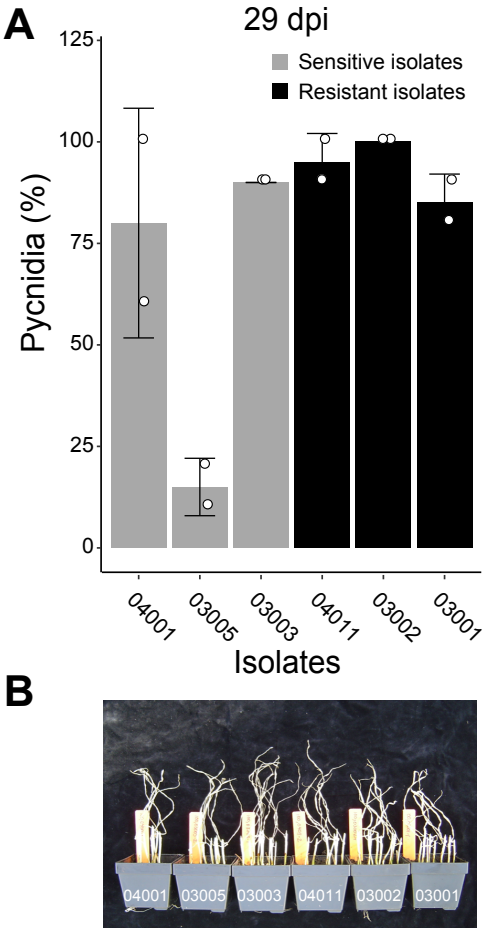


Figure S9. Responses of seedlings of wheat cv. Taichung 29 at 20 dpi (left panels) and 29 dpi (right panels) after inoculation with resistant (04001, 03005 and 03003) or sensitive (04011, 03002 and 03001) isolates of *Zymoseptoria tritici* after a pre-treatment (48h prior to inoculation) with the full recommended rate of Amistarä. A) Percent pycnidia based on visual observations. **B)** Overall view of seedlings.

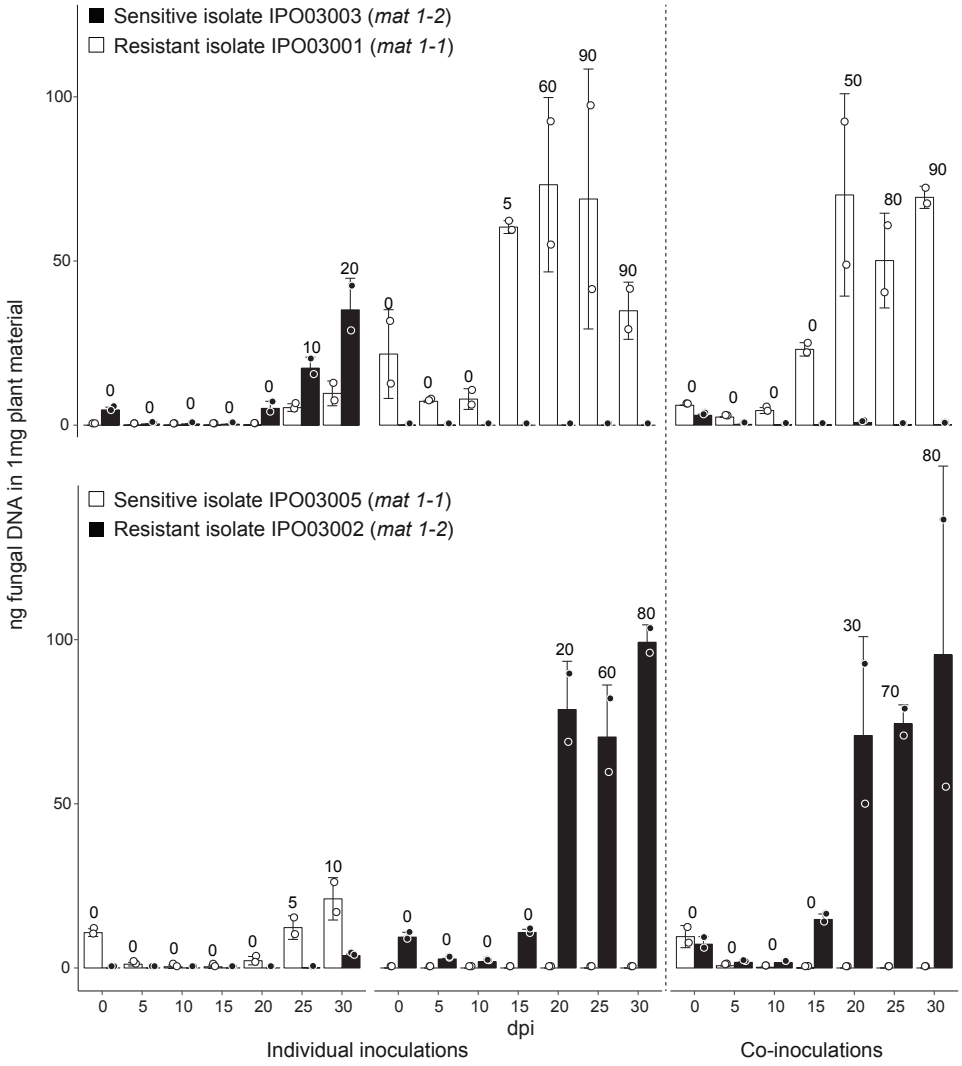


Figure S10. Fungal biomasses and percent pycnidia of parental isolates of *Zymoseptoria tritici* inoculated individually and in mixtures on cv. Taichung 29 after preventative treatment (48 h prior to inoculation) with the full recommended field rate of azoxystrobin at 0, 5, 10, 15, 20, 25, and 30 dpi. Left to right: A, IPO03003 (sensitive), IPO03001 (resistant), and mixture of IPO03001 and IPO03003. B, IPO03005 (sensitive), IPO03002 (resistant), and mixture of IPO03002 and IPO03005. Pycnidial percentages based on visual observations are shown above each bar.

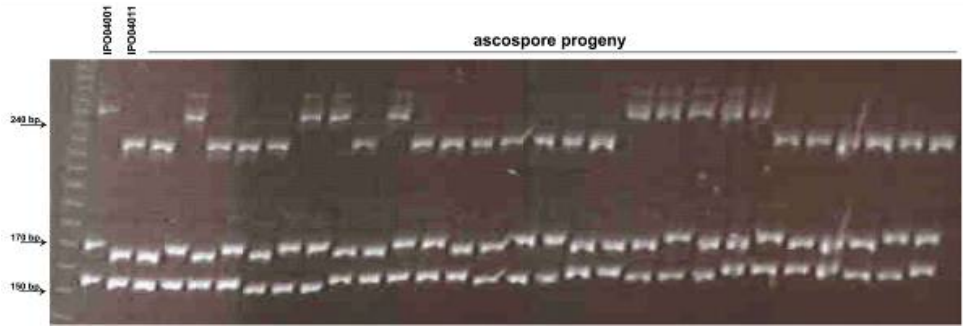


Figure S11. Example of SSR genotyping of progeny from crosses between *Zymoseptoria tritici* strains with opposite azoxystrobin phenotypes (resistant and sensitive). Isolates IPO04001 (sensitive) and IPO04011 (resistant) generated off spring on wheat seedlings preventatively treated with Amistar™ in doses ranging from 0-200%. Multi-plexed PCRs using the differentiating SSR marker primer sets ag-0003, tcc-0006 and tcc-0008 revealed recombinant SSR profiles in progeny.

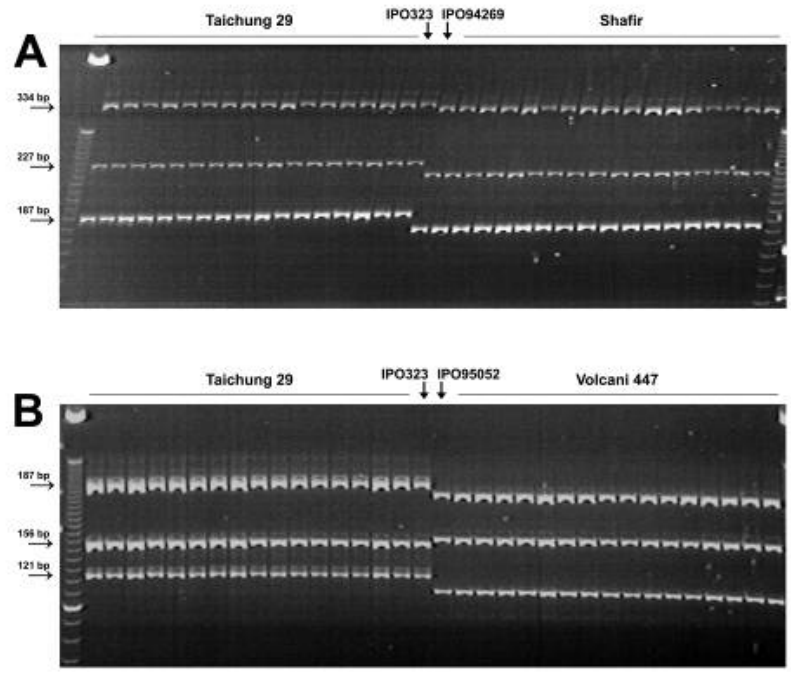


Figure S12. SSR genoyping of re-isolations of asexual pycnidial isolates from wheat leaves that were co-inoculated with *Zymoseptoria tritici* isolates. A) From mixtures of IPO323 and IPO94269 on cvs. Taichung 29 and Shafir using SSR markers (from top to bottom) ac-0001,

ggc-0001, and caa-0002. **B)** From mixtures of IPO323 and IPO95052 on cvs. Taichung 29 and Volcani 447 using SSR markers (from top to bottom) ag-0003, ag-0006, and ac-0007.

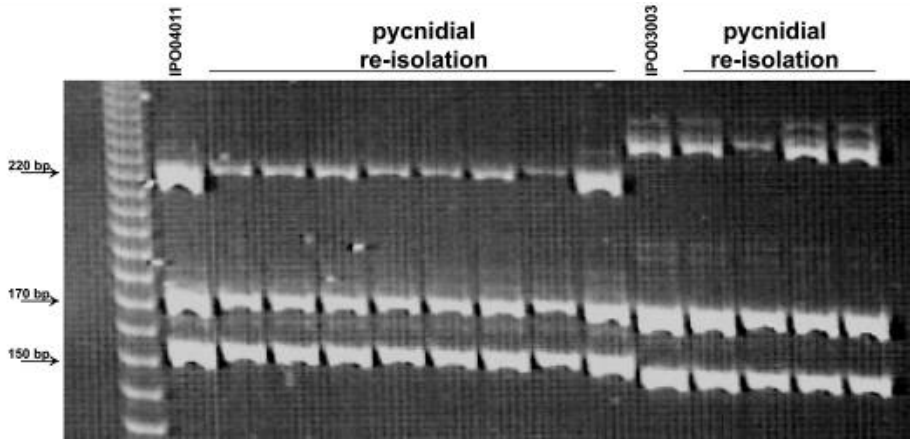


Figure S13. SSR genotyping of pycnidial isolates of *Zymoseptoria tritici* recovered from wheat seedlings preventatively treated with Amistar™ at half and full doses. All SSR patterns are clonal like the sensitive pycnidial isolates IPO04001 or IPO03003. Multi-plexed PCRs using the differentiating SSR marker primer sets ag-0003, tcc-0006 and tcc-0008 revealed no recombinant SSR patterns in recovered pycnidial isolates.

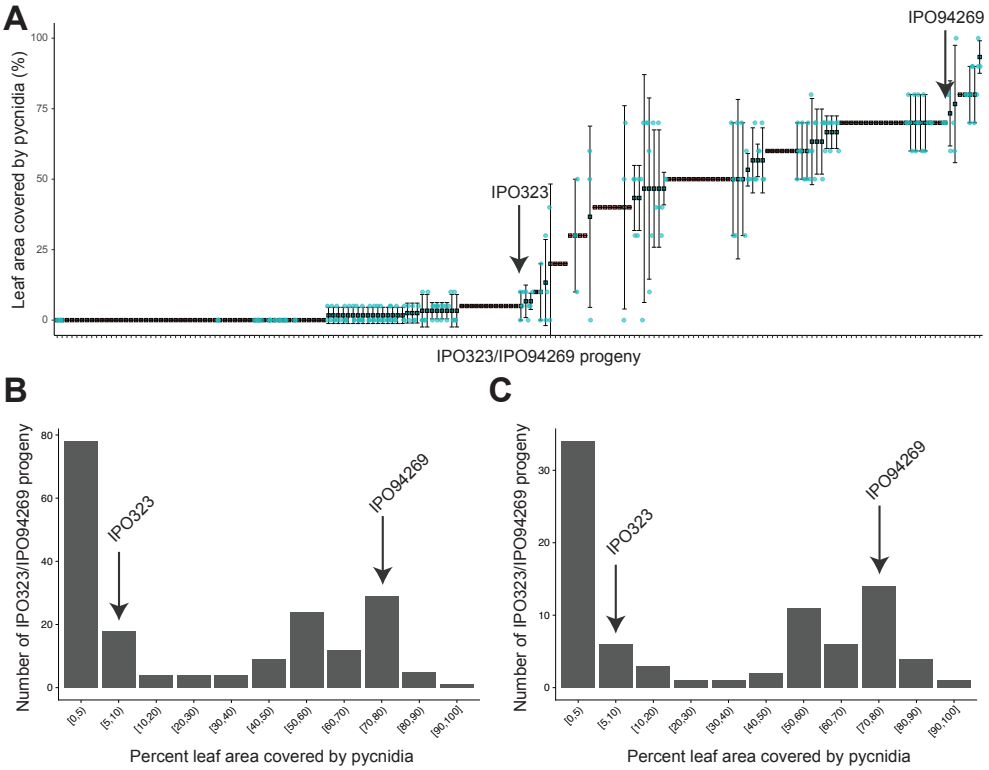
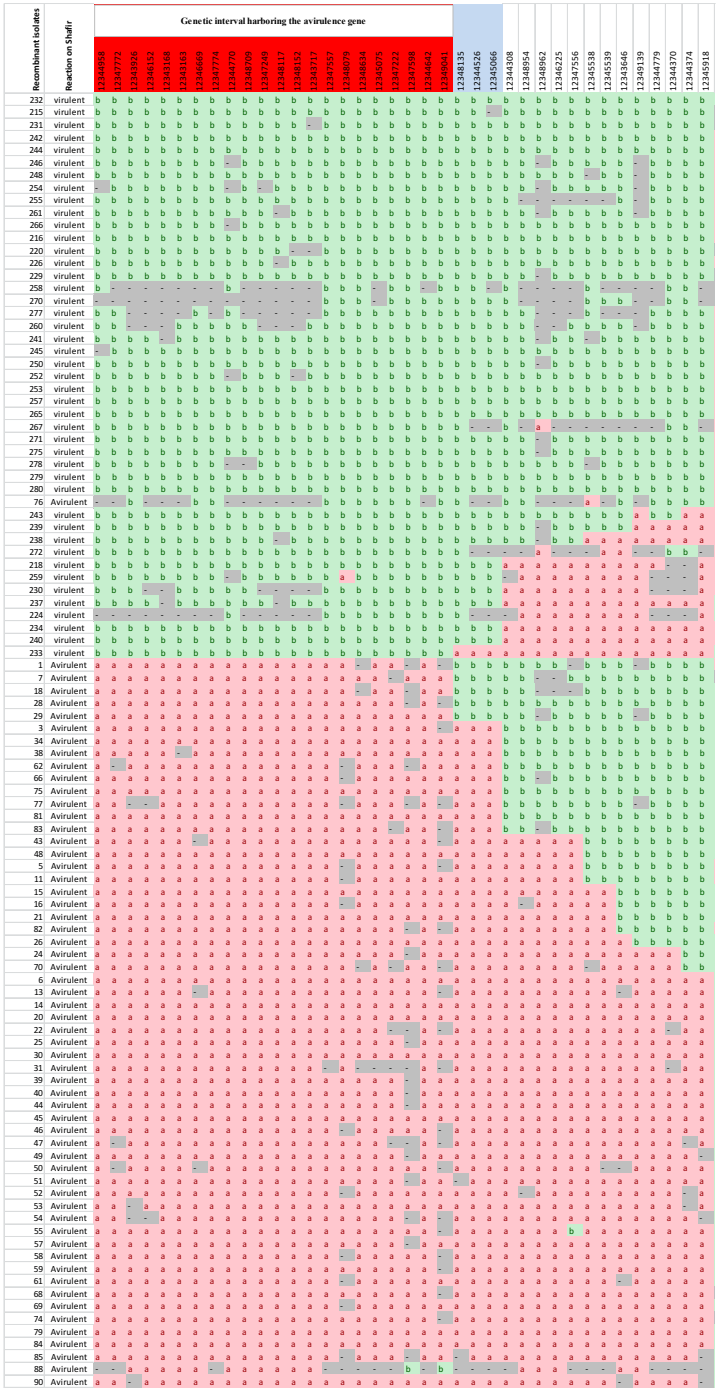


Figure S14. Segregation of the F1 IPO323/IPO94269 progeny isolates for pycnidia development on cv. Shafir carrying the *Stb6* resistance gene at 21 days post-inoculation. **(A)** Average leaf area covered by pycnidia for each IPO323/IPO94269 progeny isolates and their parental isolates, is shown. Individual isolates were ordered along the x-axis based on their average leaf area covered by pycnidia. Isolates for which more than one up to three independent experiments have been performed are highlighted in green (whiskers indicate standard deviations), while isolates used in a single experiment are shown in red. Distributions of pycnidia development (median leaf area covered) for **(B)** all isolates and **(C)** isolates with >1 independent experiment are shown. Arrows indicate average parental pycnidia development on cv. Shafir.

Figure S15. (right) Graphical genotyping of DArTseq markers and (a)virulence to 'Shafir'. By sorting the progeny isolates from the F1-population from IPO323 x IPO94260 according to the positions of recombination events on chromosome 5 of Zymoseptoria tritici, the avirulence gene AvrStb6 could be positioned between the blue flanking markers as shown on the picture. The isolates are haploid, and therefore have either the maternal locus from the avirulent parent or the paternal locus from the virulent parent. Red colour indicates loci inherited from the avirulent parent IPO323, whereas loci inherited from the virulent parent IPO94269 are displayed in green. Grey dashes indicates missing values. The yellow markers fully co-segregated with the (a)virulence. We only display the isolates that showed recombination near the (a)virulence locus.


Genetic interval harboring the avirulence gene
Recombinant isolates
Reaction on Shafir
12344958
12347722
12343926
12346152
12343168
12343163
12346669
12347774
12344770
12348709
12347249
12348117
12348152
12343717
12347557
12348079
12348634
12345075
12347222
12347598
12346442
12349841
12348135
12344526
12345066
12344308
12348954
12348962
12346225
12347556
12345538
12345539
12343646
12349139
12344779
12344370
12344374
12345918

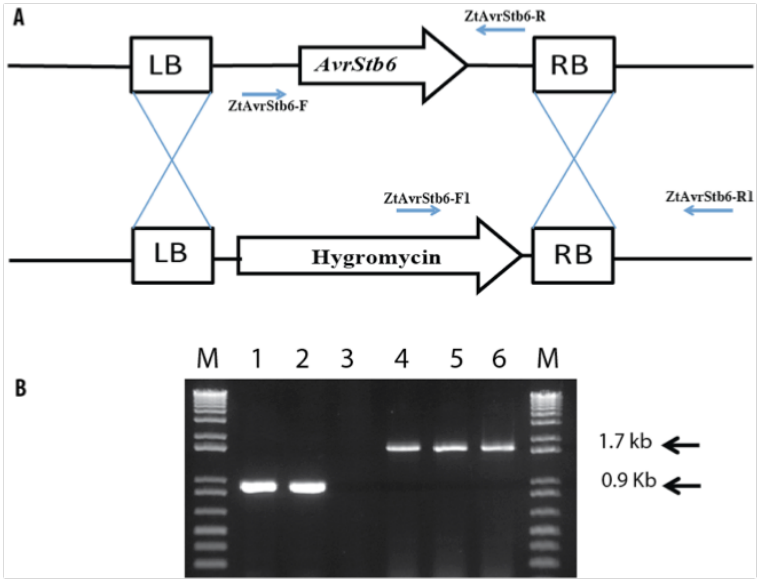


Figure S16. Replacement strategy for *AvrStb6* in *Zymoseptoria tritici*. A) Diagram showing the replacement by the hygromycin phosphotransferase (*hph*) resistance cassette through homologous recombination. The dotted line depicts the flanking regions used for homologous recombination. **B)** Identification of replacement mutants by PCR; Lane M, 1-kb-plus ladder. The gel shows the correct amplification (900 bp), using primers ZtAvrStb6-F and ZtAvrStb6-R, of *AvrStb6* in IPO323 (Lane 1) and the ectopic strain IPO323E*AvrStb6*#2 (Lane 2), but no amplification in the mutant strain IPO323Δ*AvrStb6*#33 (Lane 3). Amplification with primers ZtAvrStb6-F1 and ZtAvrStb6-R1, which are located in the middle of the *hph* gene and downstream of *AvrStb6,* respectively, result in a 1.7 kb amplicon in the three independent replacement mutants IPO323Δ*ZtStb6*#14, IPO323Δ*ZtStb6*#19 and IPO323Δ*ZtStb6*#33.

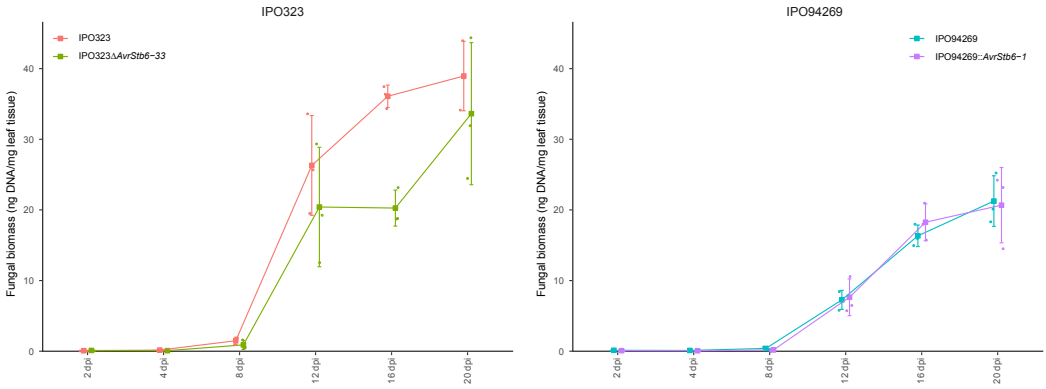


Figure S17. Fungal biomass quantifications of *Zymoseptoria tritici* in the susceptible wheat cv. Taichung 29 at 2, 4, 8, 12, 16 and 20 days post inoculation. A) Fungal biomass comparison of *Z. tritici* IPO323 (WT) and the knock-out strain IPO323Δ*AvrStb6*-33*,* **B)** Fungal biomass comparison of *Z. tritici* IPO94269 (WT) and the *AvrStb6* random integration strain IPO94269::*AvrStb6*-1.

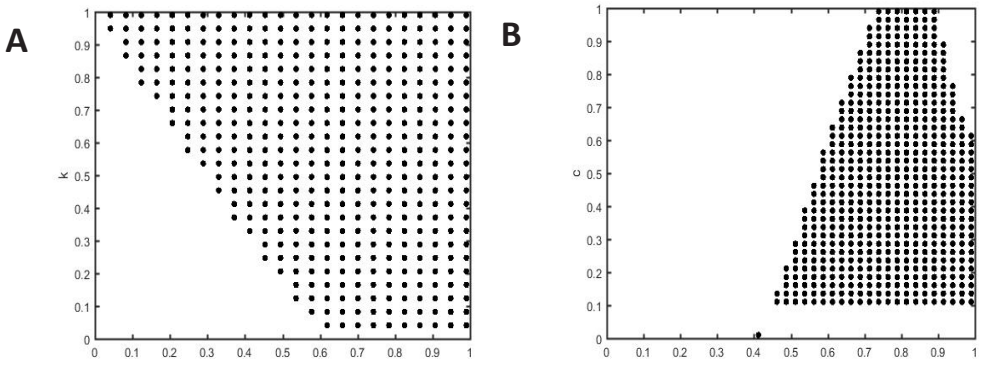


Figure S18. a. The parameter space σ vs k. The dotted area of the graph shows the pair of (σ, k) parameters where the internal equilibrium point is stable. In this plot, the fitness cost to the plant by being infected was 0.2, the fitness cost to the plant of resistance was 0.1. b. The parameter space σ vs c. The dotted area of the graph shows the pair of (σ, c) parameters where the internal equilibrium point is stable. In this plot, the fitness cost to the pathogen of being virulent was 0.2, the fitness cost to the plant of resistance was 0.1.

Table S1. Differential set of cultivars carrying mapped genes for resistance to septoria tritici blotch (*Stb*) of wheat that have been reported in winter and spring wheat cultivars.

	Stb1	*Stb2*	*Stb3*	*Stb4*	*Stb5*	*Stb6*	*Stb7*	*Stb8*	*Stb9*	*Stb10*	*Stb11*	*Stb12*	*Stb13*	*Stb14*	*Stb15*	*StbSm3*	*Stb16q*	*Stb17*	*Stb18*	*StbWW*	*TmStb1*	References
Bulgaria 88[a]	■					■																(Adhikari et al. 2004c)
Veranopolis[a]		■				■																(Liu et al. 2013)
Israel 493[a]			■			■																(Goodwin et al. 2011)
Tadinia[a]				■		■																(Adhikari et al. 004a; Somasco et al. 1996)
CS/synthetic(6x) 7D					■																	(Arraiano et al. 2001b)
Shafir						■																(Brading et al. 2002)
Estanzuela Federal							■															(McCartney et al. 2003)
M6 Synth(w7984)								■														(Adhikari et al. 2003)
Courtot									■													(Chartrain et al. 2009)
Kavkaz - K45002						■	■			■		■										(Chartrain et al. 2005a)
TE9111[b]						■	■				■											(Chartrain et al. 2005c)
Salamouni													■	■		■						(Cowling 2006; Cuthbert 2001)
Arina[a]						■									■							(Arraiano et al. 2007)
M3 (Synthetic)																	■	■				[16]
Balance						■													■			[17]
WW1842, WW2449, WW2451																				■		[18]
MDR043 (*T. monococcum*)																					■	[19]

[a]These lines also carry *Stb6* (Chartrain et al. 2005b), [b] These lines also carry *Stb6* and *Stb7*(Chartrain et al. 2005a; Chartrain et al. 2005c).

Table S2. Summary information about the *Zymoseptoria tritici* isolates that were used in the *in planta* crossing protocol on wheat cultivars with various levels of resistance.

Isolate	Year	Origin	Host	Virulent on	Mating type
IPO001	unknown	Netherlands	Bread Wheat[1]	Bread Wheat	*mat 1-1*
IPO323	1981	Netherlands	Bread Wheat	Bread Wheat	*mat 1-1*
IPO87019	1987	Uruguay	Bread Wheat	Bread Wheat	*mat 1-2*
IPO88004	1988	Ethiopia	Durum Wheat	Durum Wheat	*mat 1-2*
IPO94269	1994	Netherlands	Bread Wheat	Bread Wheat	*mat 1-2*
IPO95054	1995	Algeria	Bread Wheat	Bread Wheat	*mat 1-2*
IPO95050	1995	Algeria	Durum Wheat	Durum Wheat	*mat 1-1*
IPO95052	1995	Algeria	Durum Wheat	Durum Wheat	*mat 1-2*

[1]derived from a hexaploid derivative of a cross between bread wheat and wild emmer wheat (*T. dicoccoïdes*, AABB, 2n=28).

Table S3. Crosses between isolates of *Zymoseptoria tritici* using an *in planta* protocol of co-inoculation of two strains with opposite mating types on both bread wheat and durum wheat cultivars and on barley. V = virulent; A = avirulent.

Isolates Crossed[1]	Cultivar	Species	Progeny
bread wheat strains			
IPO323 (V) x IPO94269 (V)	Obelisk	Bread Wheat	yes
IPO323 (V) x IPO94269 (V)	Taichung 29	Bread Wheat	yes
IPO323 (A) x IPO94269 (V)	Shafir	Bread Wheat	yes
IPO323 (A) x IPO94269 (A)	Volcani 447	Durum Wheat	no
IPO323 (A) x IPO94269 (A)	Topper 33	Barley	no
IPO001 (V) x IPO94269 (V)	Obelisk	Bread Wheat	yes
IPO001 (V) x IPO94269 (A)	Lakhish	Bread Wheat	yes
IPO001 (A) x IPO94269 (V)	Clement	Bread Wheat	yes
IPO323 (A) x IPO87019 (A)	Kavkaz-K4500	Bread Wheat	no
IPO323 (A) x IPO88004 (A)	Veranopolis	Bread Wheat	no
durum wheat strains			
IPO95050 (V) x IPO95052 (V)	Volcani 447	Durum Wheat	yes
IPO95050 (A) x IPO95052 (A)	Obelisk	Bread Wheat	no

Table S3. *continued*

bread wheat x durum wheat strains			
IPO323 (A) x IPO95052 (V)	Inbar	Durum Wheat	yes
IPO323 (V) x IPO95052 (A)	Obelisk	Bread Wheat	yes
IPO323 (A) x IPO95052 (A)	Shafir	Bread Wheat	no
IPO94269 (A) x IPO95050 (V)	Inbar	Durum Wheat	yes
IPO94269 (V) x IPO95050 (A)	Obelisk	Bread Wheat	yes
IPO95054 (A) x IPO95050 (V)	Inbar	Durum Wheat	yes
IPO95054 (V) x IPO95050 (A)	Obelisk	Bread Wheat	yes

[1]Mutiple crosses, on at least two pots with wheat seedlings. The number of ascospores retrieved per cross was not counted or estimated, but is usually >1,000. We only observed significant reductions in progeny size in crosses between *Z. tritici* isolates with reduced numbers of dispensable chromosomes, but this is not addressed in the current paper.

Table S4. Wild type *Zymoseptoria tritici* isolates (shaded) and progeny isolates from crosses between IPO323 and IPO94269 on different wheat cultivars and their host specificity, phenotypes on cv. Shafir that carries *Stb6,* and the mutations in *AvrStb6.*

Isolate number	Origin	Host specificity	Avirulent/virulent on cv. Shafir (*Stb6*)	1st/2nd mutation (amino-acid)
IPO323	The Netherlands	Bread wheat	avirulent	GGG (Glycine)/GAC (Aspartic acid)
IPO95052	Algeria	Durum wheat	avirulent	GGG (..)/GAC (..)
IPO94141	The Netherlands	Bread wheat	avirulent	GGG (..)/GAC (..)
IPO94020	The Netherlands	Bread wheat	avirulent	GGG (..)/GAC (..)
IPO2159	Iran	Bread wheat	avirulent	GGG (..)/GAC (..)
IPO98022	France	Bread wheat	avirulent	GGG (..)/GAC (..)
IPO94032	The Netherlands	Bread wheat	avirulent	GGG (..)/GAC (..)
IPO91009	Tunisia	Durum wheat	Not tested	GGG (..)/GAC (..)
IPO91004	Syria	Durum wheat	Not tested	ATA (Isoleucine)/AAC (Asparagine)
IPO91003	Syria	Durum wheat	Not tested	GGG (Glycine)/GAC (Aspartic acid)
8S	Iran	Isolate from *Phalaris minor*	Not tested	GGG (..)/GAC (..)
3S	Iran	Isolate from *P. minor*	Not tested	GGG (..)/GAC (..)
Ps1	Iran	Isolate from *P. paradoxa*	Not tested	GGG (..)/GAC (..)
IPO94269	The Netherlands	Bread wheat	virulent	ATA (Isoleucine)/AAC (Asparagine)
5	USA	Bread wheat	virulent	ATC (..)/AAC (..)
3	USA	Bread wheat	virulent	ATC (..)/AAC (..)
IPO86013	Turkey	Bread wheat	virulent	ATA (..)/AAC (..)
IPO92034	Algeria	Bread wheat	virulent	TTG (Leucine)/GGC (Glycine)

Table S4. *continued*

Isolate number	Origin	Host specificity	Avirulent/virulent on cv. Shafir (*Stb6*)	1st/2nd mutation (amino-acid)
IPO94077	The Netherlands	Bread wheat	virulent	ATA (Isoleucine)/AAC (Asparagine)
IPO88018	Ethiopia	Bread wheat	virulent	ATA (..)/AAC (..)
IPO92004	Portugal	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(86T)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (Glycine)/GAC (Aspartic acid)
323x94269(96T)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(97)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(82T)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(78)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(70)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(89)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(47)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(57)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(49)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(95)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(88)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(90 T)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(34)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(58)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(81)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(87)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(51)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(21)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)

Table S4. *continued*

Isolate number	Origin	Host specificity	Avirulent/virulent on cv. Shafir (*Stb6*)	1st/2nd mutation (amino-acid)
323x94269(66)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(71T)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(75)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(54)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(53)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(64T)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(42)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(79)	Progeny from IPO323/IPO94269[1]	Bread wheat	avirulent	GGG (..)/GAC (..)
323x94269(59)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (Isoleucine)/AAC (Asparagine)
323x94269(48)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(55)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x9426(46)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(36)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(43)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(65)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(37)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(40T)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(94T)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(68)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(62)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(92)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(63)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)

Table S4. *continued*

Isolate number	Origin	Host specificity	Avirulent/virulent on cv. Shafir (*Stb6*)	1st/2nd mutation (amino-acid)
323x94269(73)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(83)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(98)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(91)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(50)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(52)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(38)	Progeny from IPO323/IPO94269[1]	Bread wheat	virulent	ATA (..)/AAC (..)
323x94269(21)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (Glycine)/GAC (Aspartic acid)
323x94269(34)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(42)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(47)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(49)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(54)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(57)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(58)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(64S)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(70)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(71S)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(75)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(78)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(79)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(81)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)

Table S4. *continued*

Isolate number	Origin	Host specificity	Avirulent/virulent on cv. Shafir (*Stb6*)	1st/2nd mutation (amino-acid)
323x94269(82S)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(84)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(86S)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(87)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(88)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(89)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(90S)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(95)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(96S)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (..)/GAC (..)
323x94269(36)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (Isoleucine)/AAC (Asparagine)
323x94269(37)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(38)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(40S)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(43)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(46)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(48)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(50)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(52)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(55)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(59)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(62)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(63)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)

Table S4. *continued*

Isolate number	Origin	Host specificity	Avirulent/virulent on cv. Shafir (*Stb6*)	1st/2nd mutation (amino-acid)
323x94269(65)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(66)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(68)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(73)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(83)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(91)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(92)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(98)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (..)/AAC (..)
323x94269(51)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ACT (Threonine)/CTT (Leucine)
323x94269(53)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (Glycine)/GAT (Aspartic acid)
323x94269(56)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	CCT (Leucine)/ACC (Threonine)
323x94269(93)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	GGG (Glycine)/GGA (Glycine)
323x94269(94S)	Progeny from IPO323/IPO94269[2]	Bread wheat	Not tested	ATA (Isoleucine)/AAA (Lysine)

[1]Progeny derived from a cross between *Z. tritici* isolates IPO323 and IPO94269 on the susceptible cvs. Taichung or Obelisk;
[2]Progeny derived from a cross between *Z. tritici* isolates IPO323 and IPO94269 on the differential cv. Shafir.

Table S5. Summary information about the *Zymoseptoria tritici* isolates used in the *in planta* crossing protocol under various levels of preventative strobilurin applications.

Pycnidial isolate	Location	Year	Strobilurin application	Strobilurin phenotype	Mating type
IPO03001 (BCS3R)[a]	Germany	2003	unknown	Resistant	*mat1-1*
IPO03002 (BCS8S)[b]	Germany	2003	unknown	Resistant	*mat1-2*
IPO03003 (BCS16S)	Germany	2003	unknown	Sensitive	*mat1-2*
IPO03005 (BCS17S)	Germany	2003	unknown	Sensitive	*mat1-1*
IPO04001	Netherlands	2004	No	Sensitive	*mat1-1*
IPO04011	Netherlands	2004	No	Resistant	*mat1-2*
IPO323 (reference)	Netherlands	1981	No	Sensitive	*mat1-1*
IPO94269 (reference)	Netherlands	1994	No	Sensitive	*mat1-2*

[a]isolate code from Bayer CropScience; [b]phenotype of isolate was mislabeled. Sequence information and additional Phenotyping confirmed its resistance to strobilurin.

Table S6. Probes and primers for *Zymoseptoria tritici* used in this study.

Name	Sequence (5' to 3')
MAT1-1 F	CCGCTTTCTGGCTTCTTCGCACTG
MAT1-1 R	TGGACACCATGGTGAGAGAACCT
MAT1-2 F	GGCGCCTCCGAAGCAACT
MAT1-2 R	GATGCGGTTCTGGACTGGAG
StrobSNP2fwd	CTTATGGTCAAATGTCTTTATGATG
StrobSNP1rvs	GGTGACTCAACGTGATAGC
StrobSNPrcF7	CAATAAGTTAGTTATAACTGTTGCGG
StrobSNPrcR1	CTATGCATTATAACCCTAGCGT
Mmat1P3	FAM- CGCAGTCTGCTTTGAATGAGAAGTTATC –Darquencher
Mmat1F3	GGCATTTCGCAGTATGTG
Mmat1R3	CTGCGCATTTCTCGTC
Mmat2P4	YY- CCTCGCAAGCCATCGGAGA -Darquencher
Mmat2F7	GCATCCGGGATACCAGTA
Mmat2R7	CTTGGTCATGCGACGTT
ag-0003 F	ACTTGGGGAGGTGTTGTGAG
ag-0003 R	ACGAATTGTTCATTCCAGCG
gca-0004 F	TAACGGTAACGGCAACAACC

Table S6. *continued*

Name	Sequence (5' to 3')
gca-0004 R	GTGTACCCTTGAATCGCAGC
tcc-0008 F	AAAAGACATGACGCCCGAC
tcc-0008 R	ACGAGGAATAATCGCGGAAC
ag-0006 F	TAACCAACACCAGGGGAATG
ag-0006 R	CATCAGTTGTCAGCGAATGG
ag-0009 F	GACTCCATTTACCTGTGGCG
ag-0009 R	TGTGAAGGACACGCAAAGAG
tcc-0006 F	ATCTGGACACCATCCACCAG
tcc-0006 R	GTAGGTGGGAGGGTTCATGC
ac-0001 F	CACCACACCGTCGTTCAAG
ac-0001 R	CGTAAGTTGGTGGAGATGGG
ggc-0001F	GATACCAAGGTGGCCAAGG
ggc-0001R	CACGTTGGGAGTGTCGAAG
caa-0002 F	TCTGCAGAGATCCCGTTACC
caa-0002 R	ATCCATCACATGACGCACAC
ac-0007 F	TGCTCGCAAGACATAAAACG
ac-0007 R	CTCTTAGCATTGGTCGGTGG
ZtAvrSt6-F	TTCCACACTTCTTTCCACAACTCC
ZtAvrSt6-R	CATGCAATGGAGGTATGTATGGG
ZtAvrStb6-F1	GTACACTTGTTTAGAGGTAATCCTTC
ZtAvrStb6-R1	GTCGTCGTCGTCGCAATTGATAA
Q- ZtAvrSt6-F	TTCCAGGACGGGCAATATC
Q- ZtAvrSt6-R	AGCCACAACCAAGAATGACC
Mt-SSR-F	CTCAGTTCAAGTCTGAGTGC
Mt-SSR-R	GACGCACGCATTTCCACTCTA

Table S7. Natural and agricultural systems for Leonard's model and EEP model.

	Leonard's model	EEP model
Natural system	$n_{i+1} = \frac{n_i(1-k)}{{q_i}^2 + n_i((1-{q_i}^2)-k)}$	$n_{i+1} = \frac{(1-q_i^2)(1-k)n_i\left[\frac{\sigma+(1-k-\sigma)n_i}{2\sigma+(1-k-2\sigma)n_i}\right]+q_i^2(1-k)n_i}{{q_i}^2+n_i((1-{q_i}^2)-k)}$
	$p_{i+1} = \frac{p_i[1-d-n_i c(1-k)]}{1-c+n_i kc+(1-{q_i}^2)[c-d-n_i c]}$	$p_{i+1} = \frac{p_i[1-d-n_i c(1-k)]}{1-c+n_i kc+(1-{q_i}^2)[c-d-n_i c]}$
Agricultural system	$n_{i+1} = \frac{n_i(1-k)}{q^2 + n_i((1-q^2)-k)}$	$n_{i+1} = \frac{(1-q^2)(1-k)n_i\left[\frac{\sigma+(1-k-\sigma)n_i}{2\sigma+(1-k-2\sigma)n_i}\right]+q^2(1-k)n_i}{q^2+n_i((1-q^2)-k)}$
	$q_{i+1} = q_i = q$	$q_{i+1} = q_i = q$

REFERENCES

Adhikari, T.B., Anderson, J.M. & Goodwin, S.B. Identification and molecular mapping of a gene in wheat conferring resistance to *Mycosphaerella graminicola*. *Phytopathology* 93, 1158-1164 (2003).

Adhikari, T.B. *et al.* Molecular mapping of the *Stb4* gene for resistance to septoria tritici blotch in wheat. *Phytopathology* 94, 1198-1206 (2004a).

Adhikari, T.B. *et al.* Molecular mapping of *Stb1*, a potentially durable gene for resistance to septoria tritici blotch in wheat. *Theoretical and Applied Genetics* 109, 944-953 (2004b).

Arraiano, L.S., Worland, A.J., Ellerbrook, C. & Brown, J.K.M. Chromosomal location of a gene for resistance to septoria tritici blotch (*Mycosphaerella graminicola*) in the hexaploid wheat 'Synthetic 6x'. *Theoretical and Applied Genetics* 103, 758-764 (2001).

Arraiano, L.S. *et al.* A gene in European wheat cultivars for resistance to an African isolate of *Mycosphaerella graminicola*. *Plant Pathology* 56, 73-78 (2007).

Brading, P.A., Verstappen, E.C.P., Kema, G.H.J. & Brown, J.K.M. A gene-for-gene relationship between wheat and *Mycosphaerella graminicola*, the septoria tritici blotch pathogen. *Phytopathology* 92, 439-445 (2002).

Chartrain, L., Sourdille, P., Bernard, M. & Brown, J.K.M. Identification and location of *Stb9*, a gene for resistance to septoria tritici blotch in wheat cultivars Courtot and Tonic. *Plant Pathology* 58, 547-555 (2009).

Chartrain, L. *et al.* Genetics of resistance to septoria tritici blotch in the Portuguese wheat breeding line TE 9111. *Theoretical and Applied Genetics* 110, 1138-1144 (2005c).

Cowling, S.G. University of Manitoba (2006).

Cuthbert, R. University of Manitoba (2011).

Goodwin, S.B. & Thompson, I. Development of isogenic lines for resistance to septoria tritici blotch in wheat. *Czech Journal of Genetics and Plant Breeding* 47, S98–S101 (2011).

Jing, H. *et al.* Phenotypic and genetic analysis of the *Triticum monococcum*: *Mycosphaerella graminicola* interaction. *The New Phytologist* 179, 1121-1132 (2008).

Liu, Y., Zhang, L., Thompson, I.A., Goodwin, S.B. & Ohm, H.W. Molecular mapping re-locates the *Stb2* gene for resistance to septoria tritici blotch derived from cultivar Veranopolis on wheat chromosome 1BS. *Euphytica* 190, 145-156 (2013).

McCartney, C.A., Brûlé-Babel, A.L., Lamari, L. & Somers, D.J. Chromosomal location of a race-specific resistance gene to *Mycosphaerella graminicola* in the spring wheat ST6. *Theoretical and Applied Genetics* 107, 1181-1186 (2003).

Raman, R. *et al.* Molecular mapping and physical location of major gene conferring seedling resistance to septoria tritici blotch in wheat. *Molecular Breeding* 24, 153-164 (2009).

Somasco, O.A., Qualset, C.O. & Gilchrist, D.G. Single-gene resistance to septoria tritici blotch in the spring wheat cultivar 'Tadinia'. *Plant Breeding* 115, 261-267 (1996).

Tabib Ghaffary, S.M. *et al.* New broad-spectrum resistance to septoria tritici blotch derived from synthetic hexaploid wheat. *Theoretical and Applied Genetics* 124, 125-142 (2012).

Tabib Ghaffary, S.M. *et al.* Genetic analysis of resistance to septoria tritici blotch in the French winter wheat cultivars Balance and Apache. *Theoretical and Applied Genetics* 123, 741-754 (2011).

Chapter 7

General discussion: Effector repertoires of plant pathogens revealed by comparative genomics

ABSTRACT

Plant pathogens cause substantial crop losses, thereby challenging agricultural efforts to meet the current and future global food security needs of the growing world population. During the infection process, pathogens secrete effector proteins that enable host plant colonization. Over the last decade, phytopathologists have sequenced the genomes of many plant pathogens and conducted genome comparisons of multiple isolates of their species of interest, or of a strain of their species with strains of related species, to investigate effector gene dynamics and evolution. Such studies have revealed that effector genes often localize within dynamic genomic regions and shown how pathogens adapt their effector gene repertoires to their hosts. Moreover, comparative genomic studies allowed the determination of effector gene repertoires of plant pathogens and the subsequent use of such effectors to screen for recognition specificities in host plant germplasm. In this final chapter of my thesis, I will discuss some of the approaches to discover effectors, how pathogens adapt their effector catalogs to their hosts, and how the discovery of effectors can be exploited to develop durable resistance breeding strategies.

INTRODUCTION

Plant diseases, caused by pathogens and pests, impact food security and are a continuous and a major challenge for agriculture worldwide (Fisher et al. 2012). To establish their infection, pathogens secrete an arsenal of effector proteins, many of which aim to modulate host physiology, including components of the host immune system (Cook et al. 2015; Jones and Dangl 2006). However, in turn, effectors may become recognized by host immune receptors that activate strong defence responses aimed at re-instalment of immunity (Cook et al. 2015; Jones and Dangl 2006). Consequently, pathogens need to mutate or purge recognized effectors, or evolve novel ones to suppress immunity and successfully colonize the plant (Cook et al. 2015; Jones and Dangl 2006).

Over the last decade, genome comparisons of multiple isolates of a particular species of interest, or of their species of interest with related species, identified effector genes and investigated effector gene dynamics and evolution (Duplessis et al. 2011; Kim et al. 2016; Soanes et al. 2008; Sperschneider et al. 2014; Stukenbrock et al. 2011). Such effector genes are often localized within highly dynamic genomic regions that facilitate effector gene evolution to enable pathogens to escape or overcome host immunity (Dong et al. 2015; Raffaele and Kamoun 2012). Moreover, effector gene repertoires determine the host range of a particular plant pathogen (Ailloud et al. 2015; Baroncelli et al. 2016; Dong et al. 2014; Liao et al. 2016; van Dam et al. 2016). Here, I discuss some of the approaches to identify effector genes and reveal how pathogens adapt their effector gene repertoires to their host plants. Moreover, I discuss how the discovery of effectors can be exploited to develop novel durable disease management strategies.

EFFECTORS OF PLANT PATHOGENS REVEALED BY COMPARATIVE GENOMICS

Effector genes are often not randomly distributed over the genome (Haas et al. 2009; Raffaele et al. 2010; Rouxel et al. 2011). Rather, effector genes are typically localized within plastic genomic regions that are enriched for repeats and that exhibit structural variations such as presence/absence polymorphisms or small- and large-scale rearrangements (Dong et al. 2015; Raffaele and Kamoun 2012). These plastic genomics regions may occur on separate chromosomes that are generally smaller than the core chromosomes and are not required for normal growth and physiology, often referred to as conditionally dispensable chromosomes (CDCs), in particular pathogen species such as in the tomato wilt pathogen *Fusarium oxysporum* f. sp. *lycopersici* (Ma et al. 2010; Vlaardingerbroek et al. 2016), the Brassicaceae blackleg pathogen *Leptosphaeria maculans* (Balesdent et al. 2013; Rouxel et al. 2011), and in pathogenic *Alternaria* species (Akagi et al. 2009; Akamatsu et al. 1999). Alternatively, plastic genomic regions are embedded within core chromosomes, as shown for the maize smut pathogen *Ustilago maydis* (Kamper et al. 2006; Schirawski et al. 2010), but also for the broad host range pathogen *Verticillium dahliae* (de Jonge et al. 2013; de Jonge et al. 2012; Faino et

al. 2016). By comparing the genomes of multiple isolates of a pathogen of interest, core effector genes that are shared by all strains (de Jonge et al. 2013; Hemetsberger et al. 2015; Marshall et al. 2011; Meile et al. 2018; Stergiopoulos et al. 2010) and lineage-specific (LS) effector genes that are only shared by a subset of strains can be identified (de Jonge et al. 2013; Schmidt et al. 2013). In this thesis, we performed comparative genomics of 21 *V. dahliae* strains and determined the core and LS effector repertoires of these strains (Figure 1A) (Chapter 3). Intriguingly, we observed that *V. dahliae* strains that are able to infect the same host plant harbour highly divergent LS effector repertoires with not even a single shared LS effector (Chapter 3). This has led us to suggest that different *V. dahliae* strains are able to infect the same host plant by utilizing different effectors, suggesting redundancy among the activities of effector repertoires between lineages (Chapter 3). Furthermore, we assessed the expression of core *V. dahliae* effector genes, as core effectors can play essential roles on a multitude of hosts (Guyon et al. 2014; Hemetsberger et al. 2015; Santhanam et al. 2013; Yin et al. 2017), and observed differential expression of *V. dahliae* core effector genes between host plants (Chapter 3). Moreover, we compared features of *V. dahliae* core and LS effector genes and observed that LS effector genes localize significantly closer to transposable elements (TEs) when compared with core effector genes (Chapter 3), which may hint towards accelerated evolution of the LS effector repertoires facilitated by TEs (Seidl and Thomma 2017). Similarly, comparative genomics of 136 *Zymoseptoria tritici* isolates revealed that LS effector genes localize significantly closer to TEs when compared with core effector genes, suggesting rapid evolution of *Z. tritici* LS effector genes (Chapter 5). However, thus far, all identified *Z. tritici* effector genes are localized within core genomic regions (Figure 1B) (McDonald et al. 2015), while effector genes localized within core and LS regions have been shown to play a role in *V. dahliae* virulence on various host plants (de Jonge et al. 2013; de Jonge et al. 2012; Kombrink et al. 2016; Santhanam et al. 2013). Thus, comparative genomics is a powerful tool that aids the identification of core and LS effector repertoires and can provide insights into the evolution of virulence of the pathogen of interest.

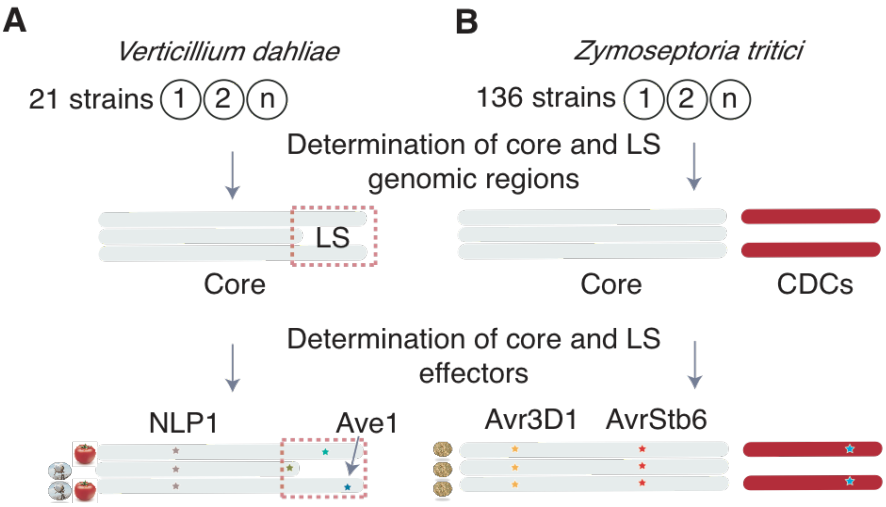


Figure 1. Effector gene repertoires revealed by comparative genomics. A) Effector genes (star) localized within the core genome (e.g. *NLP1*) or within lineage-specific (LS) genomic regions (e.g. *Ave1*) are indicated for *Verticillium dahliae* strains. Host plants, for example cotton and tomato, that the *V. dahliae* strain is capable to infect are indicated by the cotton flower and tomato symbol, respectively. **B)** Effector genes localized on core or on conditionally dispensable chromosomes (CDCs) are indicated for *Zymoseptoria tritici* strains. All thus far known effector genes, for example *AvrStb6* and *Avr3D1*, are localized on core chromosomes.

Once an effector gets recognized by a matching plant immune receptor, turning the effector into an avirulence (Avr) factor, pathogens need to eliminate or modify that effector to overcome host immunity (Cook et al. 2015; Jones and Dangl 2006). To this end, deletion of the entire effector gene (Balesdent et al. 2013; de Jonge et al. 2012; Hartmann et al. 2017; Schurch et al. 2004; Stukenbrock and McDonald 2007), or sequence modifications, even as subtle as a single nucleotide polymorphism (SNP) (Bhadauria et al. 2015; Joosten et al. 1994), may occur in the protein coding sequence of the effector gene. The deletion of an *Avr* effector gene can typically be revealed by comparative genomics between virulent and avirulent strains to identify genomic regions that are absent in virulent strains while present in avirulent strains. For example, comparative genomics between *V. dahliae* strains that are virulent and avirulent on tomato plants that carry the Ve1 immune receptor revealed a 50 kb LS region that is present in avirulent strains and absent in virulent ones, harboring the effector gene *Ave1* (de Jonge et al. 2012). This effector was subsequently experimentally shown to be the avirulence effector that is recognized by the tomato immune receptor Ve1 (de

Jonge et al. 2012). The AVRFOM2 effector of the melon wilt fungus *Fusarium oxysporum* f. sp. *melonis* that is recognized by melon Fom-2 was identified using a similar approach (Schmidt et al. 2016). However, even when recognition escape is mediated by a SNP in an *Avr* effector gene, comparative genomics between virulent and avirulent strains may be used to identify the avirulence effector gene. For example, for the tomato leaf mold fungus *Cladosporium fulvum*, comparative transcriptomics was performed on avirulent as well as virulent strains during infection of tomato plants carrying the Cf-5 immune receptor, identifying 44 *in planta*–induced effector genes (Mesarich et al. 2014). Subsequently, sequence comparisons between these effector genes revealed two that carry a SNP that is shared by virulent strains but absent in avirulent ones and that alters the protein-coding sequence (Mesarich et al. 2014). One of these effectors was termed Avr5 because follow-up experiments revealed that it is recognized by the tomato immune receptor Cf-5 (Mesarich et al. 2014). Similarly, we conducted comparative genomics between *Z. tritici* isolates that are virulent and avirulent on the wheat cultivar Shafir that carries the resistance protein Stb6 and identified a polymorphic genomic region of 250 kb (Kema et al. 2018; Chapter 6). This genomic region harbors the effector gene *AvrStb6* that occurs in nearly all *Z. tritici* isolates (except for a single isolate) and displays non-synonymous SNPs in strains that are virulent on Shafir (Kema et al. 2018; Chapter 5 and 6). Thus, sequence modifications or deletions of *Avr* effector genes are common in plant pathogens (Dodds et al. 2006; Gout et al. 2007; Iida et al. 2015; Poppe et al. 2015; Schurch et al. 2004), which highlights the importance of comparative genomics to reveal such mutated or purged effector genes in a pathogen of interest.

PATHOGENS ADAPT THEIR EFFECTOR GENE REPERTOIRES TO THEIR HOST PLANTS

The co-evolutionary arms race between plants and pathogens, in which pathogens continuously diversify their effector repertoires to avoid recognition by plant immune receptors while plants simultaneously evolve novel receptors to intercept pathogens, may lead to specialized pathogens that are capable to cause disease on a particular host (Kanzaki et al. 2012; Liao et al. 2016). However, some pathogens are still capable to cause disease on a multitude of hosts (Cook et al. 1980; Fradin and Thomma 2006; Kirk et al. 2009; Michielse and Rep 2009), yet it remains unknown how such pathogens adapt their effector gene repertoires to be competitive in arms races with diverse host plants.

Over the past few years, multiple studies have suggested that the host range of a particular plant pathogen is governed by its effector gene repertoire (Chiapello et al. 2015; Liao et al. 2016; Ma et al. 2010; van Dam et al. 2016). Thus, plant pathogens may harbour specialized effectors that exert their activities only on particular hosts. The role of specialized pathogen effectors was shown for the closely related oomycete pathogens *Phytophtora infestans*, which is pathogenic on members of the genus *Solanum*, and *P.*

mirabilis, which is pathogenic on *Mirabilis jalapa* (Dong et al. 2014). These two pathogens harbour genes that encode apoplastic EPIC effectors that exert their activities only on their respective hosts (Dong et al. 2014). Similarly, the *V. dahliae* effector protein NLP2 (necrosis-and ethylene-inducing-like protein 2) promotes virulence of *V. dahliae* strain JR2 on tomato and *Arabidopsis thaliana*, but not on *Nicotiana benthamiana* (Santhanam et al. 2013). In this thesis, we furthermore showed that the Sun1 effector contributes to aggressiveness of *V. dahliae* strain 85S only on sunflower, but not on *N. benthamiana* or *A. thaliana* (Chapter 4). Moreover, plant pathogens may harbour groups of specialized effectors that allow them to establish disease on particular hosts. For example, *formae speciales* of *F. oxysporum*, which indicates on which host plant an isolate is pathogenic, harbour groups of specialized effectors that contribute to pathogen virulence only on a particular host (Ma et al. 2010; van Dam et al. 2016). However, other plant pathogens may harbour a number of effectors that contribute to disease development on various hosts. For example, the effector protein NLP1 (necrosis-and ethylene-inducing-like protein 1) of *V. dahliae* has been shown to promote virulence on multiple host plants including tomato, *N. benthamiana*, and *A. thaliana* (Santhanam et al. 2013). Collectively, these examples illustrate that pathogens adapt their effector gene repertoires to specific or various host plants to establish disease.

EFFECTORS CAN BE USED TO SCREEN FOR RESISTANCE IN HOST PLANTS

Breeding for disease resistance can be achieved through utilizing effectors of plant pathogens (Lauge et al. 1998; Vleeshouwers and Oliver 2014). This can be achieved by transient expression of candidate effector genes in plant leaves by *Agrobacterium*-mediated expression and by a virus vector such as Potato virus X (Chapman et al. 1992; Janssen and Gardner 1990). Subsequently, plants can be screened for the occurrence of cell death–associated defence reaction, known as the hypersensitive response, which indicates effector recognition by a matching plant immune receptor (Jones and Dangl 2006). To screen for resistance, it may not be advisable to use LS effectors, as these are strain specific and could be absent in a subset of pathogen strains that may immediately break resistance. Alternatively, core effectors can be used as these may be essential for particular plant pathogens and may therefore less easily be lost. For example, the core *C. fulvum* ECP2 effector protein was used to screen for hypersensitive response among a collection of tomato genotypes, which led to the identification of the matching immune receptor Cf-ECP2 (Lauge et al. 1998). As this immune receptor recognizes a potentially important virulence factor, it may be utilized to aid for durable resistance against *C. fulvum* (Lauge et al. 1998). Thus, identification and exploitation of plant resistance proteins that recognize core effectors might potentially lead to durable pathogen resistance.

Over the past few years, many resistance genes have been identified and deployed in crop plants to limit disease development (Fu et al. 2009; Gómez-Gómez and Boller

2000; Hann and Rathjen 2007; Hurni et al. 2013; Saintenac et al. 2018; Zhang et al. 2017). However, resistance is often broken as pathogens evolve mechanisms to eliminate or modify particular effectors to overcome resistance. One mechanism to diversify effector repertoires is through sexual reproduction (Seidl and Thomma 2014), where the recombination of effector gene repertoires between two parents may lead to novel allele combinations allowing pathogens to evade the host immunity. In *Z. tritici*, sexual reproduction can even occur with an avirulent parent carrying the avirulence effector gene *AvrStb6* on the resistant host that carries *Stb6*, retaining *AvrStb6* in subsequent populations (Chapter 6). However, for various plant pathogens sex has never been observed, yet these pathogens are able to overcome host resistance. For example, the asexual pathogen *C. fulvum* overcomes resistance of tomato plant carrying the Cf-9 immune receptor through deletion of the avirulence effector gene *Avr9* (van Kan et al. 1991). Similarly, deletion of the avirulence effector gene *Ave1* allows strains of the asexual pathogen *V. dahliae* to overcome resistance of tomato plants carrying the Ve1 immune receptor (de Jonge et al. 2012). Thus, deletion of avirulence effector genes is one of the mechanisms that allows sexual as well as asexual plant pathogens to evade host immunity. Over the years, plant breeders have learned that the deployment of single resistance genes may lead to rapid degeneration of resistance due to the rapid emergence of resistance-breaking pathogen strains. Thus, novel resistance sources should be deployed wisely in order to use them in a sustainable manner. One of the strategies to deploy resistance in a sustainable manner relies on stacking multiple *R* genes directed against the same pathogen strain in a single variety, thus increasing the difficulty for a pathogen to overcome resistance (Michelmore et al. 2013). Thus, efforts should be made to simultaneously identify multiple resistance sources against the same pathogen that can be combined into the crop of interest.

REFERENCES

Ailloud, F., Lowe, T., Cellier, G., Roche, D., Allen, C., and Prior, P. 2015. Comparative genomic analysis of *Ralstonia solanacearum* reveals candidate genes for host specificity. BMC Genomics 16:270.

Akagi, Y., Akamatsu, H., Otani, H., and Kodama, M. 2009. Horizontal chromosome transfer, a mechanism for the evolution and differentiation of a plant-pathogenic fungus. Eukaryotic Cell 8:1732-1738.

Akamatsu, H., Taga, M., Kodama, M., Johnson, R., Otani, H., and Kohmoto, K. 1999. Molecular karyotypes for *Alternaria* plant pathogens known to produce host-specific toxins. Current Genetics 35:647-656.

Balesdent, M.-H., Fudal, I., Ollivier, B., Bally, P., Grandaubert, J., Eber, F., Chèvre, A.-M., Leflon, M., and Rouxel, T. 2013. The dispensable chromosome of *Leptosphaeria maculans* shelters an effector gene conferring avirulence towards *Brassica rapa*. New Phytologist 198:887-898.

Baroncelli, R., Amby, D. B., Zapparata, A., Sarrocco, S., Vannacci, G., Le Floch, G., Harrison, R. J., Holub, E., Sukno, S. A., Sreenivasaprasad, S., and Thon, M. R. 2016. Gene family expansions and contractions are associated with host range in plant pathogens of the genus *Colletotrichum*. BMC Genomics 17:555.

Bhadauria, V., Maclachlan, R., Pozniak, C., and Banniza, S. 2015. Candidate effectors contribute to race differentiation and virulence of the lentil anthracnose pathogen *Colletotrichum lentis*. BMC Genomics 16:628.

Chapman, S., Kavanagh, T., and Baulcombe, D. 1992. Potato virus X as a vector for gene expression in plants. Plant Journal 2:549-557.

Chiapello, H., Mallet, L., Guerin, C., Aguileta, G., Amselem, J., Kroj, T., Ortega-Abboud, E., Lebrun, M. H., Henrissat, B., Gendrault, A., Rodolphe, F., Tharreau, D., and Fournier, E. 2015. Deciphering genome content and evolutionary relationships of isolates from the fungus *Magnaporthe oryzae* attacking different host plants. Genome Biology and Evolution 7:2896-2912.

Cook, D. E., Mesarich, C. H., and Thomma, B. P. H. J. 2015. Understanding plant immunity as a surveillance system to detect invasion. Annual Review of Phytopathology 53:541-563.

Cook, R., Sitton, J., and Waldher, J. 1980. Evidence for *Pythium* as a pathogen of direct-drilled wheat in the Pacific Northwest. Plant Disease 64:61-61.

de Jonge, R., Bolton, M. D., Kombrink, A., van den Berg, G. C. M., Yadeta, K. A., and Thomma, B. P. H. J. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Research 23:1271-1282.

de Jonge, R., Peter van Esse, H., Maruthachalam, K., Bolton, M. D., Santhanam, P., Saber, M. K., Zhang, Z., Usami, T., Lievens, B., Subbarao, K. V., and Thomma, B. P. H. J. 2012. Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. Proceedings of the National Academy of Sciences of the United States of America 109:5110-5115.

Dodds, P. N., Lawrence, G. J., Catanzariti, A.-M., Teh, T., Wang, C.-I. A., Ayliffe, M. A., Kobe, B., and Ellis, J. G. 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. Proceedings of the National Academy of Sciences of the United States of America 103:8888-8893.

Dong, S., Raffaele, S., and Kamoun, S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Current Opinion in Genetics & Development 35:57-65.

Dong, S., Stam, R., Cano, L. M., Song, J., Sklenar, J., Yoshida, K., Bozkurt, T. O., Oliva, R., Liu, Z., Tian, M., Win, J., Banfield, M. J., Jones, A. M. E., van der Hoorn, R. A. L., and Kamoun, S. 2014. Effector specialization in a lineage of the Irish potato famine pathogen. Science 343:552-555.

Duplessis, S., Cuomo, C. A., Lin, Y.-C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly, D. L., Hacquard, S., Amselem, J., Cantarel, B. L., Chiu, R., Coutinho, P. M., Feau, N., Field, M., Frey, P., Gelhaye, E., Goldberg, J., Grabherr, M. G., Kodira, C. D., Kohler, A., Kües, U., Lindquist, E. A., Lucas, S. M., Mago, R., Mauceli, E., Morin, E., Murat, C., Pangilinan, J. L., Park, R., Pearson, M., Quesneville, H., Rouhier, N., Sakthikumar, S., Salamov, A. A., Schmutz, J., Selles, B., Shapiro, H., Tanguay, P., Tuskan, G. A., Henrissat, B., Van de Peer, Y., Rouzé, P., Ellis, J. G., Dodds, P. N., Schein, J. E., Zhong, S., Hamelin, R. C., Grigoriev, I. V., Szabo, L. J., and Martin, F. 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proceedings of the National Academy of Sciences of the United States of America 108:9166-9171.

Faino, L., Seidl, M. F., Shi-Kunne, X., Pauper, M., van den Berg, G. C. M., Wittenberg, A. H. J., and Thomma, B. P. H. J. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome Research 26:1091-1100.

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484:186-194.

Fradin, E. F., and Thomma, B. P. 2006. Physiology and molecular aspects of *Verticillium* wilt diseases caused by *V. dahliae* and *V. albo-atrum*. Molecular Plant Pathology 7:71-86.

Fu, D., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X., Sela, H., Fahima, T., and Dubcovsky, J. 2009. A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. Science 323:1357-1360.

Gómez-Gómez, L., and Boller, T. 2000. FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. Molecular Cell 5:1003-1011.

Gout, L., Kuhn, M. L., Vincenot, L., Bernard-Samain, S., Cattolico, L., Barbetti, M., Moreno-Rico, O., Balesdent, M.-H., and Rouxel, T. 2007. Genome structure impacts molecular evolution at the AvrLm1 avirulence locus of the plant pathogen *Leptosphaeria maculans*. Environmental Microbiology 9:2978-2992.

Guyon, K., Balagué, C., Roby, D., and Raffaele, S. 2014. Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. BMC Genomics 15:336.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., Cakir, C., Carrington, J. C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M. A., Fugelstad, J., Gilroy, E. M., Gnerre, S., Green, P. J., Grenville-Briggs, L. J., Griffith, J., Grunwald, N. J., Horn, K., Horner, N. R., Hu, C. H., Huitema, E., Jeong, D. H., Jones, A. M., Jones, J. D., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzon, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., Savidor, A., Schornack, S., Schwartz, D. C., Schumann, U. D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D. J., Sykes, S., Thines, M., van de Vondervoort, P. J., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Hann, D. R., and Rathjen, J. P. 2007. Early events in the pathogenicity of *Pseudomonas syringae* on *Nicotiana benthamiana*. The Plant Journal 49:607-618.

Hartmann, F. E., Sánchez-Vallet, A., McDonald, B. A., and Croll, D. 2017. A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. ISME Journal 27:2991-1204.

Hemetsberger, C., Mueller, A. N., Matei, A., Herrberger, C., Hensel, G., Kumlehn, J., Mishra, B., Sharma, R., Thines, M., Hückelhoven, R., and Doehlemann, G. 2015. The fungal core effector Pep1 is conserved across smuts of dicots and monocots. New Phytologist 206:1116-1126.

Hurni, S., Brunner, S., Buchmann, G., Herren, G., Jordan, T., Krukowski, P., Wicker, T., Yahiaoui, N., Mago, R., and Keller, B. 2013. Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. Plant Journal 76:957-969.

Iida, Y., van 't Hof, P., Beenen, H., Mesarich, C., Kubota, M., Stergiopoulos, I., Mehrabi, R., Notsu, A., Fujiwara, K., Bahkali, A., Abd-Elsalam, K., Collémare, J., and De Wit, P. J. G. M. 2015. Novel mutations detected in avirulence genes overcoming tomato Cf resistance genes in Isolates of a Japanese population of *Cladosporium fulvum*. PLoS One 10:e0123271.

Janssen, B.-J., and Gardner, R. C. 1990. Localized transient expression of GUS in leaf discs following cocultivation with Agrobacterium. Plant Molecular Biology 14:61-72.

Jones, J. D. G., and Dangl, J. L. 2006. The plant immune system. Nature 444:323-329.

Joosten, M. H. A. J., Cozijnsen, T. J., and De Wit, P. J. G. M. 1994. Host resistance to a fungal tomato pathogen lost by a single base-pair change in an avirulence gene. Nature 367:384-386.

Kamper, J., Kahmann, R., Bolker, M., Ma, L.-J., Brefort, T., Saville, B. J., Banuett, F., Kronstad, J. W., Gold, S. E., Muller, O., Perlin, M. H., Wosten, H. A. B., de Vries, R., Ruiz-Herrera, J., Reynaga-Pena, C. G., Snetselaar, K., McCann, M., Perez-Martin, J., Feldbrugge, M., Basse, C. W., Steinberg, G., Ibeas, J. I., Holloman, W., Guzman, P., Farman, M., Stajich, J. E., Sentandreu, R., Gonzalez-Prieto, J. M., Kennell, J. C., Molina, L., Schirawski, J., Mendoza-Mendoza, A., Greilinger, D., Munch, K., Rossel, N., Scherer, M., Vranes, M., Ladendorf, O., Vincon, V., Fuchs, U., Sandrock, B., Meng, S., Ho, E. C. H., Cahill, M. J., Boyce, K. J., Klose, J., Klosterman, S. J., Deelstra, H. J., Ortiz-Castellanos, L., Li, W., Sanchez-Alonso, P., Schreier, P. H., Hauser-Hahn, I., Vaupel, M., Koopmann, E., Friedrich, G., Voss, H., Schluter, T., Margolis, J., Platt, D., Swimmer, C., Gnirke, A., Chen, F., Vysotskaia, V., Mannhaupt, G., Guldener, U., Munsterkotter, M., Haase, D., Oesterheld, M., Mewes, H.-W., Mauceli, E. W., DeCaprio, D., Wade, C. M., Butler, J., Young, S., Jaffe, D. B., Calvo, S., Nusbaum, C., Galagan, J., and Birren, B. W. 2006. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. Nature 444:97-101.

Kanzaki, H., Yoshida, K., Saitoh, H., Fujisaki, K., Hirabuchi, A., Alaux, L., Fournier, E., Tharreau, D., and Terauchi, R. 2012. Arms race co-evolution of *Magnaporthe oryzae AVR-Pik* and rice *Pik* genes driven by their physical interactions. Plant Journal 72:894-907.

Kema, G. H. J., Mirzadi Gohari, A., Aouini, L., Gibriel, H. A. Y., Ware, S. B., van den Bosch, F., Manning-Smith, R., Alonso-Chavez, V., Helps, J., Ben M'Barek, S., Mehrabi, R., Diaz-Trujillo, C., Zamani, E., Schouten, H. J., van der Lee, T. A. J., Waalwijk, C., de Waard, M. A., De Wit, P. J. G. M., Verstappen, E. C. P., Thomma, B. P. H. J., Meijer, H. J. G., and Seidl, M. F. 2018. Stress and sexual reproduction affect the dynamics of the wheat pathogen effector AvrStb6 and strobilurin resistance. Nature Genetics 23:678.

Kim, K.-T., Jeon, J., Choi, J., Cheong, K., Song, H., Choi, G., Kang, S., and Lee, Y.-H. 2016. Kingdom-wide analysis of fungal small secreted proteins (SSPs) reveals their potential role in host association. Frontiers in Plant Sciences 7:186.

Kirk, W. W., Abu-El Samen, F., Tumbalam, P., Wharton, P., Douches, D., Thill, C. A., and Thompson, A. 2009. Impact of different US genotypes of *Phytophthora infestans* on potato seed tuber rot and plant emergence in a range of cultivars and advanced breeding lines. Potato Research 52:121-140.

Kombrink, A., Rovenich, H., Shi-Kunne, X., Rojas-Padilla, E., van den Berg, G. C. M., Domazakis, E., de Jonge, R., Valkenburg, D.-J., Sánchez-Vallet, A., Seidl, M. F., and Thomma, B. P. H. J. 2016. *Verticillium dahliae* LysM effectors differentially contribute to virulence on plant hosts. Molecular Plant Pathology 61:103.

Lauge, R., Joosten, M. H. A. J., Haanstra, J. P. W., Goodwin, P. H., Lindhout, P., and De Wit, P. J. G. M. 1998. Successful search for a resistance gene in tomato targeted against a virulence factor of a fungal pathogen. Proceedings of the National Academy of Sciences of the United States of America 95:9014-9018.

Liao, J., Huang, H., Meusnier, I., Adreit, H., Ducasse, A., Bonnot, F., Pan, L., He, X., Kroj, T., Fournier, E., Tharreau, D., Gladieux, P., and Morel, J.-B. 2016. Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. eLife 5:e19377.

Ma, L.-J., van der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M.-J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., Houterman, P. M., Kang, S., Shim, W.-B., Woloshuk, C., Xie, X., Xu, J.-R., Antoniw, J., Baker, S. E., Bluhm, B. H., Breakspear, A., Brown, D. W., Butchko, R. A. E., Chapman, S., Coulson, R., Coutinho, P. M., Danchin, E. G. J., Diener, A., Gale, L. R., Gardiner, D. M., Goff, S., Hammond-Kosack, K. E., Hilburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C. D., Koehrsen, M., Kumar, L., Lee, Y.-H., Li, L., Manners, J. M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S.-Y., Proctor, R. H., Regev, A., Ruiz-Roldan, M. C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D. C., Turgeon, B. G., Wapinski, I., Yoder, O., Young, S., Zeng, Q., Zhou, S., Galagan, J., Cuomo, C. A., Kistler, H. C., and Rep, M. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367-373.

Marshall, R., Kombrink, A., Motteram, J., Loza-Reyes, E., Lucas, J., Hammond-Kosack, K. E., Thomma, B. P. H. J., and Rudd, J. J. 2011. 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 Physiology 156:756-769.

McDonald, M. C., McDonald, B. A., and Solomon, P. S. 2015. Recent advances in the *Zymoseptoria tritici*-wheat interaction: insights from pathogenomics. Frontiers in Plant Sciences 6:102.

Meile, L., Croll, D., Brunner, P. C., Plissonneau, C., Hartmann, F. E., McDonald, B. A., and Sánchez-Vallet, A. 2018. A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance to septoria tritici blotch. New Phytologist 65:512.

Mesarich, C. H., Griffiths, S. A., van der Burgt, A., Ökmen, B., Beenen, H. G., Etalo, D. W., Joosten, M. H. A. J., and de Wit, P. J. G. M. 2014. Transcriptome sequencing uncovers the *Avr5* avirulence gene of the tomato leaf mold pathogen *Cladosporium fulvum*. Molecular Plant-Microbe Interactions 27:846-857.

Michelmore, R. W., Christopoulou, M., and Caldwell, K. S. 2013. Impacts of resistance gene genetics, function, and evolution on a durable future. Annual Review of Phytopathology 51:291-319.

Michielse, C. B., and Rep, M. 2009. Pathogen profile update: *Fusarium oxysporum*. Molecular Plant Pathology 10:311-324.

Poppe, S., Dorsheimer, L., Happel, P., and Stukenbrock, E. H. 2015. Rapidly evolving genes are key players in host specialization and virulence of the fungal wheat pathogen *Zymoseptoria tritici* (*Mycosphaerella graminicola*). PLoS Pathogens 11:e1005055.

Raffaele, S., and Kamoun, S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nature Reviews Microbiology 10:417-430.

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

Rouxel, T., Grandaubert, J., Hane, J. K., Hoede, C., van de Wouw, A. P., Couloux, A., Dominguez, V., Anthouard, V., Bally, P., Bourras, S., Cozijnsen, A. J., Ciuffetti, L. M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin, S. B., Gout, L., Glaser, N., Linglin, J., Kema, G. H. J., Lapalu, N., Lawrence, C. B., May, K., Meyer, M., Ollivier, B., Poulain, J., Schoch, C. L., Simon, A., Spatafora, J. W., Stachowiak, A., Turgeon, B. G., Tyler, B. M., Vincent, D., Weissenbach, J., Amselem, J., Quesneville, H., Oliver, R. P., Wincker, P., Balesdent, M.-H., and Howlett, B. J. 2011. Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by repeat-Induced point mutations. Nature Communications 2:202.

Saintenac, C., Lee, W.-S., Cambon, F., Rudd, J. J., King, R. C., Marande, W., Powers, S. J., Bergès, H., Phillips, A. L., Uauy, C., Hammond-Kosack, K. E., Langin, T., and Kanyuka, K. 2018. Wheat receptor-kinase-like protein Stb6 controls gene-for-gene resistance to fungal pathogen *Zymoseptoria tritici*. Nature Genetics 51:291.

Santhanam, P., van Esse, H. P., Albert, I., Faino, L., Nürnberger, T., and Thomma, B. P. H. J. 2013. Evidence for functional diversification within a fungal NEP1-like protein family. Molecular Plant-Microbe Interactions 26:278-286.

Schirawski, J., Mannhaupt, G., Munch, K., Brefort, T., Schipper, K., Doehlemann, G., Di Stasio, M., Rossel, N., Mendoza-Mendoza, A., Pester, D., Muller, O., Winterberg, B., Meyer, E., Ghareeb, H., Wollenberg, T., Munsterkotter, M., Wong, P., Walter, M., Stukenbrock, E., Guldener, U., and Kahmann, R. 2010. Pathogenicity determinants in smut fungi revealed by genome comparison. Science 330:1546-1548.

Schmidt, S. M., Houterman, P. M., Schreiver, I., Ma, L., Amyotte, S., Chellappan, B., Boeren, S., Takken, F. L. W., and Rep, M. 2013. MITEs in the promoters of effector genes allow prediction of novel virulence genes in *Fusarium oxysporum*. BMC Genomics 14:1-21.

Schmidt, S. M., Lukasiewicz, J., Farrer, R., van Dam, P., Bertoldo, C., and Rep, M. C. 2016. Comparative genomics of *Fusarium oxysporum* f. sp. *melonis* reveals the secreted protein recognized by the Fom-2 resistance gene in melon. New Phytologist 209:307-318.

Schurch, S., Linde, C. C., Knogge, W., Jackson, L. F., and McDonald, B. A. 2004. Molecular population genetic analysis differentiates two virulence mechanisms of the fungal avirulence gene *NIP1*. Molecular Plant-Microbe Interactions 17:1114-1125.

Seidl, M. F., and Thomma, B. P. H. J. 2014. Sex or no sex: Evolutionary adaptation occurs regardless. BioEssays 36:335-345.

Seidl, M. F., and Thomma, B. P. H. J. 2017. Transposable elements direct the coevolution between plants and microbes. Trends in Genetics:1-10.

Soanes, D. M., Alam, I., Cornell, M., Wong, H. M., Hedeler, C., Paton, N. W., Rattray, M., Hubbard, S. J., Oliver, S. G., and Talbot, N. J. 2008. Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis. PLoS One 3:e2300.

Sperschneider, J., Ying, H., Dodds, P., Gardiner, D., Upadhyaya, N. M., Singh, K., Manners, J. M., and Taylor, J. 2014. Diversifying selection in the wheat stem rust fungus acts predominantly on pathogen-associated gene families and reveals candidate effectors. Frontiers in Plant Sciences 5.

Stergiopoulos, I., van den Burg, H. A., Okmen, B., Beenen, H. G., van Liere, S., Kema, G. H. J., and De Wit, P. J. G. M. 2010. Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. Proceedings of the National Academy of Sciences of the United States of America 107:7610-7615.

Stukenbrock, E. H., Bataillon, T., Dutheil, J. Y., Hansen, T. T., Li, R., Zala, M., McDonald, B. A., Wang, J., and Schierup, M. H. 2011. The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. Genome Research 21:2157-2166.

Stukenbrock, E. H., and McDonald, B. A. 2007. Geographical variation and positive diversifying selection in the host-specific toxin SnToxA. Molecular Plant Pathology 8:321-332.

van Dam, P., Fokkens, L., Schmidt, S. M., Linmans, J. H. J., Kistler, H. C., Ma, L.-J., and Rep, M. 2016. Effector profiles distinguish formae speciales of *Fusarium oxysporum*. Environmental Microbiology 18:4087-4102.

van Kan, J. A., van den Ackerveken, G. F., and de Wit, P. J. 1991. Cloning and characterization of cDNA of avirulence gene *avr9* of the fungal pathogen *Cladosporium fulvum*, causal agent of tomato leaf mold. Molecular Plant-Microbe Interactions 4:52-59.

Vlaardingerbroek, I., Beerens, B., Schmidt, S. M., Cornelissen, B. J. C., and Rep, M. 2016. Dispensable chromosomes in *Fusarium oxysporum* f. sp. *lycopersici*. Molecular Plant Pathology 17:1455-1466.

Vleeshouwers, V. G., and Oliver, R. P. 2014. Effectors as tools in disease resistance breeding against biotrophic, hemibiotrophic, and necrotrophic plant pathogens. Molecular Plant-Microbe Interactions 27:196-206.

Yin, J., Gu, B., Huang, G., Tian, Y., Quan, J., Lindqvist-Kreuze, H., and Shan, W. 2017. Conserved *RXLR* effector genes of *Phytophthora infestans* expressed at the early stage of potato infection are suppressive to host defense. Frontiers in Plant Sciences 8:2155.

Zhang, W., Chen, S., Abate, Z., Nirmala, J., Rouse, M. N., and Dubcovsky, J. 2017. Identification and characterization of *Sr13*, a tetraploid wheat gene that confers resistance to the Ug99 stem rust race group. Proceedings of the National Academy of Sciences of the United States of America 114:E9483-E9492.

Summary
Acknowledgements
About the author
List of publications
Education statement

SUMMARY

During host colonization, plant pathogens secrete molecules that enable host colonization, also known as effector proteins. Over the last decade, considerable attention in research on plant-microbe interactions has focused on understanding how pathogens adapt their effector gene repertoires to their host plants. **Chapter 1** is a literature study in which I describe the adaptation of the fungal pathogen *Zymoseptoria tritici* that is specialized on wheat plants, and compare that to the fungal pathogen *Verticillium dahliae* that is able to colonize a broad range of host plants. Additionally, I describe the main research questions and objectives of my PhD research, and provide an outline of the thesis.

Chapter 2 reviews how effector genes of plant pathogens can be discovered using genomics-based approaches. More specifically, I describe recent advances in genome sequencing technologies, genome assembly strategies, and gene annotations can reveal complete effector gene repertoires of plant pathogens. Moreover, I describe how the knowledge on effectors can be exploited to develop sustainable resistance breeding strategies.

Research on various pathogens has revealed that effector genes often localize within dynamic regions of their genomes, which is thought to facilitate accelerated evolution of effector (gene) catalogues. In **Chapter 3**, we determined effector genes localized within the core and lineage-specific (LS) genomic regions of a collection of *V. dahliae* strains, and determined the ability of these strains to infect a collection of plant species that belong to various plant families. We observed that *V. dahliae* strains that are able to infect the same host plant harbour highly divergent LS effector repertoires. Moreover, not even a single LS effector is shared by the strains that are able to infect the same host species. This has led me to suggest that different *V. dahliae* strains infect the same host plant by utilizing different effectors, suggesting redundancy among the activities of effector repertoires between lineages. Furthermore, I assessed the expression of core *V. dahliae* effector genes, and observed considerable differential expression between hosts. I conclude that the variability within LS effector gene repertoires as well as within the expression of core effector genes of *V. dahliae* strains may be instrumental for immunity evasion on particular host plants, allowing the various strains to be competitive in the co-evolution with their hosts.

V. dahliae strains are generally characterized by their ability to infect a broad host range, but individual strains display differential capacity to infect particular hosts. I reasoned that this differential host specificity is linked to the presence of host-specific effectors that contribute to disease establishment. In **Chapter 4**, I conducted comparative genomics of a sunflower pathogenic strain with a collection of non-pathogenic strains. Two effector genes were identified that exclusively occur in the sunflower pathogenic strain and that are highly induced during host colonization. These two identified candidate effector

genes are identical copies that likely evolved via a segmental duplication. Functional analysis revealed that this duplicated effector gene quantitatively contributes to *V. dahliae* virulence on sunflower, but not on other host plants.

Z. tritici strains exhibit host specificity, as they generally infect either the bread wheat (BW) or durum wheat (DW). In **Chapter 5**, I aimed to explore genome-wide differences underpinning host specificity in this fungal pathogen. Whole-genome sequencing of a worldwide collection of 136 *Z. tritici* isolates and phenotyping assays on a set of BW and DW cultivars were conducted, revealing extensive genetic diversity amongst *Z. tritici* isolates. Isolates originating from the Middle East, the centre of origin of wheat, showed an increased genetic diversity compared to isolates obtained from other geographical locations. Assessment of genome-wide differences between BW and DW isolates revealed four effector genes carrying non-synonymous single nucleotide polymorphisms in BW isolates. Presumably, extensive genetic diversity among *Z. tritici* isolates is indicative of its evolutionary potential to infect and overcome resistance in wheat plants, which has implications for the control of this fungal wheat disease.

In **Chapter 6**, I conducted comparative population genomics combined with genetic mapping in a collection of virulent and avirulent *Z. tritici* isolates on the bread wheat cultivar Shafir that carries the resistance gene *Stb6*. A polymorphic genomic region that determines avirulence was identified, leading to the identification of the avirulence effector gene *AvrStb6*. Recognition of this effector by Stb6 triggers host resistance. The paradigm states that deployment of Stb6 leads to removal of this avirulence effector gene from pathogen populations. However, when we crossed *Z. tritici* on wheat, we found that sex occurs even with an avirulent parent, thus leading to the maintenance of *AvrStb6* in subsequent populations. This observation suggests that the sexual behaviour of *Z. tritici* has major implications for disease control.

Chapter 7 is the general discussion of this PhD thesis and describes some of the approaches to identify effector genes, how pathogens adapt their effector gene repertoires to their host plants, and how the knowledge on pathogen effectors can be used establish durable and broad-spectrum resistance.

ACKNOWLEDGEMENTS

These are the last words in my PhD thesis, which I hope you enjoyed reading. To start, I would like to thank Allah for helping me to successfully finish this thesis. Without him I am nothing and hopefully he will accept this thesis research in the balance of my good deeds.

Surely, I have not established this thesis on my own, but rather with the support of many people including my supervisors, family, colleagues, and friends. To this end, I would like to thank my promotor Bart for his supervision. Without you, Bart, surely I couldn't have done it. I still remember when you guided me to start my PhD with my daily supervisor and co-promotor, Michael, whom I also appreciate very much for his daily support. Michael took me by the hand, and bit by bit he taught me how to carry out my research independently. Thank you, Michael, for that. Also, I would like to thank my other co-promotor Gert. I had very nice conversations with you, Gert, on science, religion, and culture. Definitely, I will miss these fruitful conversations.

I would like to thank my colleagues in the Laboratory of Phytopathology who helped me directly or indirectly to make my thesis into a success. Particularly, I would like to thank my colleagues in the bioinformatics office including Jasper, Sander, Claudio, and Xiaoqian for the nice and fruitful scientific discussions. Surely, in that 'hidden' and 'independent' office we managed to generate a very friendly, cosy, and nice work place and I miss working with all of you. This office also used to house Luigi for a couple of years, and of course I would like to thank him for everything. Additionally, I would like to thank Jinling for providing the collaboration on Chapters 3 and 4. I would also like to thank Guillaume, Anne, and Marc-Henri for the lovely discussions that we had over Skype on Chapter 5. I would like to thank Paul for his help on Chapter 5. I also would like to thank Amir and Lamia for providing me the data that I needed for the bioinformatics analysis conducted in Chapter 6 and I would like to thank Els for teaching me how to do the phenotyping on wheat leaves. I would like to thank Harold, Jan, Francine, Klaas, and Matthieu for their input in the project during the Friday seminars. Lastly, thank you Ali for your administrative help.

I miss my friends in Wageningen a lot. You, my friends, were always there whenever I needed you. Thank you so much Ali, Asmaa, Mohab, Ahmed, Muniroh, Daniel, Shahid, Fajar, Confidence, Kostas, Miguel, Anna, Elizabeth, Kiflom, Aisha, and Eyup. Without you, it would have not been as much fun to be in Wageningen.

To my lovely parents, may Allah reward you with goodness (jazakom Allah khairan). I could not have done this without your support. Every step that I make towards my success in this life definitely goes back to you. You raised me very well, you always cared about me, and you did everything to make me a successful person. I will always be very grateful to you.

ABOUT THE AUTHOR

Hesham A.Y. Gibriel was born in Cairo, Egypt on the 11th of August 1988. After completing his Bachelor studies (2005-2009) with a specialization in Plant Biotechnology at the Faculty of Agriculture, Cairo University, Egypt, he worked as a product manager at Alliance-Global Egypt. Subsequently, Hesham moved to Italy in 2011 to do an MSc in Plant Pathology with the specialization in Plant Virology. After obtaining his MSc degree in 2014, he started his PhD research in the Laboratory of Phytopathology of Wageningen University in the Netherlands under the supervision of Prof. Dr. Bart Thomma. He performed comparative genomics to study host adaptation of the filamentous ascomycete fungi *Zymoseptoria tritici* and *Verticillium dahliae*, which resulted this PhD thesis. Since 2018 Hesham is working as a researcher at University College Dublin, performing comparative genomics on *Z. tritici*.

LIST OF PUBLICATIONS

Gibriel, H. A. Y., Thomma, B. P. H. J., and Seidl, M. F. 2016. The age of effectors: genome-based discovery and applications. **Phytopathology** 106:1206-1212.

Kema, G. H. J.[#], Mirzadi Gohari, A.[#], Aouini, L.[#], **Gibriel, H. A. Y.**[#], Ware, S. B., van den Bosch, F., Manning-Smith, R., Alonso-Chavez, V., Helps, J., Ben M'Barek, S., Mehrabi, R., Diaz-Trujillo, C., Zamani, E., Schouten, H. J., van der Lee, T. A. J., Waalwijk, C., de Waard, M. A., De Wit, P. J. G. M., Verstappen, E. C. P., Thomma, B. P. H. J., Meijer, H. J. G., and Seidl, M. F. 2018. Stress and sexual reproduction affect the dynamics of the wheat pathogen effector AvrStb6 and strobilurin resistance. **Nature Genetics** 23:678.

Education Statement of the Graduate School

Experimental Plant Sciences
Issued to: Hesham Gibriel

The Graduate School EXPERIMENTAL PLANT SCIENCES

Date:	**20 February 2019**
Group:	**Laboratory of Phytopathology**
University:	**Wageningen University & Research**

1) Start-Up Phase		*date*
▶	**First presentation of your project**	
	Comparative population genomics combined with genetic mapping promotes effector discovery in the fungal wheat pathogen *Zymoseptoria tritici*	27-Nov-2015
▶	**Writing or rewriting a project proposal**	
	Comparative pathogenomics of ascomycete fungi	Jul-2015
▶	**Writing a review or book chapter**	
	The age of effectors: genome-based identification and future applications. Phytopathology 106:1206-1212. DOI: 10.1094/PHYTO-02-16-0110-FI	Mar-2016
▶	**MSc courses**	
	PHP-30306 Plant-Microbe Interactions	May-June-2015
	PHP-30806 Molecular Aspects of Bio-interactions	Jan-Feb-2016
	Subtotal Start-Up Phase	*15.0 credits**

2) Scientific Exposure		*date*
▶	**EPS PhD student days**	
	EPS PhD student days 'Get2Gether'	29-30-Jan-2015
	EPS PhD student days 'Get2Gether'	28-29-Jan-2016
▶	**EPS Theme Symposia**	
	EPS Theme 2 Symposium 'Interactions between Plants and Biotic Agents' + WCS Day	20-Feb-2015
	EPS Theme 2: Interactions between Plants and Biotic Agents	22-Jan-2016
	EPS Theme 2: Interactions between Plants and Biotic Agents	23-Jan-2017
	EPS Theme 4: Genome Biology + 'WURomics: Technology-Driven Innovation for Plant Breeding'	15-16-Dec-2016

►	**National meetings (e.g. Lunteren days) and other National Platforms**	
	Annual meeting 'Experimental Plant Sciences', Lunteren	13-14-Apr-2015
	BioSB meeting, Lunteren	4-5-Apr-2017
►	**Seminars (series), workshops and symposia**	
	DTL focus meeting: New sequencing technology: Nanopore and PacBio	08-Mar-2016
	Seminar: Dr. Theo van der Lee, Comparative genomics of plant pathogens: cryptic sex and the escape of some chromosomes to a genome defence system	15-Dec-2016
	Applications and Challenges of Oxford Nanopore Sequencing in the Life Science Industry	14-Apr-2016
	BiosB seminar: Bioinformatics, genomics and metagenomics applied to farm animal health and food security by Mick Watson	20-Apr-2016
	B-wise seminar: Marco Bink, Reconstruction of genome ancestry blocks in complex plant populations	02-Feb-2016
	B-Wise seminar: Sven Warris, Electronic annotation: where we are, where we want to be, and semantics to get there	05-Apr-2016
	B-wise seminar: Sander Rodenburg, Reconstruction of a metabolic network for the plant pathogen *Phytophthora infestans*	06-Dec-2016
	Seminar: Dr. Sophie Nadot, Perianth evolution in Ranunculaceae: are petals ancestral in the family?	20-May-2016
	Seminar: Dr. Martin Cann, The immune receptor Rx1 remodels chromatin and chromatin interactors in immunity	11-Jul-2017
	Seminar: Frantisek Marec, From sex chromosomes to sex determination in Lepidoptera	25-Oct-2017
	Rewriting our genes meeting	30-Sep-2016
►	**Seminar plus**	
►	**International symposia and congresses**	
	Zymoseptoria community meeting, Kiel, Germany	05-06-Sep-2017
	COST sustain, Banylus sur Mer, France	17-19-Feb-2016
►	**Presentations**	
	Cost sustain (talk)	17-Feb-2016
	EPS theme4 (talk)	16-Dec-2016
	BioSB (poster)	04-Apr-17
	B-Wise seminar (talk)	04-Nov-17

►	**IAB interview**	
►	**Excursions**	
	Subtotal Scientific Exposure	*10.9 credits**

3) In-Depth Studies		*date*
►	**EPS courses or other PhD courses**	
	Postgraduate course: Genome assembly	28-29-Apr-2015
	Next Generation Sequencing data analysis	29-31-Aug-2016
►	**Journal club**	
	Phytopathology literature discussion group	2014-2018
►	**Individual research training**	
	Subtotal In-Depth Studies	*4.5 credits**

4) Personal Development		*date*
►	**Skill training courses**	
	Scientific writing	Feb-Apr-2015
	PhD competence assessment	17-19-Feb-2015
	EPS Introductory Course	22-Sep-2015
	Career Orientation	4-25-Oct-2016
►	**Organisation of PhD student days, course or conference**	
►	**Membership of Board, Committee or PhD council of EPS**	
	Subtotal Personal Development	*3.8 credits**

TOTAL NUMBER OF CREDIT POINTS*	**34.2**

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

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

This work was carried out in the Laboratory of Phytopathology, Wageningen University, The Netherlands.

Financial support from Wageningen University for printing this thesis is greatly appreciated.

Cover & layout design: Loes Kema | gvo.nl
Printed by: GVO drukkers & vormgever, Ede (NL) | gvo.nl